

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université Mouloud Mammeri, Tizi-Ouzou



Faculté Génie Electrique et Informatique  
Département Automatique

**Actes de la**

**Troisième Conférence internationale sur la Vision Artificielle**

**CVA' 2015**

Tizi-Ouzou le 12, 13 et 14 avril 2015



La Sittelle kabyle (endémique)



Avec le parrainage de l'Agence  
Thématique de Recherche  
en Science et Technologie



Chihab Diffusion



Touring Voyage Recherche

Laboratoire Vision Artificielle et Automatique des Systèmes



### Présidents d'Honneur

**Pr. Naceur Eddine HANNACHI**, Recteur de l'Université de Mouloud  
Mammeri Tizi-Ouzou

**Pr. Mohamed-Saïd BELKAÏD**, Doyen de la faculté génie électrique et  
informatique de l'UMMTO

### Président de la Conférence

**Pr. Moussa DIAF**, FGEl, UMMTO

### Comité d'Organisation

Ourida Megherbi : Présidente

M.Tahar Bensidhoum

Sadia Alkama

Takfarinas Chelli

Fadhila Boudjemaa

Ouardia Chilali

Kherraz Khedidja

Ouali Yousfi

Ouiza Boukendour

Ouiza Hamdous

Merzouk Younsi

Samir Yesli

## Comité de programme

Pr. Patrick SIARRY (UPEC, France)  
Pr. Boudjemaa MANSOURI (CHU Bab-El-Oued, DZ)  
Pr. Mounir BENAÏSSA (ENIT, Tunisie)  
Pr. François CABESTAING (Univ. Lille1, France)  
Pr. Raymond CHIONG (Univ. Newcastle, Australie)  
Pr. Mourad FAKHFAKH (Univ. Sfax, Tunisie)  
Pr. Zahia GUESSOUM (Univ. Paris 6, France)  
Pr. Abdenour HADID (Univ. Oulu, Finlande)  
Pr. Amine NAIT-ALI (UPEC, France)  
Pr. Eric PETIT (UPEC, France)  
Pr. Mohamed SLIMANE (Univ. Tour, France)  
Pr. A.TALEB-AHMED (Univ. Valenciennes, France)  
Pr. Jean-Philippe URBAN (Univ. Mulhouse, France)  
Pr. Mhania GUERTI (ENP, Algérie)  
Pr. Abderrezak GUESSOUM (Univ. Blida, Algérie)  
Pr. Slimane LARABI (USTHB, Algérie)  
Pr. Malika MIMI (Univ. Mostaganem, Algérie)  
Pr. Abdelouahab MOUSSAOUI (Univ. Sétif, Algérie)  
Pr. Youcef SMARA (USTHB, Algérie)  
Pr. Hicham TEBBIKH (Univ. Guelma, Algérie)  
Pr. Rachid AHMED-OUAMAR (UMMTO, Algérie)  
Pr. Soltane AMEUR (UMMTO, Algérie)  
Pr. Nacereddine BENAMROUCHE (UMMTO, Algérie)  
Pr. Arezki BENFDILA (UMMTO, Algérie)  
Pr. Moussa DIAF (UMMTO, Algérie)  
Pr. Saïd DJENNOUNE (UMMTO, Algérie)  
Pr. Kamal Hammouche (UMMTO, Algérie)  
Pr. Mourad LAGHROUCHE (UMMTO, Algérie)  
Pr. Zohra MAZOUZI (UMMTO, Algérie)  
Pr. Rezki ZIANI (UMMTO, Algérie)

# SOMMAIRE

Abdelaziz Mokrani, Roland Baddeley and Mokhtar Nibouche <b>Estimation of moving objects from a sequence of images</b>	01
<i>El Hannachi Sid Ahmed and Benamrane Nacéra</i> <b>Compression des Images Médicales Fixes par RDCT et Codeurs Imbriqués</b>	06
<i>Hameurlaine Messaoud, Moussaoui Abdelouahab and Cherroun Hadda.</i> <b>Partial volume correction in brain MRI segmentation.</b>	12
Azeddine Benlamoudi, Abdelkrim Ouafi, Djamel Samai, Abdelmalik, Taleb-Ahmed, Abdenour Hadid and Salah Eddine Bekhouche <b>Face Spoofing Detection From Single Images Using Active Shape Models with Stasm And LBP</b>	18
Asma Manel Kherchi and Slimane Larabi <b>Head Pose Estimation from Depth Map.</b>	23
Benouis Mohamed, Senouci Mohamed and Benkkadour Mohamed Kamel Eldine <b>A Novel technique for Human face recognition using WPD &amp; LBP &amp; Bi-dimensional Subspace Analysis</b>	28
Abdelhak Bouhali and Daoud Berkani <b>Combinaison du Filtrage Linéaire et le Seuillage Adaptatif d'Ondelettes pour le Dé-bruitage d'Images Satellitaires</b>	34
Gacem Amel, Berrached Nasr Eddine and Merad Boudia Sami <b>Urban Roads Extraction from Very High Resolution Satellite Images</b>	42
Malika Bendouda and Berrached Nasreddine <b>Multi-scale approach for urban road extraction using remote sensing image</b>	47
Mahi Habib and Kaouadji Mounia <b>Méthodologie de mise en place de règles de connaissance pour la classification des objets sur des images à THRS</b>	51
Soraya Cheriguene and Nabiha Azizi <b>Constructing ensembles of classifiers using Random Subspace and Static Classifiers Selection Paradigms.</b>	56
Djeddou Ferhat and Hammoudi Abderazek <b>Discrete Composite Differential Evolution Algorithm Practical Engineering Problems Optimization</b>	61
Kaouter Labeled and Hadria Fizazi <b>Utilisation de l'Algorithme de Recherche Coucou pour le Regroupement des Données de Télédétection</b>	65

Rabhi Seddik and Fizazi Hadria	73
<b>Etude comparative entre l'algorithme de Firefly et les comportements collectifs des fourmis dans la Classification des Images satellitaires</b>	
Ghania Khensous, Messabih Belhadri, Chouarfia Abdallah and Bernard Maigret	80
<b>Comparison of Metaheuristic Algorithms for Unconstrained Global Optimization</b>	
Lamiche Chaabane and Moussaoui Abdelouahab	85
<b>An Hybrid Method to Improve Segmentation of Brain MR Images</b>	
Belhadj Sabrina, Benali Abdelali, Ahmed-Foitih Zoubir and Taleb-Ahmed Abdelmalik	90
<b>EEG-fMRI fusion approaches and its applications.</b>	
Masmoudi Mostefa, L. Kaddour El Boudadi, Khelf Mohamed, Dahane Amine and A Loukil	94
<b>L'utilisation de la réalité virtuelle pour la telerobotique</b>	
Oulad Kaddour Mohamed and Benatchba Karima	99
<b>SVM based gender classification</b>	
Leila Belkhdja, Djamila Hamdadou, Abdelhamid Keddi and Smail Fadhel	103
<b>Application à la localisation des zones surchauffées dans les installations électriques par thermographie infrarouge</b>	
Dahia Issaad, A. Medjahed, S. Amara, M. Bouafia, N. Djaker, M. Lamy De La Chapelle	107
<b>La spectroscopie de corrélation de fluorescence pour l'analyse ultra-sensibles des systemes biologiques</b>	
Kahina Slimani, Soltane Ameer and Zohra Ameer	111
<b>Extraction des joints de grains par une approche multi-échelles</b>	
Mehdi Cherifi, Lahdir Mourad and Soltane Ameer	115
<b>L'efficacité de la transformée de Radon vis-à-vis du bruit de quantification : Etude comparative</b>	
Bendani Djazia, Chalabi Zidelkhir and Berrached Nasreddine	122
<b>Système de Diagnostic de l'Infarctus du Myocarde par le Réseau de Neurones LVQ et la Carte SOM-LVQ</b>	
Mohamed Brahim and Moussaoui Abdelouahab	128
<b>MACHINE LEARNING FOR MALWARE DETECTION USING API CALLS</b>	
Ouiza Boukendour and Fatiha Huidene	134
<b>Identification Automatique d'Objets Microscopiques de Formes Complexes. Application aux Grains de Pollen</b>	
Mazari Chahrazad, Boussenou Tassadit et Diaf Moussa	139
<b>Réseaux Bayésiens appliqués à une fusion de données multi-capteurs en vue d'un contrôle intelligent d'un microclimat</b>	

Boussenou Tassadit, Mazari Chahrazad et Diaf Moussa	145
<b>Optimisation de paramètres dans une serre agricole par fusion de données multi capteurs</b>	
Youness CHAWKI, Mohammed OUANAN et Brahim AKSASSE	149
<b>Estimation des fréquences 3-D par la méthode ESPRIT 3-D dans un bruit Gaussien Coloré</b>	

# Estimation of moving objects from a sequence of images

Abdelaziz Mokrani<sup>#</sup>, Roland Baddeley<sup>\*</sup>, Mokhtar Nibouche<sup>\*</sup>

<sup>#</sup>*Institute of Electrical and Electronic Engineering, Algeria.*

*Mokaziz3@yahoo.fr*

<sup>\*</sup>*Experimental Psychology, University of Bristol, UK.*

*Roland.Baddeley@bristol.ac.uk*

<sup>\*</sup>*FET- Engineering, Design and Mathematics, University of the West of England, UK.*

*Mokhtar.Nibouche@uwe.ac.uk*

**Abstract**— Actually, object tracking is a challenging machine vision problem. The proposed paper is devoted to probe the particle filter model proposed by Fearnhead [5] for tracking. In other words, we investigate a tracking method to estimate moving objects in videos.

**Keywords**— tracking, Bayesian inference, resampling, Dirichlet process, particle filter, images.

## I. INTRODUCTION

Machine vision and machine learning are two fields of artificial intelligence. They're very useful to control for example nanorobots. The successful implementation of these types of robots has still difficulties. However, one of the objectives of such robots is to track targeted object. In this case, the role of the nanorobot is to capture images in different possible places. Once the image is captured, then processed, the programmed nanorobot must be able to track object, for example, to deliver drugs. This research project aims to investigate the particle filter proposed by Fearnhead [5] to track objects. Our paper regroups four sections. The opening one discusses briefly some of the more important works in tracking field and different approaches used. This section shows us our motivation to apply particle filter method in robotics. Section 2 introduces essential concepts in probabilistic modelling regarding to our research project. The formal statistical concepts, the notion of sampling are described and the algorithm is discussed as well. Also, the flowchart of particle filter is presented. Section 3 describes our application of the particle filter. It discusses the results obtained during the implementation process. The analysis focuses on the necessary parameters to get good results. We finish by a general conclusion of our work and possible works which permit to improve our system in the future.

## II. TRACKING INVESTIGATED OVER MANY YEARS

The estimation is a crucial issue in many areas: signal processing, automatic, machine vision...and so on. In the field of computer vision, many applications require the detection and tracking of moving objects in a sequence of images.

Traditionally, the problem can be formulated by Bayesian inference methods. But the type of Kalman filter and particle filter provide beneficial solutions.

Let's firstly show the history of the developments in computer vision by using different approaches to track objects in a sequences of images. Jojic et al [11] proposed a technique for automatically learning layers of "flexible sprites". Bowden et al [2] and Sami et al [18] based on non-linear statistical models to interpret pose and motion from 2D images. Gilles et al [8] used learning switching dynamic models. Qian et al [15] proposed a very interesting model based on Bayesian algorithms for tracking motion of moving objects. We remark that these works have relied on either mixture or switching models.

In urban environment, we can mention the two works: Schönherr et al [19] based on object edges in their research work. They used edges detection because generally they are less sensitive to changes compared to others features.

Barth et al [1] estimated in their work the pose and motion parameters. Also, they presented a real-time multi-filter approach.

The applications of statistical methods are largely used in biology. However, the science, biostatistics, bases on experiments to collect, analyse data and then interpret the results. Statistical test can be applied in biology as well when the results are not clear. Recently, Ramanathan et al [16] presented algorithms to track the behaviour of the cells and they segmented the cells from images then tracked them.

In all, solving tracking problems which are relevant in different fields differ in type of background models.

## III. PROBABILISTIC MODELLING

Probability theory is a mathematical branch to express data having lots of uncertainty, noise... and so forth. In other words, probabilistic models are useful to express uncertainties. For more details, the book of Christopher [4] is a very good reference. We use them in order to visualise models in high level for complex problems. Another reason is that to derive desirable solutions. In more general terms, there are many considerable techniques.

### A. Bayesian inference

Bayesian inference permitting to process incomplete information formally or more precisely it is useful to process unknown number of components. To illustrate this approach, we shall introduce some notation. Bibliographically, today there are several books which have been written on Bayesian inference, Gudmund [10] and Garthwaite et al [7]. They present excellent introductions to the theoretical Bayesian reference. Let's first illustrate how this method works. The idea in this approach is to suppose that parameters have a priori distribution. In all, there are several useful models:

- Single-parameter models, for example the Binomial and Poisson. In such models, the parameter is one dimensional.
- Multiparameter models, for example multinomial model. However, the distribution can be either continuous or discrete. For example, Poisson, Binomial and multinomial distributions are discrete. On the other hand, Student-t and Dirichlet distributions are continuous.

Hierarchically, a prior distribution can be constructed basing on historical experiments. In other words, from observed parameters we can estimate prior distribution basing on the Bayes' rule:  $p(\theta | Y) \propto p(Y | \theta)p(\theta)$ , where:

$p(\theta)$  represents the priori distribution and  $p(Y | \theta)$  represents the data distribution.

$p(\theta | Y)$  represents the posteriori density.

For a fixed data  $Y$ . The function of the parameter  $\theta$  which represents the quantity of interest  $p(Y | \theta)$ , is called likelihood function.

Before going any further, in statistical inference, hypothesis testing is useful to make decisions. The next figure shows and summarises the relation between unobserved states and measurements.

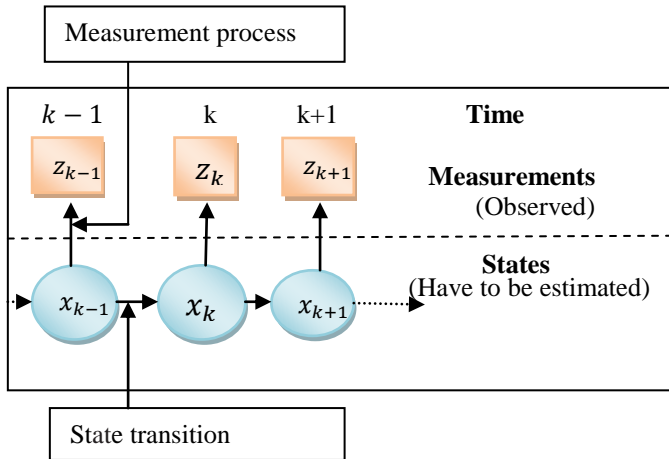
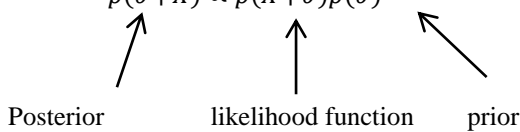


Fig. 1 State transition and Measurement process steps, Matthias [13]

If  $\theta$  represents a quantity of interest and  $X$  represents a random sample which describes all things we know about  $\theta$ , then, we describe the uncertainty of  $\theta$  by:

$$p(\theta | X) \propto p(X | \theta)p(\theta)$$



We assume  $y$  depends on  $X$  only through  $\theta$ . The prediction is:  $p(y | X) = \int p(y | \theta)p(\theta | X)d\theta$ .

In all, in Bayesian inference, the uncertainties have a joint posterior distribution. However, the condition is on observed data.

### B. Nonparametric Bayesian model

Basically, a nonparametric model can be defined as a not parametric model. This last is just a model using certain defined parameters. For instance, Dirichlet process is a nonparametric model for unknown parameters.

### C. Dirichlet process

Dirichlet process is a class of nonparametric Bayesian methods. Ferguson [6] defined Dirichlet process as a probability measure on the space of all probability measures. However, he defined the two desirable proprieties for these measures:

- Large support of a priori distribution;
- A posterior distribution conditionally to the samplers observed should be "analytically manageable".

### D. Particle filters

Density estimation is a statistical procedure for estimating the probability distribution function basing on observed data. Many approaches have been suggested. For example Martin et al [12] proposed the nonparametric approach kernel density estimation. However, density algorithms permit to estimate posterior density and by implementing Bayesian recursive equations to estimate unknown density, then we talk about particle filter. Over the past decade, the number of papers within the field of machine vision using statistical approaches such particle filter has grown significantly. Particle filter is a statistical estimation technique which has been widely used. There are many papers and books written on the topic of particle filtering since it was first coined in the early 1990s. It was introduced by Gordon et al [9]. Briefly, the particle filter algorithm consists of three steps: particles computation, weight computation, and resampling algorithm.

Resampling is the mathematical process of selecting individuals from the original samples. In stratified random sampling, the probability distribution function  $h(x)$  is given by the formula:

$$h(x) = \sum_{i=1}^N \beta_i h_i(x).$$

Where:

$N$  represents strata, and  $B$  is the standard beta distribution.

$$\sum_{i=1}^N \beta_i = 1.$$

An important characteristic of a particle is its weight. In resampling, the idea is to propagate a subset of particles. Our nonparametric method of statistical inference permits to update the particles having largest weight. The idea of Fearnhead-clifford resampling algorithm is to leave a chance to some low-weight Gaussians to survive. The next figure shows graphically the resampling idea.



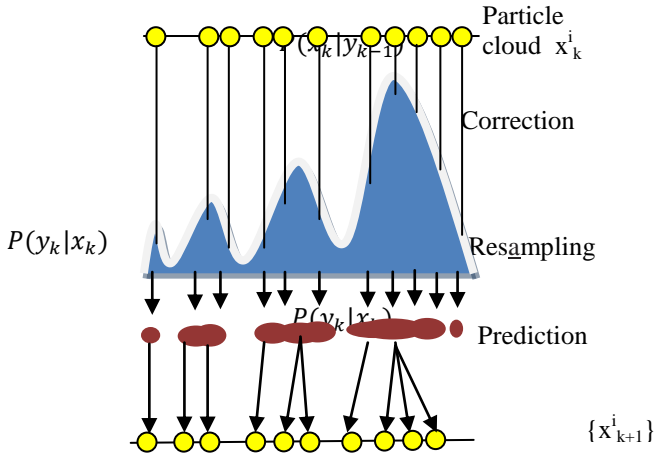


Fig. 2 Particle filter with sampling and resampling, Ramazan et al [17]

To summarise, the next figure illustrate the particle filter flowchart.

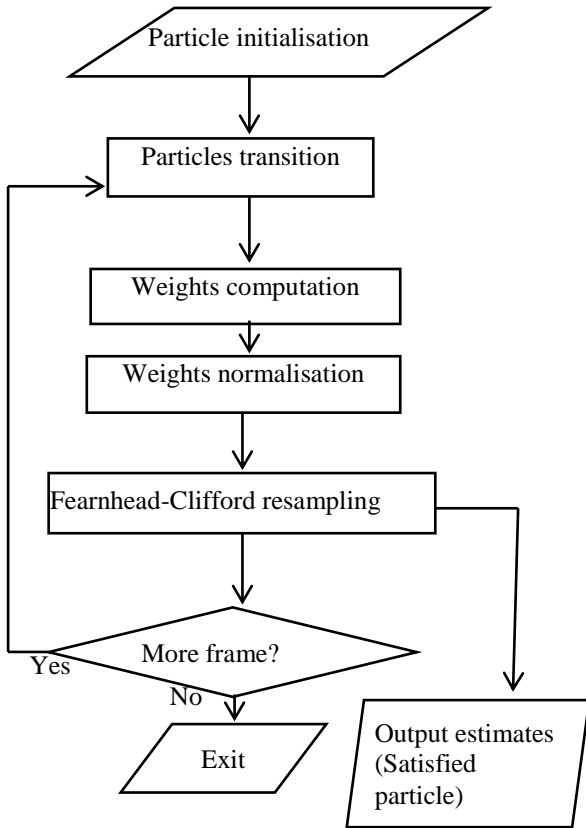


Fig. 3 Our particle filter flowchart

Many models of particle filter have been proposed. Basing on Fearnhead-Clifford resampling Algorithm we define our useful algorithm in the next figure.

1. First step, Initialisation:

For  $i=1$  to  $N$

- Sample  $X_0^{(i)} \sim \pi_0(X_0)$

- And set the weight  $w_0^{(i)} = \frac{p(X_0^{(i)})}{\pi_0(X_0^{(i)})}$

Calculate  $w_0 = \sum_{i=0}^N w_0^{(i)}$ .

2. Second step, iteration:

- For  $t=1, 2, \dots, N$  do
- For  $i=1, \dots, N$  do

- ❖ Generate  $X_t^{(i)}$  according to  $q_t(X_t | \tilde{x}_{t-1}^{(i)}, Z_t)$

- ❖ Set  $w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(Z_t | X_t^{(i)}) p(X_t^{(i)} | X_{t-1}^{(i)})}{q_t(X_t^{(i)} | X_{t-1}^{(i)}, Z_t)}$  with  $\sum_{i=1}^N w_t^{(i)} = 1$

3. Third step, resampling:

- After getting the normalised weights, we calculate

$$c: M = \sum_{i=1}^N \min(cw_t^{(i)}, 1).$$

- For  $i=1$  to  $N$

- ❖ If  $w_t^{(i)} > \frac{1}{c}$ , we keep  $i^{(th)}$  particle with its weight to form a set of  $A$  particles.
- ❖ Else, we keep  $i^{(th)}$  particle with its weight to form a set of  $(M - A)$  particles.
- ❖ Use the resampling algorithm of Carpenter et al [3] to resample the rest particles in the set  $M - A$ , and each particle of this last assigned a weight  $\frac{1}{c}$ .

Fig.4 Our particle filter algorithm

#### E. Our research work

Particle filter is a statistical approach. However, in our work we investigate particle filter of Fearnhead [5] regarding to its performance. But, it requires high performance of hardware (cameras, sensors...). For this main reason, the free VHumanEva dataset Moshe et al [14] is used. In other words, in our application we exploit a storage database. A dataset is a set of data where we can retrieve that data. In machine vision, the data are a collection of images and videos. Actually, several datasets are available and each has particular format and they are generally captured for a very specific purpose. In all, more than twenty datasets are available and they are free to use them and to test a proposed approach. We limit here to mention just the useful one for object tracking for implementing our algorithm. VHumanEva is a database of a variety of datasets including human poses and motions estimation. It contains approximately 80.000 frames and 56 videos combining four subjects. However, it contains six motions: walking, jogging, throwing/catching, gesture and boxing. Most videos are synchronised with 3D and they are captured by 1.3 M-pixel cameras.

But, in real time, the robot captures images, and then it is able to track moving objects. The next figure shows our simple system.

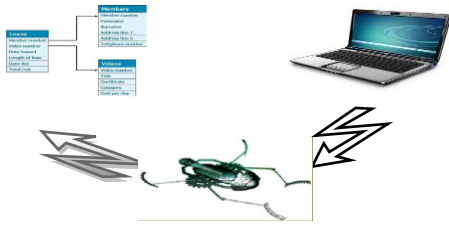


Fig. 5 Our system overview

#### IV. CONCLUSION

The project goal was to track any object for a given video in VHumanEva dataset. In other words, it was to make a position of a moving object in each frame. However, the study looked at different recognition algorithms with a particle filter. The different algorithms studied all showed the benefits and disadvantages. Some algorithms are easy to use and to understand but they have some limitations. The algorithm of colour recognition was the first important algorithm. This algorithm is simple and convenient but limited because the objects can lose their depth. The histogram algorithm which permitted us to compare two histograms was the most difficult to implement and the results are not very conclusive, because it was not easy to find a suitable description of a video. The latter algorithm covers the resampling of the particles with some improvements to increase the performance of our application to propagate the particles. It has the same constraint as the method of recognition of the camera movement. In all each algorithm studied is more effective than others in specific cases, out of their simplicity or accuracy.

Our algorithm was tested on numerous videos. However, we performed a series of experiments using the Weizmann Dataset. We define the motion by comparing two frames, between the reference histogram and the histogram of each particle basing on the HSV model. This difference then represents the weight of this last. In Markovian process the transition of probabilities between two histograms is given by the next formula:

$$\exp(-20d(H_0, H_i))$$

However, the objective is to compare our reference histogram,  $H_0$ , with the  $i^{th}$  histogram,  $H_i$ , of the  $i^{th}$  particle.

The number of particles plays an important role in our application. That is to say, by changing it we may get different results, and let's say, it may have influence on our program. In the next figure we show the results obtained on two different videos when the number of particles is 200.

In the future, it will be very interesting to investigate our project in the real environment. The program, running at the moment, can also be improved. Firstly, by improving its speed at execution, simplifying, and improving the different techniques. Secondly, by facilitating its use, there are a large number of variables that may play an important role to recognise moving objects.



Fig. 6 Noticeable results with number of particles 200

#### ACKNOWLEDGMENT

I would like to take this opportunity to express acknowledgements to my parents to complete this work. Many thanks also to Dr Roland Baddeley and Dr Nibouche Moukhtar for helpful discussions, supports, and invaluable advices.

In order to evaluate our application with different conditions, we calculate the error of the target region. However, we show in percentage the error between the reference regions and target regions. The next figure presents the error of our algorithm. In other words, it shows the accuracy and the reproducibility of our system which are, in average, 17 and 3 respectively.

#### REFERENCES

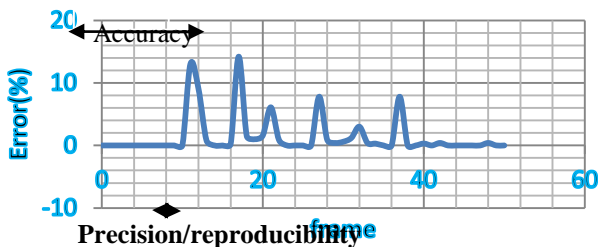


Fig. 7 The error at tracking regions for each frame.

- [1] A. Barth A. and U. Franke, *Tracking oncoming and turning vehicles at intersections*. 13th International IEEE Conference on Intelligent Transportation Systems, pp 861 – 868, Sept 2010.
- [2] R. Bowden, T.A. Mitchell, and M. Sarhadi, *Cluster based non-linear principle component analysis*. Electronic letters, vol33, no22, pp1858 – 1859, 1997.
- [3] J. Carpenter, P. Clifford, and P. Fearnhead, *An improved particle filter for non-linear problems*. IEEE proceedings-Radar, Sonar and Navigation, vol 46, no1, pp2–7, 1999.
- [4] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer. 2006.
- [5] P. Fearnhead, *Particle Filters for Mixture Models with an Unknown Number of Components*. Kluwer Academic Publishers, Statistics and Computing , vol 4, no1, pp.11-21, 2004.
- [6] T. S. Ferguson, *A Bayesian analysis of some nonparametric problems*. The Annals of Statistics, vol 1, pp 209-230, 1973.
- [7] D. Gamerman, *Markov chain Monte Carlo : stochastic simulation for Bayesian inference*. London : Chapman & Hall, 1997.
- [8] C. Gilles, N. Jacinto, and M. Jorge, *Learning switching dynamic models for objects tracking*. Pattern Recognition, vol37, pp 1841 – 1853, 2003.
- [9] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*. IEEE Proceedings on Radar and Signal Processing, vol140, no2, pp 107 – 113, 1993.

- [10] R.I. Gudmund, *Bayesian statistical inference*. Quantitative Applications in the Social Sciences, 1984.
- [11] N. Jovic and B.J. Frey, *Learning flexible sprites in video layers*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, pp 199-206, 2001.
- [12] E. B. Martin, and A.J. Morris, *Non-parametric confidence bounds for process performance monitoring charts*. Journal of Process Control, vol 6, no 6, pp 349–358, 1996.
- [13] M. Matthias, *Particle Filters, a tutorial*. Institut fur Angewandte Physik, J.W.Goethe-Universitat Frankfurt. 2003.
- [14] B. Moshe, G. Lena, S. Eli, I. Michal, and B. Ronen, *Actions as Space-Time Shapes*. Tenth IEEE International Conference on Computer Vision (ICCV), vol2, pp 1395 – 1402, 2005.
- [15] G. Qian, R. Chellappa, and Q. Zheng, *Bayesian Algorithms for Simultaneous Structure from Motion Estimation of Multiple Independently Moving Objects*. IEEE Transactions on Image processing, vol14, no1, pp 94-109, 2005.
- [16] K. Ramanathan, L.L.S. Ong, H.H. Asada, and M.H. Ang, *Automated tracking of biological cells in an “in-vitro” environment using active contours and distance measures*. IEEE 5th International Conference on Cybernetics and Intelligent Systems. pp 241 – 246, 2011.
- [17] H. Ramazan, A.N. Mohammad, and T. Mohammad, *A Multi Swarm Particle Filter for Mobile Robot Localization*. International Journal of Computer Science Issues, vol 7, Issue 3, no 2, 2010.
- [18] R. Sami, P. Alexandra, and G. Shaogang, *Multi-View Non-linear Active Shape Model using Kernel PCA*. Proceedings of the tenth British Machine Vision Conference, 1999.
- [19] K. Schönherr, B. Giesler, S. Wahju, and A. Knoll, *Robust tracking and monocular 3D reconstruction of line-based features in urban environments*. 13th International IEEE Conference on Intelligent Transportation Systems, pp1685-1690, 2010.

# Compression des Images Médicales Fixes par RDCT et Codeurs Imbriqués

El Hannachi Sid Ahmed Benamrane Nacéra

*Département d'informatique*

*Faculté des mathématiques et informatique*

*Université des Sciences et de la Technologie d'Oran, Mohamed Boudiaf*

*Oran, ALGERIE*

sidahmed.elhannachi@univ-usto.dz

nacera.benamrane@univ-usto.dz

**Résumé—** L'objectif de base de la compression des images médicales est de réduire le débit binaire et d'améliorer l'efficacité de compression pour la transmission rapide et le stockage de ces images, tout en maintenant une qualité d'image approuvable pour un diagnostic ou une thérapie. Ce papier propose une technique efficace destinée à appliquer une compression adaptative des images médicales. Il s'agit d'une application d'une Transformée Réversible en Cosinus Discrète (RDCT) avec un codeur EZW modifié, conçus à compresser sans perte la partie Région d'Intérêt (ROI) à des taux de compression élevés. La partie hors région d'intérêt est codée par l'algorithme SPIHT. Les résultats expérimentaux montrent l'efficacité de l'approche proposée et sa performance en termes de taux de compression et qualité de reconstruction en conservant le contenu de la partie ROI.

**Mots clés—** compression, images médicales, ROI, RDCT, EZW, SPIHT.

## I. INTRODUCTION

Les techniques d'imagerie médicale sont plus en plus sophistiquées, rapides et précises et plus l'imagerie est précise, plus les données acquises sont volumineuses et complexes, il est donc indispensable de disposer d'outils performants pour la transmission et le stockage de ces énormes quantités d'information. Parallèlement, la compression d'images numériques a connu une évolution incessante à celle des techniques d'imageries. Elle permet de réduire la taille d'une image dans le but d'augmenter la capacité des supports de stockages.

En particulier, dans le domaine médical l'image se caractérise par son utilisation en traitements de diagnostics. Cela explique l'importance toute particulière de la compression des images médicales.

Actuellement, la compression des images médicales est effectuée sans ou presque sans perte d'informations, pour assurer l'intégrité des données et éviter les erreurs de diagnostic dues à la dégradation de la qualité de l'image après décompression notamment dans les parties diagnostics appelées régions d'intérêt. Cependant, ce type de compression (sans perte) n'offre pas de réduction significative du volume

de ces données. Dans ce contexte, la compression avec pertes en imagerie médicale présente également de nombreux avantages pour d'autres applications que le diagnostic.

Afin de trouver un bon compromis entre le taux de compression et sur la validation de la qualité diagnostique d'images compressées, il paraît d'avantage de combiner les deux techniques de la compression sur une image médicale après avoir déterminé la région d'intérêt (Region Of Interest : ROI). Ainsi, après une étape d'extraction de la ROI dans une image médicale (Zone tumorale dans une image IRM du cerveau par exemple), une compression sans perte sera appliquée sur cette région d'intérêt tandis que la partie restante appelée arrière plan sera compressée avec perte.

JPEG2000 est la seule norme de compression d'images à couvrir avec un seul système l'ensemble des besoins de compression d'images médicales. Cette méthode JPEG2000 qui effectue une compression sans perte dans les régions d'intérêts et une compression avec perte ailleurs, pourrait être la solution pour fournir un codage d'image efficace et précis pour le domaine médical. La gestion de ROI se fait par la technique du Maxshift, implantée entre la quantification et le codage entropique. Elle consiste à appliquer un décalage binaire vers la droite aux coefficients associés au fond de l'image (Arrière plan), les coefficients de ROI étant codés sur les bits de poids fort [1].

Dans une image médicale une zone d'intérêt peut être détectée manuellement ou dans d'une manière semi-automatique ou bien automatique.

En général, les régions d'intérêt dans une image médicale prennent des formes aléatoires. Dans le cas d'une extraction manuelle, l'opérateur doit sélectionner des régions rectangulaires ou elliptiques qui doivent contenir les régions d'intérêt. En cas d'une extraction automatique on fait appelle à une méthode de segmentation ou classification. La détection de la ROI a été étudiée pendant plusieurs années et les algorithmes existants incluent des extractions de la ROI basées sur les attributs de bas niveau de l'image (Contours, critère d'homogénéité spatiale) [2], ou l'exigence d'une spécification manuelle de la forme de la ROI (Autres propriétés géométriques) par d'éventuels utilisateurs [3].

Plusieurs travaux existants dans la littérature ont discuté l'utilisation de la ROI pour améliorer la compression d'image médicale. Dans [4] un algorithme de compression d'image médicale Multi-ROIs avec conservation des contours est proposé. Se basant sur l'opérateur de segmentation de Canny, l'algorithme extrait l'information contour de l'image, et alors divise l'image en régions d'intérêt (ROIs) et non-ROIs. Au niveau des ROIs, une compression sans perte basée JPEG2000. Le codeur SPIHT a été utilisé pour le codage de la partie non-ROIs.

Dans [5], une approche basée-contexte régions d'intérêt pour compresser des images médicales en particulier images vasculaires, a été implémenté. L'image vasculaire est divisée en régions d'intérêt prioritaires et région secondaires. Ces régions ont été détectées automatiquement par l'algorithme de croissance de régions. Basée sur l'algorithme JPEG2000, les PROI ont été codées sans perte et les SROI avec perte. Un travail proposé dans [6] présente une méthode basée ROI avec l'algorithme SPIHT, la méthode utilise plusieurs régions en définissant les ROI par un masque appliqué sur la sous-bande basse fréquence qui permet de réduire la quantité de bits nécessaires pour le masque. Dans [7] une méthode de compression est présentée pour les images tomodensitométriques du foie (CT). L'approche comprend une compression ROI (sélectionnée manuellement) en utilisant un codeur sans perte de Huffman tandis que le reste de l'image (arrière-plan) est compressé par plusieurs techniques presque sans perte soient: SPIHT, EZW et DCT-RLE.

Ainsi, le but de toutes méthodes de compression des images médicales est d'améliorer le taux de compression et le PSNR tout en conservant l'information relative au diagnostic, contenue dans la région d'intérêt.

Dans ce contexte, ce papier propose une approche de compression sélective d'image médicale. La région d'intérêt sélectionnée subit une compression sans perte en appliquant une transformée DCT réversible RDCT. Réarrangés en sous-bandes, les coefficients de la RDCT sont codés par un l'algorithme EZW modifié en mode sans perte. L'arrière plan, qui ne nécessite pas d'être codé en pleine résolution, est compressé à l'aide de la technique SPIHT.

La section qui suit, décrit brièvement les techniques de base utilisées dans l'approche. Dans la section3, l'approche proposée est étalée. Des résultats expérimentaux et des discussions sont présentés dans la section4. Une conclusion ainsi que quelques perspectives sont données dans la dernière section

## II. MÉTHODES DE BASE UTILISÉES

### A. Transformée en Cosinus Discrete DCT

Le passage par la DCT a été l'idée majeure pour la compression JPEG [8]. Ce processus appartient à une classe d'opérations mathématiques, tout comme la transformée de Fourier. L'application de la DCT, ou d'une Transformée de Fourier fait passer l'information de l'image du domaine spatial en une représentation identique dans le domaine fréquentiel. Pourquoi ce changement de domaine est-il si

intéressant ? Parce qu'une image classique admet une grande continuité entre les valeurs des pixels. Les hautes fréquences étant réservées à des changements rapides d'intensité du pixel. Ainsi on parvient à représenter l'intégralité de l'information de l'image sur très peu de coefficients, correspondant à des fréquences plutôt basses.

La DCT s'applique à une matrice carrée. Le résultat produit est représenté dans une matrice de même dimension. Les basses fréquences se trouvant en haut à gauche de la matrice, et les hautes fréquences en bas à droite.

La formule de calcul de la DCT sur une matrice  $N \times N$  est donnée par :

$$DCT(i, j) = \frac{1}{\sqrt{2}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x, y) \cos\left(\frac{(2x+1)i\pi}{2N}\right) \cos\left(\frac{(2y+1)j\pi}{2N}\right) \quad (1)$$

La formule inverse est comme suit :

$$P(x, y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j) DCT(i, j) \cos\left(\frac{(2x+1)i\pi}{2N}\right) \cos\left(\frac{(2y+1)j\pi}{2N}\right) \quad (2)$$

Avec :

$$C(x) = \frac{1}{\sqrt{2}} \quad \text{si } x = 0, \quad 1 \quad \text{si } x > 0 \quad (3)$$

Rappelons que nous avons opté à implémenter une version réversible de la transformée en DCT afin d'appliquer une compression sans perte au niveau de la région d'intérêt. Cette partie sera détaillée dans les prochaines sections.

### B. Les codeurs imbriqués

Les ondelettes remportent un vif succès dans le domaine de la compression, et la littérature regorge d'études sur ce sujet. La représentation en ondelette d'une image  $y$  est vue comme un ensemble de coefficients ayant une structure hiérarchique en arbre.

Il existe à l'heure actuelle un grand nombre de méthodes de codage spécifiquement adaptées à la structure et aux propriétés des représentations multi échelles. Parmi les plus récentes et les plus efficaces, nous distinguons les codeurs EZW (Embedded image coding using Zerotrees of Wavelet coefficients) et SPIHT (Set Partitioning In Hierarchical Trees). Ces algorithmes de codage sont basés sur la structure naturelle en arbre de la décomposition multi-échelle et sur la dépendance inter-échelles des coefficients. Nous rappelons brièvement ci-dessous les principes de ces algorithmes de codage et nous renvoyons le lecteur à [9] et [10] pour une présentation détaillée.

1) *EZW*: C'est un algorithme simple de compression d'image remarquablement efficace, ayant la propriété que les bits du train de bits sont générés par ordre d'importance, produisant un code entièrement intégré. Cet algorithme effectue donc une succession d'approximations par quantifications imbriquées afin de permettre un codage progressif. L'information est ordonnée par précision, amplitude, échelle et localisation spatiale, et est embarquée dans une Zerotree spécifiant la carte de signification des coefficients. Le Zerotree représente l'arbre hiérarchique entre un coefficient d'ondelette et les coefficients correspondant à la même zone de l'image dans la sous-bande de même

orientation fréquentielle du niveau de résolution supérieur. Un nœud du Zerotree permet d'indiquer si tous les coefficients qui appartiennent au sous arbre sont inférieurs au seuil de quantification, et le cas échéant évite de les coder.

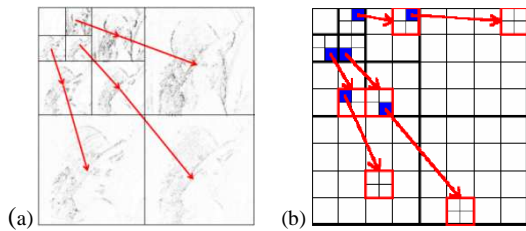


Fig. 1 (a) Dépendance entre sous-bandes (b) Hiérarchie des coefficients

L'algorithme parcourt les coefficients suivant l'ordre donné par la figure 1(a) à la recherche de ceux significatifs par rapport au seuil courant. En utilisant la hiérarchie donnée par la figure 1(b), l'algorithme produit alors une carte marquant la position des coefficients significatifs ainsi que leur signe.

Les étapes du LZW sont comme suit:

- ✓ Choisir un seuil  $T_i$
- ✓ Coder l'arbre en élagant les nœuds insignifiants ( $<T$ ) et dont la descendance est insignifiante par les symboles : PODS, NEG, IZ, ZTR
- ✓ Affiner les coefficients déjà codés afin d'avoir une précision  $T_i/2$
- ✓ Répéter avec  $T_i = T_{i-1}/2$

2) *SPIHT*: Tout comme le codeur EZW, cette méthode applique une quantification par approximations successives pour améliorer la précision de la représentation des coefficients d'ondelettes et pour faciliter le codage imbriqué. On effectue donc une partition des sous-bandes en sous-ensembles regroupant des coefficients non significatifs main selon un procédé différent de l'algorithme EZW. En effet, des coefficients non significatifs dans les basses fréquences ont de fortes chances d'avoir des enfants non significatifs dans les hautes fréquences, ceux possédant une énergie importante ont généralement des enfants possédant moins d'énergie. Le tri partiel réorganise ainsi les coefficients d'une branche en fonction de leur bit de poids le plus fort (réduction de l'erreur quadratique) à l'aide d'un parcours en profondeur des nœuds possédant un descendant significatif dans l'arbre hiérarchique contrairement à EZW qui effectue un parcours en largeur. Le codeur SPIHT est encore très utilisé comme référence pour l'évaluation d'autres méthodes et comme base à d'autres algorithmes

### III. APPROCHE PROPOSÉE

Dans le cas des images médicales, l'information pertinente est souvent diluée dans une masse d'information et qui doit être impérativement sauvegardée lors d'une opération de compression. Ainsi, les algorithmes de compression d'images médicales doivent être efficaces, adaptés et intelligents. Pour répondre à ces exigences, nous proposons un schéma de compression basé région d'intérêt. Les étapes de cette approche sont illustrées par la Fig.2.

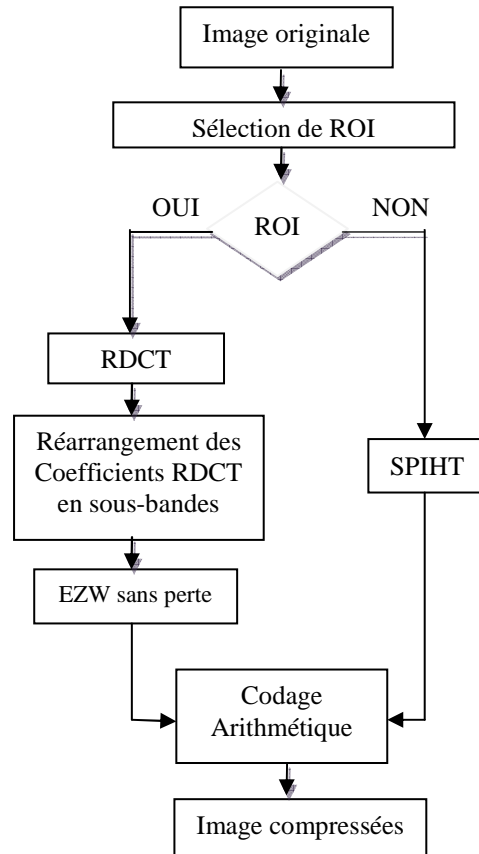


Fig.2 Schéma de codage proposé

#### A. Codage de la partie ROI

Nous avons appliqué la DCT Réversible (RDCT) sur la partie ROI dans un premier temps. Ensuite, les coefficients obtenus sont réorganisée en sous-bandes identiquement à la multi-résolution de la DWT [11].

Rappelons que la DCT est l'idée majeure pour la compression JPEG. En effet ce processus appartient à une classe d'opérations mathématiques, tout comme la Transformée de Fourier. Elle permet un changement de domaine d'étude, tout en gardant exactement la même fonction étudiée.

Pour une compression sans perte, la transformée réversible RDCT est évidemment intéressante car elle est simplement appliquée avec plus d'avantage et une faible consommation de

mémoire. En effet, comme la transformée en ondelette entière, ces transformations entières RDCT peuvent être obtenues en utilisant le schéma lifting utilisant un vecteur d'entrée de nombre entier et également en sortie [13].

Après avoir réarrange les coefficients DCT en sous-bandes, nous proposons une technique de compression sans perte basé sur un algorithme EZW modifié [12] pour coder la ROI.

1) *Réarrangement des coefficients RDCT*: Dans cette étape une transformée réversible RDCT est appliquée sur la ROI. Les coefficients obtenus sont par la suite réorganisés en sous-bandes similairement à une décomposition en ondelettes. Cette présentation de sous-bande est un simple réarrangement des coefficients de la DCT réversible. La Fig.3 illustre le réarrangement des blocs de 4×4 en sous-bandes. Remarquons que la plupart de l'énergie est concentrée dans la sous-bande en haute échelle, elle contient les coefficients DC des blocs RDCT. Les sous-bandes situées en échelles inférieures contiennent les coefficients AC. Par conséquent, la structure RDCT réarrangé apparait bien adaptée pour un codage d'arbre à zéro qui permettent de codage plus efficace. Le lecteur intéressée peut consulter le papier arboré dans [14] pour une étude détaillée sur le passage d'une transformée en DCT vers une présentation en sous-bande peut être obtenu à partir d'une DCT.

Pour chaque blocs de taille 8×8, on applique la relation illustrée dans l'équation suivante:

$$W[(i\%8)M/8 + i/8] [(j\%8)N/8 + j/8] = D[i][j] \quad (4)$$

$D[M][N]$ : matrice des coefficients RDCT de taille  $M \times N$

$W[M][N]$ : matrice des coefficients réarrangés de taille  $M \times N$

% Désigne l'opérateur modulo

$i=1, \dots, M, j=1, \dots, N$

A <sub>00</sub>	A <sub>01</sub>	A <sub>02</sub>	A <sub>03</sub>	B <sub>00</sub>	B <sub>01</sub>	B <sub>02</sub>	B <sub>03</sub>
A <sub>10</sub>	A <sub>11</sub>	A <sub>12</sub>	A <sub>13</sub>	B <sub>10</sub>	B <sub>11</sub>	B <sub>12</sub>	B <sub>13</sub>
A <sub>20</sub>	A <sub>21</sub>	A <sub>22</sub>	A <sub>23</sub>	B <sub>20</sub>	B <sub>21</sub>	B <sub>22</sub>	B <sub>23</sub>
A <sub>30</sub>	A <sub>31</sub>	A <sub>32</sub>	A <sub>33</sub>	B <sub>30</sub>	B <sub>31</sub>	B <sub>32</sub>	B <sub>33</sub>
C <sub>00</sub>	C <sub>01</sub>	C <sub>02</sub>	C <sub>03</sub>	D <sub>00</sub>	D <sub>01</sub>	D <sub>02</sub>	D <sub>03</sub>
C <sub>10</sub>	C <sub>11</sub>	C <sub>12</sub>	C <sub>13</sub>	D <sub>10</sub>	D <sub>11</sub>	D <sub>12</sub>	D <sub>13</sub>
C <sub>20</sub>	C <sub>21</sub>	C <sub>22</sub>	C <sub>23</sub>	D <sub>20</sub>	D <sub>21</sub>	D <sub>22</sub>	D <sub>23</sub>
C <sub>30</sub>	C <sub>31</sub>	C <sub>32</sub>	C <sub>33</sub>	D <sub>30</sub>	D <sub>31</sub>	D <sub>32</sub>	D <sub>33</sub>

Fig.3 Blocs de coefficients RDCT restructurés en sous-bandes

2) *Codage des coefficients RDCT*: Pour l'étape de codage, nous proposons quelques modifications pour simplifier l'algorithme EZW classique pour un codage sans perte et une amélioration du taux de compression.

La modification change les passes utilisées dans l'algorithme de codage. En effet, l'algorithme EZW utilisé deux passes: passe dominante et passe subordonnée. La passe dominante recherche les coefficients significatifs pour chaque plan de bits, et le la passe subordonnée quantifie alors ces coefficients significatifs trouvés.

Pour une compression sans perte, les résultats de deux passages doivent être stockés pour le décodage. Par

conséquent, le rapport de compression qui en résulte est une somme des taux des deux passes dominants et subordonnés bits.

Notons que le passe subordonnée contribue moyennement d'un tiers du débit total. Ainsi, afin d'épargner plus du taux de bits (amélioration du taux de compression), nous retirons la passe subordonnée. Contrairement à EZW de Shapiro, la phase de passe subordonnée est éliminée par le codage de valeurs résiduelles obtenues par différences successives (c'est à dire, en soustrayant le seuil subordonné à partir des coefficients de la liste subordonné) pour fournir une représentation exacte des coefficients pour chaque passage subordonnée

Le coefficient  $Im(x,y)$  dont la valeur absolue est plus grande que le seuil de courant, est reconnue comme étant un pixel significatif. Il est remplacé par une valeur résiduelle  $Res(x,y)$  dans l'image. C'est une autre façon de remplacer la phase de passe subordonnée tout en accomplissant la tâche de quantification des coefficients.

Une seconde modification consiste à ajouter deux autres symboles pour plus de commodité dans la mise en œuvre. Dans l'algorithme EZW de Shapiro, la liste dominante D est composé de quatre symboles (POS, NEG, IZ, et ZTR) chacun codé en binaire sur deux bits; ces symboles sont codés arithmétiquement avant la transmission.

Dans notre approche, en plus des quatre symboles codés proposées par Shapiro, on a ZD, qui indique que les coefficients descendants sont les Zerotree. L'autre est CN, pour indiquer les coefficients significatifs déjà numérisés.

Enfin, un codage arithmétique est appliqué sur tous le flux binaire afin d'augmenter les taux de compression.

## B. Codage de la partie Non-ROI

Un codage SPIHT est par la suite appliqué sur la partie arrière-plan (Non région d'intérêt). On a adopté la base bi-orthogonale de Daubechies dont le filtre d'analyse a neuf coefficients et le filtre dual de synthèse a sept coefficients. Cette base est généralement notée : « 9/7 Daubechies filter bank ».

Le codeur SPIHT génère une qualité d'image excellente et des performances en raison de plusieurs propriétés de l'algorithme de codage. Donc le choix de cette méthode revient essentiellement à l'efficacité de l'algorithme SPIHT qui permet d'appliquer un codage avec de forts taux de compression tout en ayant des reconstructions de qualités très acceptables. Enfin un codage arithmétique a été appliqué étape de codage entropique.

## IV. RÉSULTATS ET DISCUSSIONS

Cette section présente des résultats expérimentaux montrant l'efficacité de l'approche proposée. Nous présentons ici des tests effectués sur des images de cerveau de taille 256x256, présentées dans la Fig.4.

Afin d'évaluer la qualité de l'image reconstruite, l'erreur quadratique moyenne EQM (Equation 5), est calculée.

L'EQM est utilisée pour calculer le rapport crête signal sur bruit (PSNR) qui est donné par l'équation 5.

$$EQM = \frac{1}{N} \sum_i (\hat{x}_i - x_i)^2 \quad (5)$$

N : taille de l'image,  $x_i$  : pixels de l'image originale,  $\hat{x}_i$  : pixels de l'image reconstruite.

$$PSNR = 10 \log_{10} \frac{255 * 255}{EQM} \quad (6)$$

Le taux de Compression (TC) est donné également par  $T_{org}$  : Taille de l'image originale,  $T_{com}$  : taille de l'image compressée.

$$TC = \left(1 - \frac{T_{com}}{T_{org}}\right) \times 100 \quad (7)$$

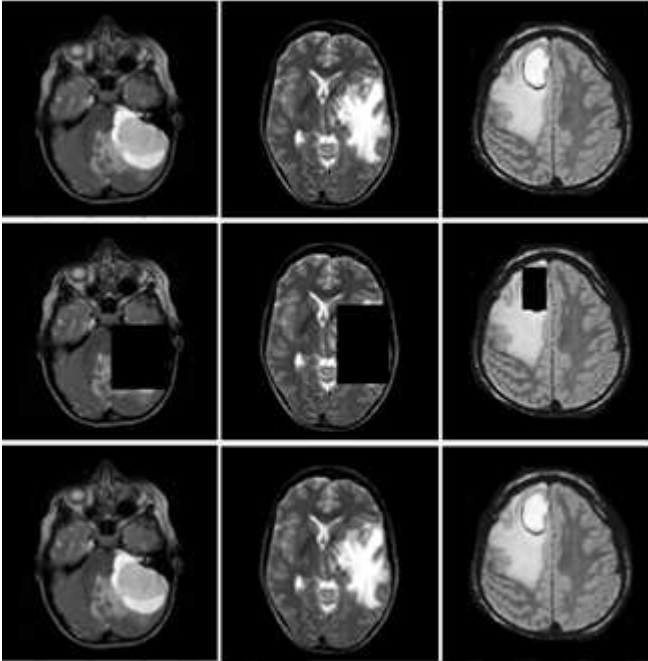


Fig.4 Résultats de compression par l'approche proposée: Image originales, sélection de ROI, images reconstruites

La sélection de la région ROI est faite sous forme rectangulaire couvrant la partie nécessaire au diagnostic (Zone tumorale). Les résultats de compression obtenus après la phase de sélection de la ROI sont présentés dans la Fig.4. Ces résultats montrent que les images reconstruites respectent la contrainte du codage sans perte au niveau de la région d'intérêt et gardent une apparence acceptable sur la partie restante.

TABLE I

RESULTATS COMPARATIFS EN MESURES DE PSNR ET TC

Images	EZW		SPIHT		Approche proposée	
	PSNR	TC	PSNR	TC	PSNR	TC
Image 1	33,14	56,86	34,22	64,58	36,25	60,29
Image 2	30,03	52,91	32,50	59,18	35,00	58,06
Image 3	31,64	50,15	33,45	60,02	34,83	55,68

La table 1 présente les résultats obtenus en termes de PSNR et taux de compression pour les trois images de test. En comparant les résultats avec ceux obtenus où une simple compression à été appliquée (sans ROI) par les codeurs SPIHT et EZW. L'approche proposée surpasse largement le codeur EZW en termes de PSNR et taux de compression. Par rapport au codeur SPIHT, on constate que notre approche aboutit à une augmentation en valeur de PSNR au détriment d'une très légère diminution en taux de compression (entre 2% et 5%).

TABLE II

RESULTATS COMPARATIFS EN TERMES DE BPP SUR ROI

Image	CALIC	JPEG2000	Approche proposée
Image1	5,15 bpp	4,27 bpp	4,31 bpp
Image2	5,18 bpp	4,24 bpp	4,35 bpp
Image3	6,11 bpp	4,75 bpp	4,23 bpp

La table 2 présente les résultats obtenus sur la partie ROI en termes de débits binaires bpp (bit par pixel), obtenus par des techniques sans perte (CALIC[15], JPEG2000[1]). Comparée à la norme JPEG2000, notre approche aboutit à des débits binaires compétitifs, nous remarquons aussi que ces résultats dépassent ceux obtenus par l'algorithme sans perte CALIC.

Cela revient essentiellement à l'efficacité de l'algorithme MEZW appliqués sur les coefficients de la RDCT qui ont été réarrangé en multi-résolution et qui semblent mieux approprié à coder des parties moyennement petites. Toutefois lors du codage avec région d'intérêt, on a aboutit à une reconstruction parfaite au niveau des zones d'intérêt dans le but de conserver les détails relatifs au diagnostic ce qui explique ce qui explique une augmentation en valeur de PSNR.

## V. CONCLUSION

Dans ce papier, une méthode basée sur un codeur EZW appliquée sur des coefficients DCT, a été présentée s'efforçant à accomplir une amélioration en termes de PSNR et taux de compression, sans détérioration de la qualité de l'image au niveau de la partie contenant de l'information relative au diagnostic médical. Cet algorithme a été comparé à d'autres



techniques comme SPIHT et EZW. Notre codeur détient la meilleure performance pour des images médicales en termes de PSNR.

Il existe des systèmes de codage hybride incluant des techniques d'analyse, de pré ou post-traitement d'images, qui sont de plus en plus courants. Dans ce contexte, une sélection automatique de la région d'intérêt se basant sur une technique de classification et une étude sur une quantification adaptative, sont des enjeux importants.

#### BIBLIOGRAPHIES

- [1] C. A. Christopoulos, A. N. Skodas, and T. Ebrahimi, "The JPEG 2000 still image coding system: An overview", *IEEE Trans. Consumer Electron.*, vol. 46, pp 1103-1127, Nov. 2000.
- [2] L. Huibao, J. Si, G.P. Abousleman, "Knowledge-based hierarchical Region-of-Interest detection", in *Proc. ICASSP'02*, IEEE International Conference, Orlando, Florida, pp 13-17, 2002 .
- [3] G.P. Nguyen, M. Worring, "An user based framework for salient detail extraction", Proceedings of the 2004 IEEE International Conference on Multimedia and Expo, ICME 2004, 27-30 June 2004, Taipei, Taiwan.
- [4] H.U. Min, Chnangjiang, Zhang, L.U. Juan, "A multi ROIs Medical image Compressin Compressin with Edge Feature Preserving", Proceedings of 2008 3rd international conference on Intelligent System and Knowledge Engineering
- [5] M. Firoozbakht, J. Dehmeshki, M. Martini, Y. Ebrahimdoost, H. Amin, M. Dehkordi, A. Youannic, SD. Qanadli, "Compression of digital medical images based on multiple regions of interest", The Fourth International Conference on Digital Society, IEEE, 2010.
- [6] R.F.L. Chavez, Y.Iano, R.S. Higa, R. Arthur and O.Saotome, "Generalized region of interest coding applied to SPIHT", *Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT)*, May Edition, 2012.
- [7] N. Londhe, S. Chawre, "Region Based Coding of Liver Cancer CT Images", *Canadian Journal on Image Processing and Computer Vision* Vol. 4 No. 1, pp. 9-15, January 2013.
- [8] K.R. Rao, P.Yip, "Discrete Transform algorithms, Advantages, Applications", Academic Press, Boston, 1990.
- [9] J. M. Shapiro. "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Transaction. Signal Processing*, vol. 41, no. 12, pp. 3445-3462. December 1993.
- [10] A. Said, W.A. Pearlman. "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", *Circuits and Systems for Video Technology, IEEE Trans. on*, vol.6, no. pp.243-250, Jun 1996.
- [11] Dj. Samai, N. Doghmane, M. Bedda and L. Lucas, "Shape-Adaptive DCT for Color Object-Based Coding", *J. Automation & Systems Engineering*, pp. 4-3, 2010.
- [12] A.Ouafi, A. Taleb-Ahmed , Z. Baarir and A. Zitouni., "A modified emdded zerotree wavelet (MEZW) algorithm for image compression", *Journal of Mathematical imaging and Vision*, vol.30, pp 298-307, 2008.
- [13] Yusong Yan, Qingyun Shi, "ReversibleInteger DCT and lossless image coding",*Software Journal*, Vol.11, No.5, pp.620-627, May, 2000
- [14] Yeong-An Jeong and Cha-Keon Cheong, " A DCT-based Embedded Image Coder using Wavelet Structure of DCT for very low bit rate video CODEC", *IEEE Trans. on Consumer Electronics*, Vol. 44, No. 3, pp.500-507, AUGUST 1998.
- [15] X. Wu, N. Memon, Context-based, adaptive, lossless image codec, *IEEE Trans. on Communications*. vol.45, no.4, pp.437-444, 1997.

# Partial Volume Correction in Brain MRI Segmentation

Messaoud HAMEURLAINE<sup>#</sup>, Abdelouahab MOUSSAOUI<sup>\*</sup>, Hadda CHERROUN<sup>#</sup>

<sup>#</sup> *faculty of sciences, Amar Thledji university,*

*Laghouat, Algeria*

<sup>1</sup> hamessainf@yahoo.fr

<sup>3</sup> hadda\_cherroun@mail.lagh-univ.dz

<sup>\*</sup> *faculty of sciences, Ferhat Abbas university,*

*Setif, Algeria*

<sup>2</sup> moussaoui\_abdel@yahoo.fr

**Abstract**—A single voxel in a medical image may be composed of mixture of tissue types, it is the partial volume effect (PVE). Many partial volume correction techniques are used in literature but what approach is best. This paper presents a validation study of statistical brain magnetic resonance image segmentation algorithms with partial volume correction. The methods are tested on simulated data for which the classification ground truth is known. Validity of the various classification methods as well as in the tissue volume is estimated with different local and global measures. Results demonstrate that the partial volume is perfectly modelled by methods that account for mixture classes and spatial information. Overall, we can't say that method is more efficient to another by a global way, but depending on the case and requirements we distinguish a more interesting one.

**Keywords**— Magnetic resonance imaging, Brain tissue models, segmentation, hidden Markov random field, partial volume, Expectation-maximization, ICM algorithm.

## I. INTRODUCTION

Accurate and robust brain tissue segmentation from magnetic resonance (MR) images is a key issue in many applications of medical image analysis for quantitative studies and particularly in the study of several brain disorders such as Alzheimer's disease or Schizophrenia [1]-[3]. Moreover, brain tissue segmentation can also be required as preliminary step of image processing algorithms such as, for instance, voxel-based morphometry [4] or image registration [5]. Manual tracing by an expert of the three brain tissue types—white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF)—is exceedingly time consuming as the volume of data involved in magnetic resonance imaging (MRI) studies is large. On the other hand, automated and reliable tissue classification is a challenging task as the intensity representation of the data typically does not allow a clear delimitation of the different tissue types present in a MRI, because of partial volume (PV) effect, image noise and intensity nonuniformities caused by magnetic field inhomogeneities. Niessen et al. [6] showed that consistently misplacing the tissue borders in a 1 mm isotropic brain MR image with only a single pixel in each slice resulted in volume errors of approximately 30%, 40%, and 60% for white matter, gray matter and cerebrospinal fluid (CSF), respectively. This

problem is also important in the segmentation of PET image [25]-[27][36].

The goal of this work is to assess the robustness and accuracy of unsupervised classification methods with estimation of partial volume. This paper is organized as follows. First, in Section II, the state of the art is presented. Then, in Section III, the methods analysed in this comparative study are summarized. In Section IV, the data set used for this assessment study, the validation measures, the classification results are presented and discussed. Finally, conclusions are in Section V.

## II. STATE OF THE ART

Statistical classification methods usually solve the estimation problem by either assigning a class label to a voxel or by the estimation of the relative amounts of the various tissue types within a voxel [7][8]. Finite Gaussian Mixture (FGM) models, that assume a Gaussian distribution for the image intensities, are widely used and their parameter estimation problem is typically solved in an expectation-maximization (EM) framework [9]-[11].

Other algorithms [7] [12] add separate classes to take into account the PV voxels and model them also by independent Gaussian densities. A more realistic model of PV than Gaussian is proposed by Santiago et al. [13][14] and it is extensively used by other authors [15]-[17]. Recently, new methods model the spatial information by a Markov random field (MRF) [18][7][19]-[21].

A nonparametric classification technique can be considered also when no well justified parametric model is known [22][23]. Tohka et al. [28] proposed an algorithm which used statistical estimators, based on the MAP estimation [18]. Chiverton and Wells [29] presented a local adaptive Gradient-controlled spatial regularizer (GSR) using a Markov Random Field to model the class membership and a Markov Chain Monte Carlo (MCMC) simulation to adapt the model to the observed data. On the other hand, PV estimation has received considerable attention and different approaches have been proposed for classification and computation of fractional content [24][27][30]-[33].

### III. METHODS

In this section, the methods used in this comparative study are presented.

TABLE I  
THE FORTH METHODS UNDER STUDY

Method	Intensity model	Spatial model
G-HMRG	5 gaussian	M. Random Field
GPV-HMRF	3Gaussian + 2 PVE	M. Random Field
NP-HMRF	Non-parametric	M.Random Field
FAST-PVE	3Gaussian + 3 PVE	M. Random Field

The first method considers the finite Gaussian mixture model to represent the image and a hidden MRF model to account for spatial prior information as in [23]. The second method uses a specific PV distribution for mixture Tissues as in [10],[11]. The third algorithm model the tissue classes by nonparametric probability densities models [26], the resulting algorithm minimizes an information theoretic quantity, called the error probability. The last approach present an extremely fast method named FAST-PVE for tissue classification and partial volume estimation using a Markov Random Field (MRF) based spatial prior and the fast ICM algorithm, which is generally applicable to any MRF-based segmentation method [31].

#### A. Image Model

The image contains  $N$  data points to be classified with  $i \in S = \{1, 2, \dots, N\}$  indexing the image voxels.  $Y$  is the random variable associated to the data features  $y_i$ , with the set of possible outcomes  $\mathcal{D}$ .  $y_i$  represents the intensity of voxel  $i$ . The classification process aims to classify the data  $S$  into one of the hidden underlying classes present in the image. The family of random variables  $X$  represents these classes.  $x = \{x_1, x_2, \dots, x_N\}$  denotes a possible configuration of  $X$ .

Assuming that all the random variables  $y_i$ , are identically and independently distributed. Then, the probability density function of the voxel intensity is :

$$P(y_i) = \sum_{x_i \in \mathcal{L}} P(x_i)P(y_i|x_i) \quad (1)$$

Where  $P(x_i)$  is the prior probability of the tissue class  $x_i$  and  $P(y_i|x_i)$  is the conditional probability density function of  $y_i$  given the tissue class  $x_i$ . The prior probability  $P(x_i)$  is used to model the spatial coherence of the images. The transition probability  $P(y_i|x_i)$  models the image intensity formation process for each tissue type. Different models are used for pure tissues and for tissue mixtures.

The probability density function of the observed intensity  $y$  for the pure tissue class is Gaussian :

$$p(y|x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(y-\mu_x)^2}{2\sigma_x^2}} \quad (2)$$

where the model parameters  $\theta_x = \{\mu_x, \sigma_x\}$  are, respectively, the mean and standard deviation of the Gaussian.

Two different models of mixture tissues are considered. The simplest one assumes that the PV can be modeled by a Gaussian distribution as proposed in [12] and [20]. In this case, (2) is used both for pure and mixture tissues. This model is used in subsection C. A more complex probability density model for mixture tissues was proposed by Santiago *et al.* [10] and used by [18] and [36]. This model is used in subsection D and F. A mixture tissue voxel has a probability density function :

$$p(y|x, \alpha) = \frac{1}{\sigma_x(\alpha)\sqrt{2\pi}} e^{-\frac{(y-\mu_x(\alpha))^2}{2\sigma_x^2(\alpha)}} \quad (3)$$

where the two pure tissues composing the voxel are denoted by  $l_1, l_2$ , and  $\alpha$  is the fraction of tissue  $l_1$  present in the mixture voxel. The mean and variance of the mixture are determined by the model parameters of the pure tissues:

$$\mu_x(\alpha) = \alpha\mu_{l_1} + (1-\alpha)\mu_{l_2} \quad (4)$$

$$\sigma_x^2(\alpha) = \alpha^2\sigma_{l_1}^2 + (1-\alpha)^2\sigma_{l_2}^2 \quad (5)$$

The probability density function for the whole PV tissue  $x$  is :

$$P(y|x) = \int_0^1 P(y|x, \alpha)P(\alpha|x)dx \quad (6)$$

choosing the correct function for  $P(\alpha|x)$  is a complex issue [37]. This model is used in subsection D and F. The most authors assume a uniform distribution [24].  $\alpha$  is approximately uniform around  $\alpha=0.5$  with peaks at  $\alpha=0$  and  $\alpha=1$ , this leads to :

$$P(y|x) = \int_0^1 P(y|x, \alpha)dx \quad (7)$$

$P(y|x)$  varies depending on the parameters  $\theta_l = \{\mu_l, \sigma_l\}$ .

Other algorithm incorporates a non-uniform partial volume density. Finally, it is also possible not to make any assumption on the shape of the probability density functions of each tissue class. Nonparametric, information theoretic alternatives are considered in [22][38][36]. This model is described in subsection E.

#### B. Spatial Distribution Model

The other term in equation (1) is  $P(x_i)$ . It describes the prior knowledge about the spatial distribution of brain tissues in the image volume. Generally this is done by using a MRF to model spatial interactions among tissue classes [39], [40].

The sites in the image  $S$  are related with a neighborhood system  $N = \{N_i, i \in S\}$ . A random field  $x$  is a MRF on  $S$  with respect to  $N$  if and only if :

$$P(x) > 0, x \in X \quad (8)$$

$$\text{and } P(x_i|x_{S-\{i\}}) = P(x_i|x_{N_i}) \quad (9)$$

According to the Hammersley-Clifford theorem, a MRF can be estimated by a Gibbs distribution :

$$P(x) = Z^{-1} e^{-U(x, \beta)} \quad (10)$$

Where  $U(x)$  is the energy function,  $\beta$  the spatial parameter and  $Z$  a normalization factor.

Several definitions of the energy function are given in the framework of image segmentation [41]. The energy function is denoted by :

$$U(x|\beta) = \sum_{i \in S} \left( V_i(x_i) + \frac{\beta}{2} \sum_{j \in N_i} V_{ij}(x_i, x_j) \right) \quad (11)$$

The external field  $V_i(x_i)$  weighs the relative importance of the different classes present in the image. A no external energy ( $V_i(x_i)=0$ ) is used [46]. The interactions between neighbors is modeled by  $V_{ij}(x_i, x_j)$ . Taking into account the distance between neighboring and to preserve thin structures, a function is in [24].

$$V_{ij}(x_i, x_j) = \frac{\delta(x_i, x_j)}{d(i, j)} \quad (12)$$

where

$$\delta(x_i, x_j) = \begin{cases} -2, & \text{if } x_i = x_j \\ -1, & \text{if they share a tissue type} \\ +1, & \text{otherwise} \end{cases} \quad (13)$$

and  $d(i, j)$  is the distance between voxels  $i$  and  $j$ . The spatial parameter  $\beta$  controls the relative influence of the spatial prior over the intensity model. In sections C, D, and E, the value of  $\beta$  is fixed empirically to 1.2 by classifying a training set [24]. Finally,  $Z(U)$  defined the normalization factor of the Gibbs distribution as:

$$Z(U) = \sum_x e^{-U(x, \beta)} \quad (14)$$

Instead of computing  $Z$ , the conditional probabilities  $P(x|x_{N_i})$  are normalized by forcing the equality:

$$\sum_{x_i} P(x_i|x_{N_i}) = 1 \quad (15)$$

### C. Gaussian Hidden Markov Random Field Model

In section C and D, the voxels are classified into  $K=5$  tissue types representing tissue types (pure tissue classes  $P=\{\text{CSF};\text{GM};\text{WM}\}$ ) and partial volume mixtures of two tissue types (mixed Classes  $M=\{\{\text{CSF};\text{GM}\};\{\text{GM};\text{WM}\}\}$ ).

In the GHMRF(Gaussian Hidden Markov Random Field) model, Gaussian distribution model each brain tissue and spatial information is taken into account. The image intensity distribution function is:

$$P(y|\theta) = \sum_x P(x|x_{N_i}) \cdot f_x(y|\theta_x) \quad (16)$$

Where  $f_x(y|\theta_x)$  is a Gaussian distribution parameterized by  $\theta_x = \{\mu_x, \sigma_x\}$ .  $P(x|x_{N_i})$  represents the locally dependent probability of the tissue class  $x_i$  depends on the voxel neighborhood  $x_{N_i}$ . The parameters  $\theta$  estimated by the maximize of the log-likelihood function.

$$\hat{\theta} = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} P(y|\theta) \quad (17)$$

An adapted version of the EM algorithm called HMRF-EM is used, as suggested in [20].

**Initialization step:** best initialization for  $\theta(\hat{\theta})$   
**Expectation step:** compute the a posteriori probabilities  $\forall x$

$$\hat{P}^{(k)}(x|y_i, \hat{\theta}) = \frac{P(y_i|\hat{\theta}_x^{(k-1)}) \cdot \hat{P}^{(k-1)}(x|x_{N_i})}{\sum_l P(y_i|l, \hat{\theta}_l^{(k-1)}) \hat{P}^{(k-1)}(l|l_{N_i})} \quad (18)$$

**Maximization step:**

$$\hat{\omega}_x^{(k)} = \hat{P}^{(k)}(x|x_{N_i}) \quad (19)$$

$$\hat{\mu}_x^{(k)} = \frac{\sum_{i \in S} \hat{P}^{(k)}(x|y_i, \hat{\theta}) y_i}{\sum_{i \in S} \hat{P}^{(k)}(x|y_i, \hat{\theta})} \quad (20)$$

$$(\hat{\sigma}_x^{(k)})^2 = \frac{\sum_{i \in S} \hat{P}^{(k)}(x|y_i, \hat{\theta}) (y_i - \hat{\mu}_x^{(k)})^2}{\sum_{i \in S} \hat{P}^{(k)}(x|y_i, \hat{\theta})} \quad (21)$$

Algorithm 1: E.M algorithm

In equation (19), the term  $\hat{P}^{(k)}(x|x_{N_i})$  is estimated as in [24] and detailed in [47].

Finally, classification is performed by the class for each voxel that maximizes the posterior probability.

### D. GPV and HMRF Model: GPV-HMRF

The second approach relies only on the intensity information. Pure tissue intensities are modelled by Gaussian distributions while mixture tissues are modelled as proposed by Santago *et al.* [13], [14]. The resulting probabilistic model is the same as equation (16), with  $f_x(y|\theta_x)$  defined either as a Gaussian for pure tissues or by the PV equation (7).

The parameter optimization is performed similarly to the GHMRF algorithm. The modified EM-algorithm becomes, as in [38] :

$$\hat{P}^{(k)}(x|y_i, \hat{\theta}) = \frac{\hat{P}(y_i|x, \hat{\theta}_x^{(k-1)}) \cdot \hat{P}^{(k-1)}(x|x_{N_i})}{\sum_l \hat{P}(y_i|l, \hat{\theta}_l^{(k-1)}) \hat{P}^{(k-1)}(l|l_{N_i})} \quad (22)$$

$$\hat{\mu}_x^{(k)} = \frac{\sum_{i \in S} \hat{P}^{(k)}(x|y_i) y_i}{\sum_{i \in S} \hat{P}^{(k)}(x|y_i)} \quad (23)$$

$$(\hat{\sigma}_x^{(k)})^2 = \frac{\sum_{i \in S} \hat{P}^{(k)}(x|y_i) (y_i - \hat{\mu}_x^{(k)})^2}{\sum_{i \in S} \hat{P}^{(k)}(x|y_i)} \quad (24)$$

In this approach, equation(23) and equation (24) are only computed for pure tissues. Besides  $\hat{P}(y_i|x, \hat{\theta}_x)$  in equation (22) is either a Gaussian or a PV distribution depending on the tissue type. The term  $\hat{P}^{(k)}(x|x_{N_i})$  is estimated as in [24] and detailed in [15].

### E. Nonparametric HMRF: NP-HMRF

Finally, it is also possible not to make any assumption on the shape of the probability density functions of each tissue class. Nonparametric, information theoretic alternatives are considered in this work [22], [23], [35]. For nonparametric classification, the posterior probability from parametric

classification,  $(y|x)$ , is replaced by an error probability which is defined as follows:

$$P_e = \sum_{x^{est}} \sum_y \sum_x P(E = 1|x^{est}, x) \cdot P(x^{est}|y)P(y|x) \cdot P(x) \quad (25)$$

Where  $x^{est}$  is a realization of random variable  $X^{est}$  which estimates  $X$  from  $Y$ . The probabilities  $P(x^{est}|y)$  and  $P(y|x)$  are estimated by Parzen-window probability density estimation [22]:

$$P(x^{est}|y) = \sum_{y_2 \in S_{x^{est}}} \frac{1}{|y_2|} G(y - y_2, \sigma_2^2) \quad (26)$$

$$P(y|x) = \sum_{y_1 \in S_x} \frac{1}{|S_x|} G(y - y_1, \sigma_1^2) \quad (27)$$

Where  $G(y - \mu, \sigma^2)$  is a Gaussian of expectation  $\mu$  and variance  $\sigma^2$ ,  $S_x$  denotes the set of voxels being classified into class  $x$ ,  $|S_x|$  is the number of elements of this set, and  $|y|$  is the number of samples with intensity  $y$ .

The probability  $P(E = 1|x^{est}, x)$  is called the distortion of the nonparametric classification algorithm, and is given by eq.28 where  $k$  is a class tissue.

$$P(E = 1|x^{est}, x) = (1 - \delta_{x^{est}, x}) \cdot \sum_k \frac{1}{|S_k|} \quad (28)$$

In complete analogy to parametric HMRF, the prior probabilities,  $P(x)$  are modeled by a Gibbs distribution equation (1). Then, the classification objective consists of determining the class label map  $\hat{X}$  that minimizes an error probability:

$$\hat{X} = \arg \min_X P(X)P_{e|X} \quad (29)$$

#### F. FAST-PVE Method

First, in the step of PV classification, voxels are tissue classified into  $K = 6$  tissue types representing the main tissue types (pure tissue classes  $P = \{\text{CSF}; \text{GM}; \text{WM}\}$ ) and partial volume mixtures of two tissue types (mixed Classes  $M = \{\{\text{Background}; \text{CSF}\}; \{\text{CSF}; \text{GM}\}; \{\text{GM}; \text{WM}\}\}$ ).

Second, in the PVC estimation step, FAST-PVE estimates the proportions of each tissue type within each voxel using the information from the PV classification step to make the estimation problem well-posed.

The PV classification is casted as an optimization problem and modelled as in subsections A and B. The model parameters must be estimated before the PV estimation [31]. The main differences are:

$$\beta = 0.1 ;$$

$$\delta(x_i, x_j) = \begin{cases} 2, & \text{if } x_i = x_j \\ 1, & \text{if they share a tissue type} \\ -1, & \text{otherwise} \end{cases} \quad (30)$$

The most central component to the speed of FAST-PVE is the fast ICM algorithm for solving the optimization problem.

The fast ICM algorithm to solve the optimization problem is presented below and detailed in [31].

In the PVC estimation step, FAST-PVE estimates the proportions  $\alpha$  of each tissue type within each voxel. the maximum-likelihood principle is used :

$$\alpha_{ij}^* = \arg \max_{\alpha \in [0,1]} \log g(x_i | \mu(\alpha), \sigma(\alpha)) \quad (31)$$

The estimate of the maximum-likelihood PVC and Parameter estimation are detail in [32].

#### IV. RESULTS ON SIMULATED DATA

We perform PV estimation on the simulated MR brain image obtained from Brain Web (<http://www.bic.mni.mcgill.ca/brainweb>). The results on the simulated data are compared to the ground truth classification. The most used measures of the validity of the classification methods are presented.

First, visual comparison of slide of each of the volumes classified by each of the methods with the corresponding slide of the ground truth classification volume is made for image with  $N=7\%$  and  $Rf=20\%$ . Visually, we can see that the result of the segmentation differs from one method to another.

Second, Results describe the percentage of voxels correctly classified in the following tests for all levels of noise and inhomogeneities are made.

For  $RF=20\%$ , FAST is more sensible to noise, GHMRF, GPV-HMRF and FAST nearly the same results, and they show a significant superiority over NP-HMRF method.(fig.1)

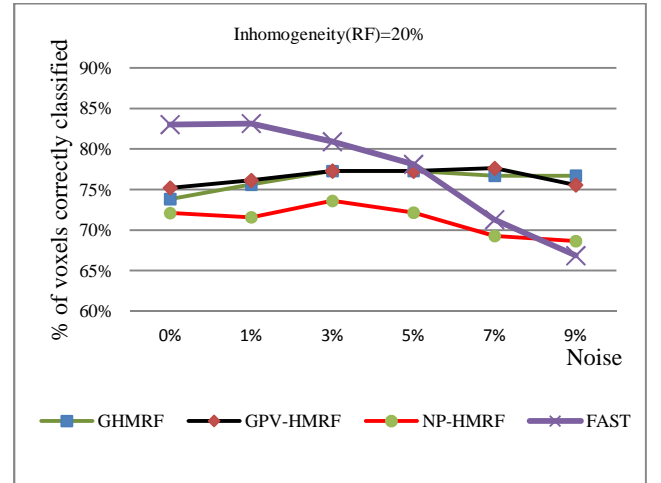


Fig.1: The percentage of voxels correctly classified  $RF = 20\%$

Third, The Dice Similarity Measure (DSM) between the ground truth and automatic segmentation is also used as the performance measure. For  $N=7\%$ , the DSM values estimated in GM+CSF mixture (fig.2), GM and GM+WM mixture tissue extraction by the all methods are not sensible to levels

of inhomogeneities, GHMRF, GPV-HMRF and NP-HMRF nearly the same results, and they show a low superiority over FAST method.

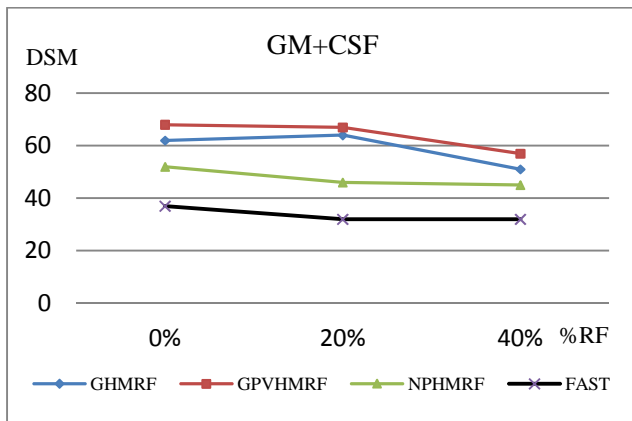


Fig. 1 : Comparison of DSM values for GM+CSF (N=7%)

Forth, a volumetric comparison for each tissue of the synthetic data is estimated. The volume of CSF, GM, And WM are estimated by counting the voxels of each class, the PV voxels also contribute to the volume of each tissue. The accuracy of the PV correction estimates was measured for each tissue type separately by root mean square error(RMSE). In fig.3 the error is shown as percentage with respect to the real volume of each tissue over all phantoms (5N0RF, 7N0RF, 5N20RF, 5N20RF, 7N20RF, 9N20RF). GHMRF and GPV-HMRF show a significant inferiority error over FAST, that can be justified by the estimation method of volume and RMSE.

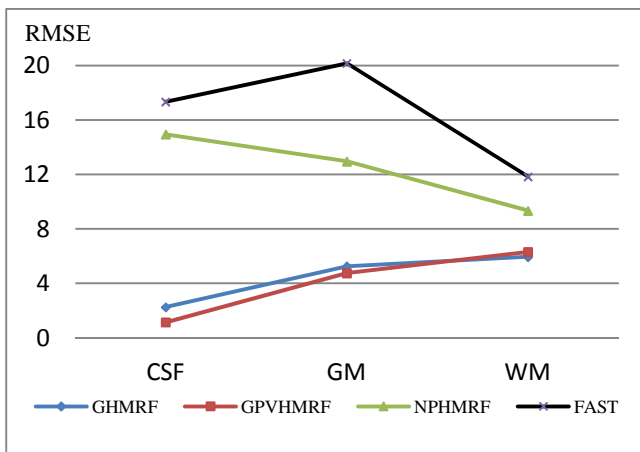


Fig.3: Root mean square error comparison

Finally, the computation time is estimated for all methods. The average time is 3-5 minutes for FAST, 10-15 minutes for GHRMF, around 20 minutes for GPV-HMRF and superior than 1 hour for NP-HMRF. FAST show a significant speed over all other methods.

Overall, we can't say that method is more efficient than another by a global way, but depending on the case and requirements we distinguish a more interesting one.

## V. CONCLUSION

In this paper, a comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction were presented. Four methods were applied on images with different levels of noise and inhomogeneities. Simulated datasets were used to perform the comparison and a number of widely used metrics were used to determine the efficacy of the respective segmentation algorithms. As the results of experiments confirm, quality segmentation obtained is variable depending on the level of noise and inhomogeneities. We can also say that the FAST method is the fastest. Overall, we can't say that method is more efficient than another by a global way, but depending on the case and requirements we distinguish a more interesting one.

## REFERENCES

- [1] S.K.Warfield, K.H.Zou, W.M.Wells, , "Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 23(7), pp. 903-921,2004.
- [2] C. Guttman, R. Benson, S. K.Warfield, X.Wei, M. Anderson, C. Hall, K. Abu-Hasaballah, J. Mugler, and L. Wolfson, "White matter abnormalities in mobility-impaired older persons," *Neurology*, vol. 54, no. 6, pp. 1277-1283, 2000.
- [3] F. Maes, K. Van Leemput, L. E. DeLisi, D. Vandermeulen, and P. Suetens, "Quantification of cerebral grey and white matter asymmetry from MRI," *Med. Image Computing Comput.-Assist. Intervention*, pp.348-357, 1999.
- [4] J. Ashburner and K. Friston, "Voxel-based morphometry: The methods," *NeuroImage*, vol. 11, pp. 805-821, 2000.
- [5] S. K. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis, "Adaptive, template moderated, spatially varying statistical classification," *Med. Image Anal.*, vol. 4, no. 1, pp. 43-55, Mar. 2000.
- [6] W. J. Niessen, K. L. Vincken, J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, "Multiscale segmentation of three-dimensional MR brain images," *Int. J. Comput. Vis.*, vol. 31, no. 2/3, pp. 185-202, 1999.
- [7] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 885-896, Oct 1999.
- [8] L. Nocera and J. C. Gee, "Robust partial volume tissue classification of cerebral MRI scans," *Proc. SPIE (Medical Imaging 1997: Image Processing)*, vol. 3034, pp. 312-322, 1997.
- [9] W.Wells, R. Kikinis, W. Grimson, and F. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 429-442, Aug. 1996
- [10] P. Schroeter et al., "Robust parameter estimation of intensity distributions for brain magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 172-186, Apr. 1998.
- [11] R. Guillemaud and M. Brady, "Estimating the bias field of MR images," *IEEE Trans. Med. Imag.*, vol. 16, no. 3, pp. 238-251, Jun. 1997.
- [12] S. Ruan, C. Jaggi, J. Xue, and J. Bloyet, "Brain tissue classification of magnetic resonance images using partial volume modeling," *IEEE Trans. Med. Imag.*, vol. 19, no. 12, pp. 172-186, Dec. 2000.
- [13] P. Santago and H. D. Gage, "Quantification of MR brain images by mixture density and partial volume modeling," *IEEE Trans. Med. Imag.*, vol. 12, no. 3, pp. 566-574, Sep. 1993.
- [14] P. Santago and H. Gage, "Statistical models of partial volume effect," *IEEE Trans. Image Process.*, vol. 4, no. 11, pp. 1531-1540, Nov. 1995.
- [15] A. Noe, S. Kovacic, and J. C. Gee. "Segmentation of cerebral mri scans using a partial volume model, shading correction, and an anatomical prior". *Proc. SPIE2001 (Medical Image Processing)* 0001;1466-1477 , 2001.
- [16] D. H. Laidlaw, K. W. Fleischer, and A. H. Barr, "Partial-volume Bayesian classification of material mixtures in MR volume data using

- voxel histograms," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 74–86, Feb. 1998.
- [17] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "A unifying framework for partial volume segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 105–119, Jan. 2003.
- [18] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage*, vol. 13, pp. 856–876, 2001.
- [19] K. Held, E. Rota Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H.-W. Müller-Gärtner, "Markov random field segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 878–886, Dec. 1997.
- [20] Y. Zhang et al., "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.
- [21] S. Ruan, B. Moretti, J. Fadili, and D. Bloyet, "Fuzzy Markovian segmentation in application of magnetic resonance images," *Comput. Vis. Image Understanding*, vol. 85, no. 1, pp. 54–69, 2002.
- [22] T. Butz, "From error probability to information theoretic signal and image processing," Ph.D. dissertation, *Signal Process. Inst., Swiss Federal Inst. Technol., Zurich, Switzerland*, Jun 2003.
- [23] T. Butz, P. Hagmann, E. Tardif, R. Meuli, and J.-P. Thiran, "A new brain segmentation framework," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 2003, vol. , *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 586–593, 2003.
- [24] M.B. Cuadra, L. Cammoun, T. Butz, O. Cuisenaire, and J.P. Thiran, "Comparison and Validation of Tissue Modelization and Statistical Classification Methods in T1-Weighted MR Brain Images," *IEEE Transactions on Medical Imaging - TMI* , vol. 24, no. 12, pp. 1548–1565, 2005.
- [25] CC Meltzer, MN Cantwell, PJ Greer, et al. "Does cerebral blood flow decline in healthy aging? APET study with partial-volume correction". *J Nucl Med.* 2000;41:1842–1848.
- [26] H. Zaidi, T. Ruest, F. Schoenahl, and M.L. Montandon "Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET" *NeuroImage* 32 (2006) 1591 – 1607
- [27] D. Gutierrez, M.L. Montandon, F. Assalb, M. Allaouaa, O. Ratiba, K.O. Lovblad, H. Zaidi, "Anatomically guided voxel-based partial volume effect correction in brain PET: Impact of MRI segmentation" *Computerized Medical Imaging and Graphics* 36 (2012) 610– 619
- [28] J. Tohka, A. Zijdenbos, A. Evans. "Fast and robust parameter estimation for statistical partial volume models in brain MRI" *NeuroImage* 2004;23(1):84–97.
- [29] J. Chiverton, K. Wells "Adaptive partial volume classification of MRI data" *Physics in Medicine and Biology* 2008;53(20):5577–94.
- [30] O. Acosta, P. Bourgeat, M.A. Zuluaga, J. Fripp, O. Salvado, S. Ourselin "Automated voxel-based 3D cortical thickness measurement in a combined lagrangian eulerian PDE approach using partial volume maps." *Medical Image Analysis* 2009;13:730–43.
- [31] J. Tohka . "FAST-PVE: Extremely Fast Markov Random Field Based Brain MRI Tissue Classification" SCIA Image Analysis, Lecture Notes in Computer Science vol 7944 pp. 266 - 276, 2013.
- [32] R.M. Brouwer , H.E. Hulshoff Pol, H.G. Schnack" Segmentation of MRI brain scans using non-uniform partial volume densities" *NeuroImage* 49 (2010) 467–477.
- [33] S. M'hiri , S. Mabrouk , F. Ghorbel "Segmentation des IRM cérébrales par une variante bootstrapée du HMRF-EM : étude préliminaire sur fantômes » *IRBM* 33 (2012) 2–10.
- [34] B.F. Hutton a,n, B.A. Thomas a, K. Erlandsson a, A. Bousse a, A. Reilhac-Laborde b, D. Kazantsev c, S. Pedemonte c, K. Vunckx d, S.R. Arridge c, S. Ourselin c « What approach to brain partial volume correction is best for PET/MRI? " *Nuclear Instruments and Methods in Physics Research A* 702 (2013) 29–33.
- [35] E. Gokcay and J. Principe, "Information theoretic clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 2, pp. 158–171, Feb. 2002.
- [36] A. Noe and J. C. Gee. "Partial volume segmentation of cerebral mri scans with mixture model clustering" presented at *Information Processing in Medical Imaging: 17th International Conference, IPMI 2001* Davis, CA, USA.
- [37] M. A. Gonzalez Ballester, A. P. Zisserman, and M. Brady, "Estimation of the partial volume effect in MRI," *Med. Image Anal.*, vol. 6, no. 4, pp. 389–405, Dec. 2002.
- [38] T. Butz, P. Hagmann, E. Tardif, R. Meuli, and J.-P. Thiran, "A new brain segmentation framework," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 2003, vol. , *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 586–593.
- [39] J. Zhang, "The mean field theory in EM procedures for Markov random fields," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2570–2583, Oct. 1992.
- [40] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, I. Karatzas and M. Yor, Eds. New York: Springer, 1995.
- [41] N. Peyrard, "Approximations de type champ moyen des modles de champ de Markov pour la segmentation de donnes spatiales," *Ph.D. Dissertation, University J. Fourier, Grenoble, France*, Oct. 2001.
- [42] G. Celeux, F. Forbes, and N. Peyrard, "EM-based image segmentation using potts models with external field," *INRIA, Paris, France, Tech. Rep.* 4456, Apr. 2002.
- [43] M. B. Cuadra, "Atlas-based segmentation and classification of magnetic resonance brain images," *Ph.D. dissertation, Signal Process. Inst., Swiss Federal Inst. Technol., Zurich, Switzerland*, Nov. 2003. thesis 2875.

# Face Spoofing Detection From Single Images Using Active Shape Models with Stasm And LBP

Azeddine Benlamoudi\*, Djamel Samai\*, Abdelkrim Ouafi†, Abdelmalik Taleb-Ahmed‡, Salah Eddine Bekhouche† and Abdenour Hadid§

\*Laboratory of LAGE, University of Ouargla, Algeria. be.azeddine@gmail.com ,Samai.djamel@gmail.com

†Laboratory of LESIA, University of Biskra, Algeria. ou\_karim@yahoo.fr, salah@bekhouche.com

‡LAMIH, UMR CNRS 8201 UVHC, University of Valenciennes, France Abdelmalik.Taleb-Ahmed@univ-valenciennes.fr

§Center for Machine Vision Research, PO Box 4500, FI-90014 University of Oulu, Finland. hadid@ee.oulu.fi

**Abstract**—Besides the recognition task, today's biometric systems need to cope with additional problem: spoofing attacks, like presenting a photo of a person(client) to camera. We study in this paper an anti-spoofing solution for distinguishing between 'live' and 'fake' faces. In our approach we focused in face detection using Viola-Jones algorithm and Active Shape Models with Stasm for locating landmarks. Then, we apply Local Binary Patterns (LBP) operator to extract the features in each region of the image. Finally, we use a nonlinear Support Vector Machine (SVM) classifier with kernel function for determining whether the input image corresponds to a live face or not. Our experimental analysis on a publicly available database NUAA, showed excellent results compared to existing methods.

**Keywords**—BIOMETRIC, SPOOFING, STASM, LBP, SVM, NUAA.

## I. INTRODUCTION

Nowadays we are experiencing an increasing demand for highly secure identification and personal verification technologies. This demand becomes even more apparent as we become aware of new security breaches and transaction frauds [1]. The main reason is that a biometric sample is a face represented in a digital image, which is intrinsically highly reproducible by several means like printed photos and electronic portable devices capable of showing images and videos (laptops and even cellular phones have nowadays wide and very good quality screens) [2].

In the context of face biometrics, an impostor tries to access the system as a valid user with three approaches [3]:

- Showing photography of a valid user
- Showing a video of a valid user, or
- Showing a 3D facial model of a valid user

Unfortunately, research in countermeasures to this type of attack has not kept-up-even if such threats have been known for nearly a decade. There seems to exist no consensus on best practices, techniques or protocols for developing and testing spoofing-detectors for face recognition[4].

Commonly cited papers refer to the problem of photo attack detection in two major complementary directions [2], [5], [6]:

- Static analysis, based on the fundamental idea that during the manufacturing process of a photo attack

a certain loss of information occurs and also peculiar noise is introduced .

- Video analysis, that tries to detect, as humans do, facial physiological clues like blinks, mouth movements and changes in facial expression.

In this paper, the proposed approach focused in face detection using Viola-Jones algorithm [7] and Active Shape Models with Stasm [8]. We used Viola-Jones algorithm to detect the face and Stasm to locate the eyes. The difference in face detection between our approach and [9], is that in ours, we use Stasm to locate the eyes and in [9] they used eye detection algorithm (aligned by the nose and the eyes). Then we applied the LBP operator to extract the features and we used a nonlinear SVM classifier to determine if the input image is real or not.

The rest of the paper is organized as follows: SectionII discusses related works on anti spoofing attacks. SectionIII presents the Database used in our tests. SectionIV describes our approach in details. The experimental results and a comparison with many related works are summarized in SectionV. Finally a conclusion and future works are given in sectionVI.

## II. RELATED WORK

Anti-spoofing for 2-D face recognition systems can be coarsely classified in 3 categories with respect to the clues used for attack detection: motion, texture analysis and liveness detection [10].

The first one interests in detecting clues generated when two dimensional counterfeits are presented to the system, for example photos or video clips [11]. Kollreider et al. [12] evaluated the trajectories of selected part of the face from a short sequence of images using a simplified optical flow analysis followed by a heuristic classifier. The same authors in [13] introduced a method to fuse these scores with liveness properties such as eye-blinks or mouth movements. Bao et al. [14] proposed the detection of attacks produced with planar media using optical flow based on motion estimation.

Exploring the input image data is to take advantage of texture patterns that may look unnatural by counter measures. Li et al. [15], used a Fourier spectra to compare the hard-copies of client faces and real accesses. Li et al. method works well for down-sampled of the print-photo attack identity, but it is fail for higher-quality sometimes.



Liveness detection tries to capture signs of life from user images by analyzing spontaneous movements that cannot be detected in photographs, such as eye blinking. The authors in [2] and [16] brought a real-time liveness detection specifically against photo-spoofing using spontaneous eye-blinks, which are supposed to occur once every 2-4 seconds in humans.

Juka et al. [17] proposed an approach based on learning texture features from single images using LBP, Gabor wavelet and Histogram of Oriented Gradients (HOG). In [18], the same authors presented a novel approach based on analyzing the texture of the facial images using Multi-Scale Local Binary Patterns (MSLBP), Local Phase Quantization (LPQ) and Gabor wavelets. Tiago et al. [6] proposed a novel countermeasure against face spoofing. This approach uses an operator called Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) that combines space and time information into a single descriptor with a multi-resolution strategy. In [19], the same authors analyzed three recently published countermeasures (Correlation with frame differences, LBP countermeasure, LBP-TOP countermeasure). Chingovska et al. [20] inspected the potential of texture features based on LBP and its variations on three types of attacks: printed photographs, photos and videos. William et al. [4] introduced an anti-spoofing solution based on a set of low-level feature descriptors exploring both spatial and temporal information using Partial Least Squares (PLS).

In this paper, we propose a new approach based on face detection using Viola-Jones algorithm for detection the face and Active Shape Models with Stasm for locating Landmark. These Landmarks help us to crop the essential part of the face, then we apply LBP operator to extract features and then SVM classifier.

### III. DATABASE AND PROTOCOL

In our work, we used the publicly available NUAA Photograph Imposter Database. The NUAA database, proposed by Tan et al. [9] comprises images extracted from videos of 15 subjects captured in three sections and contains attempts of attack based on hand-held printed photos. This dataset is divided into training and test sets. The former has 1743 live images and 1748 non-live, and the latter consists of 3362 live and 5761 non-live samples (c.f. Tab I).

TABLE I: Number of images in the training set and test set.

NUAA dataset		Session1	Session2	Session3	Total
Training Set	Client	889	854		1743
	Imposter	855	893		1748
Test Set	Client	0	0	3362	3362
	Imposter	0	0	5761	3362
Total		1744	1747	9123	12614

The data provided in this dataset consists of grayscale face images cropped using the Viola-Jones detector, and normalized to 64 x 64 pixels aligned by the nose and the eyes. Tan et al. [9] used the normalized images in their experiments to perform a direct comparison with their results. We used in our test the same set of images (c.f. fig 1) and also images that we've cropped using our method explained in ( section 4-c).



Fig. 1: . Illustration of the samples from the database. In each column (from top to bottom) samples are respectively from session 1, session 2 and session 3. In each row, the left pair are from a live human and the right from a photo.

### IV. THE PROPOSED APPROACH

Using the human face as a key to security, biometric face recognition technology has received significant attention in the past several years. So there is big problem when a person showing in front in camera the prints photo because it can look very similar to the images captured from live faces (cf. Fig. 1).

In this section, we explain our approach of anti-spoofing used to differentiate between live faces and fake ones. Fig.2 shows the block diagram of our anti-spoofing approach. The proposed method adopts face detection using Viola-Jones algorithm [7], and Active Shape Models with Stasm [8]. Viola-Jones algorithm is used to get all the face needed for Stasm, which locates landmarks. Then we used the coordinates of eyes to rotate and crop the face. In each bloc 3x3 of cropped image, we applied LBP operator [21], for describing the micro-textures. Each vector in its own transformed feature space is then fed to a non-linear SVM classifier. The individual SVM outputs determine whether there is a real face or a fake one in front of the camera. We describe below each step in detail.

#### A. Viola-Jones algorithm

The Viola and Jones algorithm is a method for detecting an object in a digital image, proposed by Paul Viola and Michael Jones in 2001 [7]. Originally invented to detect faces, it may also be used to detect other types of objects such as cars or aircraft.

We use Viola-Jones in our approach to detect the face. The question here is why we did not use Stasm directly. When we used Stasm directly in large images with small face we cannot find the correct face, so, we must apply Viola-Jones to detect the face first (c.f 2.a), because Viola is good in face detection with any pictures has faces.

#### B. Active Shape Models with Stasm

Stasm is a software package for locating landmarks using Active Shape Models (ASMs). The package comes preconfigured for locating landmarks in faces. We need Stasm only for localization of eyes (Coordinates). To rotate the face, we must depend on coordinates of eyes (c.f fig.2.b).

#### C. Crop and normalize the face

For adjust and crop the face, we need to calculate the distance between the two eyes (distance A)(c.f fig.2.c). So to

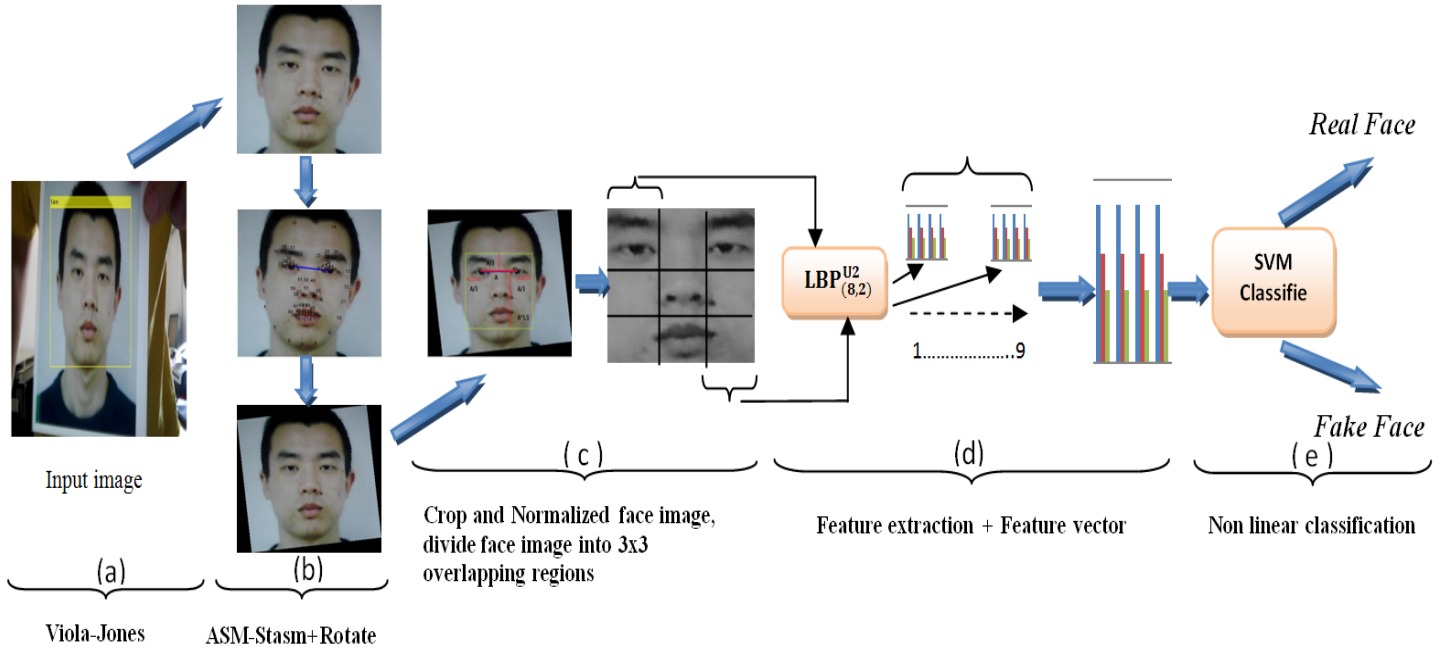


Fig. 2: The proposed approach : (a) Viola-Jones algorithm, (b) Active Shape Models with Stasm, (c) Crop and normalize the face, (d) Feature extraction using LBP and (e) Non-linear SVM classifier for determining a real face or fake.

crop the face, we used the distance  $A$ . Then, we normalized the cropped face into a  $(64 \times 64)$  pixel images. In normalized face image, we used the histogram equalization, for adjusting image intensities to enhance contrast. After that, we added noise (salt and pepper) and applied median filter on the image to denoise it. In order to extract the local features of the face image, we divided it into  $3 \times 3$  overlapping regions and applied the LBP on each bloc.

#### D. Feature extraction using LBP

The LBP operator is an image operator, which transforms an image into an array or image with more detail. The basic LBP, introduced by Ojala et al. [21], was based on the assumption that texture has locally two complementary aspects, a pattern and its strength.

The original LBP works in a  $3 \times 3$  pixel block of image. The pixels in this block are threshold by its center pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel. As the neighborhood consists of 8 pixels, a total of  $2^8=256$  different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood [22]. Fig 3 illustrates the principle of basic LBP operator.

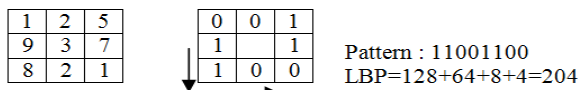


Fig. 3: The basic LBP operator.

The LBP operator used a circular neighborhood. The notation  $(P, R)$  is generally used for pixel neighborhoods to

refer to sampling points and circle of radius. So the calculation of the  $LBP_{P,R}$  codes can be easily done. The value of the LBP code of a pixel  $(x_c, y_c)$  is given by[22]:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (1)$$

where  $g_c$  corresponds to the gray value of the center pixel  $(x_c, y_c)$ ,  $g_p$  refers to gray values of  $P$  equally spaced pixels on a circle of radius  $R$ , and  $s$  defines a thresholding function as follows:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Another extension of the original LBP called uniform patterns where a uniformity measure of a pattern is used:  $U$  (pattern) is the number of bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered circular.

When, divided the face image into  $3 \times 3$  overlapping regions, we used  $LBP_{(8,2)}^{U2}$  operator on each region. The local 59-bin histograms from each region are computed and collected into a single 531-bin histogram.

#### E. Classification

A Support Vector Machine (SVM) performs classification by finding the hyper plane that maximizes the margin between two classes. The vectors (cases) that define the hyper plane called the support vectors.

In our experiments, once the enhanced histograms are computed, we use a nonlinear SVM classifier [23] with radial

basis function kernel for determining whether the input image corresponds to a live face or not. The SVM classifier is first trained using a set of positive (real faces) and negative (fake faces) samples.

## V. EXPERIMENTAL ANALYSIS

We evaluated the proposed approach on the NUAA Photograph Imposter Database [9]. In our experiments, we used Matlab2013b, beginning with Viola-Jones algorithm to locate all components of the Face images. Using Stasm on detected face image, the eyes are localized. The coordinates of the eyes are used to adjust, and then to crop the face as explain before (c.f section IV-C). All cropped faces are resized to a consistent size 64x64.

We also divided the normalized faces in 9 block with overlapping algorithm before applying  $LBP_{(8,2)}^{U2}$  to extract the local features in each region of the image. In this step, we computed the histogram of each block to get 59 bin histograms. We collected then these histograms in a simple one of a 531 bin. For classification, we used SVM classification.

We applied our approach using face detection without Stasm using the same image normalization in NUAA data bases in one hand. In other hand, we calculated the results using the Viola Jones algorithm and Active Shape Model with Stasm. Also, for 107 images not detected by Stasm, we have used a manual detection by manual calculation of coordinates of eyes needed to crop the face.

We compared our results with those of the state of art : LBP+Gabor+HOG [3], LBP [5], LPQ [5], Bad Illumination Conditions [24]. For fair comparison, we used the same protocol with other authors: 1743 live images, 1748 non-live, for train and 3362 live and 5761 non-live samples for test.

The performance of the three detection (without Stasm, with Stasm, and manual ) with texture operators LBP in terms of Receiver Operating Characteristic (ROC) curves and Detection Error Tradeoff (DET) curve are shown in (Fig.4 and Fig. 5).

From the results, we can notice that the three descriptors performed quite well. The equal error rates (EER), shown in Table II, indicates that the detection with Stasm (EER= 2.4) is better than without Stasm (EER= 3.9). For the manual detection of 107 image not detected by Stasm (EER= 0.6) gives best results. So, we have to develop an algorithm to detect automatically any coordinate of eyes.

TABLE II: Performance comparison between our proposed approach and the best results in [17], [18], [24] on the same database and using the same protocol.

Methods	Accuracy %	EER	AUC
Bad Illumination,Conditions [24]	93	8.2	-
LBP overlapping,blocks [18]	-	2.9	0.99
LBP+Gabor+HOG,[17]	98	1.1	0.999
Without_stasm	97.31	3.9	0.9930
With_stasm	98.41	2.4	0.9975
With_stasm(correction manual)	99.61	0.6	0.9998

## VI. CONCLUSION AND FUTUR WORK

Face biometric systems are vulnerable to spoofing attacks and photographs are the most common sources of spoofing attacks. Indeed, face prints usually contain printing quality defects that can be well detected. The surface properties of real faces and prints are also different.

We proposed in this work, an approach for anti-spoofing detection based on Active Shape Models with Stasm and LBP that discriminate live faces from fake ones.

Our approach tested on NUAA Photograph Imposter Database witch contains several real and fake faces showed excellent results compared to many previous works. On future work we will try to test our approach with another data base and to find a method replacing the manual method for locating landmarks on images which have hard position.

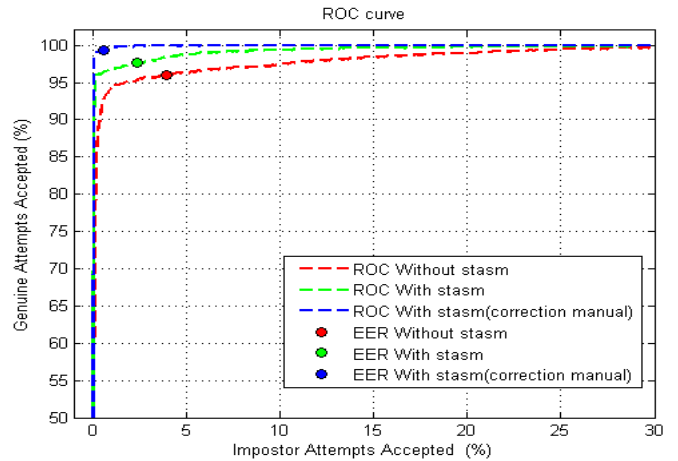


Fig. 4: Performance (ROC curves) of the proposed approach without Stasm,with Stasm, and manual.

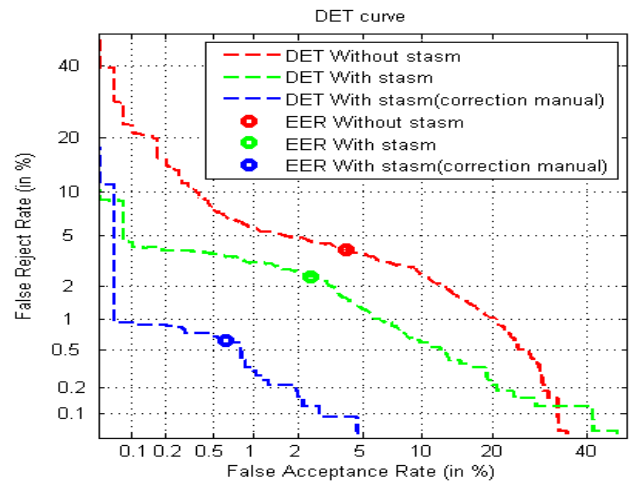


Fig. 5: Performance (DET curves) of the proposed approach without Stasm, with Stasm, and manual.



Fig. 6: Examples of anti-spoofing classification (Real: blue, Fake: yellow and wrong classification: red)

## REFERENCES

- [1] F. L. Podio, "Biometrics technologies for highly secure personal authentication," *National Institute of Standards and Technology*, <http://whitepapers.zdnet.com/search.aspx>, 2001.
- [2] G. Pan, Z. Wu, and L. Sun, "Liveness detection for face recognition," *Recent advances in face recognition*, pp. 109–124, 2008.
- [3] B. Toth and U. C. von Seelen, "Liveness detection for iris recognition," in *The Presentation Sheet of NIST Workshop, Biometrics and E-Authentication over Open Networks*, 2005.
- [4] W. R. Schwartz, A. Rocha, and H. Pedrini, "Face spoofing detection through partial least squares and low-level descriptors," in *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011, pp. 1–8.
- [5] R. Tronci, D. Muntoni, G. Fadda, M. Pili, N. Sirena, G. Murgia, M. Ristori, and F. Roli, "Fusion of multiple clues for photo-attack detection in face recognition systems," in *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011, pp. 1–6.
- [6] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, "Lbp-top based countermeasure against face spoofing attacks," in *Computer Vision-ACCV 2012 Workshops*. Springer, 2013, pp. 121–132.
- [7] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1. IEEE, 2001, pp. 1–511.
- [8] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *Computer Vision-ECCV 2008*. Springer, 2008, pp. 504–513.
- [9] X. Tan, Y. Li, J. Liu, and L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model," in *Computer Vision-ECCV 2010*. Springer, 2010, pp. 504–517.
- [10] A. Anjos and S. Marcel, "Counter-measures to photo attacks in face recognition: a public database and a baseline," in *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011, pp. 1–7.
- [11] O. Kahm and N. Damer, "2d face liveness detection: An overview," in *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the*. IEEE, 2012, pp. 1–12.
- [12] K. Kollreider, H. Fronthaler, and J. Bigun, "Evaluating liveness by face images and the structure tensor," in *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*. IEEE, 2005, pp. 75–80.
- [13] K. Klaus, H. Fronthaler, and J. Bigun, "Verifying liveness by multiple experts in face biometrics," in *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*. IEEE, 2008, pp. 1–6.
- [14] W. Bao, H. Li, N. Li, and W. Jiang, "A liveness detection method for face recognition based on optical flow field," in *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on*. IEEE, 2009, pp. 233–236.
- [15] J. Li, Y. Wang, T. Tan, and A. K. Jain, "Live face detection based on the analysis of fourier spectra," in *Defense and Security*. International Society for Optics and Photonics, 2004, pp. 296–303.
- [16] G. Pan, L. Sun, Z. Wu, and S. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcam," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.
- [17] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using texture and local shape analysis," *IET biometrics*, vol. 1, no. 1, pp. 3–10, 2012.
- [18] J. Maatta, A. Hadid, and M. Pietikainen, "Face spoofing detection from single images using micro-texture analysis," in *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011, pp. 1–7.
- [19] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, "Can face anti-spoofing countermeasures work in a real world scenario?" in *Biometrics (ICB), 2013 International Conference on*. IEEE, 2013, pp. 1–8.
- [20] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the*. IEEE, 2012, pp. 1–7.
- [21] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.
- [22] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, *Computer vision using local binary patterns*. Springer, 2011, vol. 40.
- [23] M. A. Hearst, S. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *Intelligent Systems and their Applications, IEEE*, vol. 13, no. 4, pp. 18–28, 1998.
- [24] B. Peixoto, C. Michelassi, and A. Rocha, "Face liveness detection under bad illumination conditions," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 3557–3560.

# Head Pose Estimation from Depth Map

AsmaManel KHERCHI<sup>1</sup>, Slimane LARABI<sup>2</sup>

<sup>1</sup>*Cerist, Ben Aknoun– Alger*

<sup>2</sup>*USTHB University, BP 32, EL ALIA, Alger*

<sup>1</sup>k.asmamanel@gmail.com

<sup>2</sup>slarabi@usthb.dz

**Abstract**— In this paper we propose a new method for head pose estimation using the depth sensor Kinect. Our approach infers the head pose based on the symmetry or asymmetry computed on the depth map of the face. This approach does not require the location of the nose or any other feature on the face such as has been done in many works, but uses only the depth map of the face. Two features are proposed for characterizing the pan, roll and tilt rotation of the head. The first one, measures the area of nearest region of the face relatively to the face area. The second one, it concerns the symmetry on the depth map of the face. Experiments are conducted on our acquired data. The obtained results are promising and demonstrate the usefulness of the proposed features.

**Keywords**— Depth-Map-Head pose estimation-Symmetry-Kinect

## I. INTRODUCTION

Depth estimation for faces or body is an important problem that has been largely studied in many computer vision applications such as face recognition, face and gesture recognition, face animation and analysis. This interest is due to the invariance of the depth map to illumination changes in the scene and thus the accuracy of head or body pose estimation is guaranteed.

In this paper we propose a new method for head pose estimation using the depth sensor Kinect. Our approach is based on the symmetry or asymmetry computed on the depth map of the face. Two features are proposed for characterizing the pan, roll and tilt rotation of the head. The first one concerns the relative area of nearest region of the face to the Kinect. The second one concerns the asymmetry on the depth map of the face. We will see that relatively to the computed axis located along the head, the asymmetry is in relationship with the pan angle. Our approach does not need the location of the nose or any other feature on the face such as it has been done in many works, but uses only the depth map of the face. In addition, we assume that face is located using the image and applying Viola and Jones detector.

Experiments are conducted on our acquired data. The obtained results demonstrate the usefulness of the proposed features. In the next section, a review of relevant works to this topic is given. Section 3 is devoted to our approach. First, we explain our basic principle; the algorithm is given. The final section contains the conducted tests and the obtained results.

## II. PREVIOUS WORKS

Head pose estimation, which consists to estimate the three rotation angles (Pan, Tilt, Roll) of the head (see figure 1), has been well studied but remains one of the open problems of computer vision applications due to the external environment which influences the image quality. Many states of the art were proposed where we can see the difficulty of this problem [2].

To improve the performances of head pose estimation, the depth data that may be acquired by stereovision or the depth camera (Kinect) has been added and some approaches are proposed [3, 4, 5, 6, 7].

Due to the availability of depth-sensing technologies, many works were proposed by using the depth for solving the problem of head pose estimation.

Many assumptions are made: the nose is visible [8], the head is the only object present in the field of view [2], and the 2D image data is combined with the depth [9], [10].

In the recent work [1], authors deal with depth images where other parts of the body might be visible and therefore need to discriminate which image patches belong to the head and which don't using Discriminative Random Regression Forests. In other recent work [11], a reference depth image of a human subject is obtained. The method searches the 6-dimensional pose space to find a pose from which the head appears identical to the reference view. This search is formulated as an optimization problem whose objective function quantifies the discrepancy of the depth measurements between the hypothesized views to the reference view.

Instead of other methods, we estimate head pose directly using some features extracted from the depth map without any constraint on the scene. In addition, our method deals with any head pose and there is no constraint for pan, roll or tilt rotations such as it has been assumed in the literature.

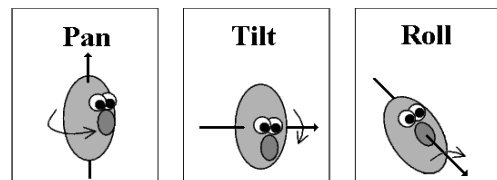


Fig. 1. The three motions defining the head pose.

### III. OUR APPROACH

#### A. Basic principle

Our aim is to estimate the three angles (Pan, Tilt and Roll) of the head from the depth map. Once the face is located on the image acquired by the Kinect by using the Viola and Jones detector [12], the depth map associated to head is divided into a set of plans  $\Pi_i$  of depths. Each plan  $\Pi_i$  is defined as the set of pixels in the depth map having a value of depth in a specific interval  $I_i = [d_{min}^i, d_{max}^i[$ . The values of  $d_{min}^i, d_{max}^i$  are deduced for a given number  $N_p$  of intervals (which is fixed by user) and knowing the maximum and minimum depths of face's pixels.

Whatever the position of the head relatively to the sensor Kinect, for  $N_p=6$ , the distance between minimal and maximal values of depths is divided into six equidistant intervals. Figure 2 illustrates the depth map with six plans colored, the closest plan is colored in red and the far one in grey.

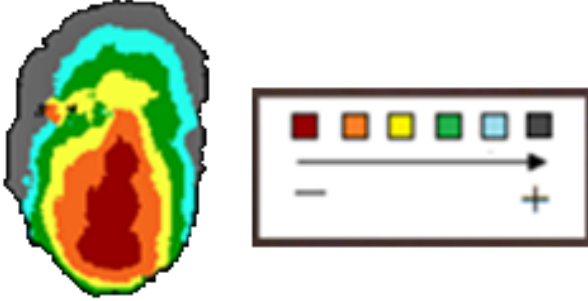


Fig. 2. Computed plans on the depth map.

A set of features are extracted from the depth map associated to the face. **The first feature  $f_1$**  concerns the area of the nearest region of the face to the sensor Kinect. The more the pan angle increases, the more the area of the nearest region increases compared to the face area. Indeed, due to the geometry of face, if the head is in front to the Kinect, the region of nose and lips is the nearest and occupies less pixels in map depth. However, when the head performs pan motion, the nearest area will concerns the cheek part and becomes largest. Figure 3 illustrates an example of three motions of pan where the nearest region colored by red color changes and follows the head motion.

Let  $S_i$  be the area of the regions of the same depth (Plan  $\Pi_i$ ,  $i=1..6$ ). Let  $S_1$  be the area of regions of first the plan (nearest plan to the Kinect). The feature  $f_1$  is defined by equation 1:

$$f_1 = \frac{S_1}{\sum_{i=1}^{N_p} S_i} \dots (1), \text{ where } N_p \text{ is the number of plans.}$$

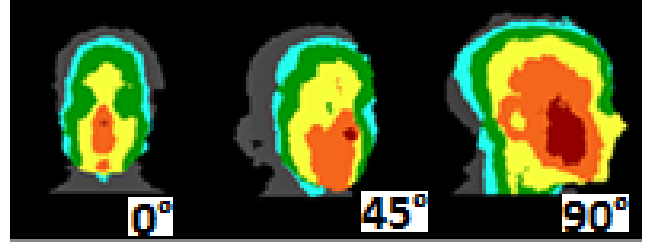


Fig.3. Depth maps of head which performs three Pan's motions, where the color red (resp. grey) indicates the nearest (resp. far) region.

This first feature characterizes also the Pan combined to the Tilt motion. Indeed, the tilt motion allows the displacement of the nearest region relatively to the bounding box encompassing the face from the center (Tilt equal zero) to the high part in case of inclination of the head towards the low, inversely in case of inclination of the head towards the high (see figure 4).

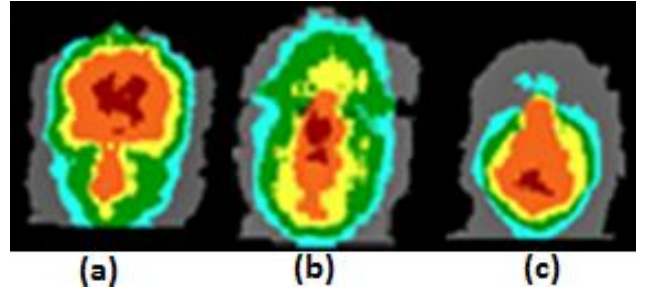


Fig. 4. Depth maps head performing tilt motion.

In order to find the correspondence between the ratios  $N_r$  and the values of Pan, Tilt angles, a set of depth maps were acquired related to these motions (see experiment section).

**The second feature  $f_2$**  concerns the symmetry on the depth map of the face. Firstly, the minimum rectangle encompassing the face is located and the lines (L), (W) which pass by the center of this rectangle, respectively parallel to its length and width are considered as the symmetry axes (see figure 5). When head is in front of the camera (pan motion is equal to zero), there is symmetry between regions of the same depths relatively to the axis (L). This property is also valid for front pose even if head performs a roll motion (see figure 5).

Let  $(S_i^l, S_i^r)$  be the areas of the pair of regions (Left and Right) of the same depth (of the plan  $\Pi_i$ ). We define the asymmetrical ratio  $f_2$  by equation 2:

$$f_2 = \sum_{i=1}^{N_p} \frac{|S_i^l - S_i^r|}{S_i^l + S_i^r} \dots (2)$$

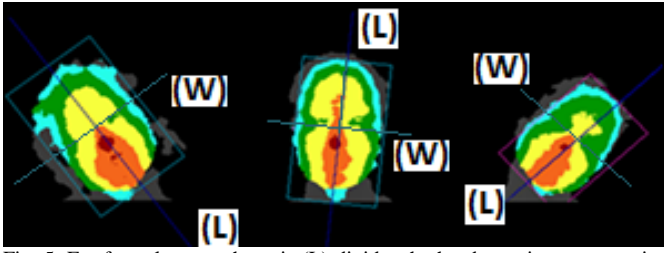


Fig. 5. For frontal poses, the axis (L) divides the head area into symmetrical regions whatever the roll rotation.

Where  $N_p = 6$  is the number of regions of different depths. The value of the ratio  $f_2$  is low when head is in front to the Kinect(theoretically equal zero). However, the more the pan rotation increases, the more  $f_2$  is high. The variation of this value when the head is performing a pan rotation for different subjects is illustrated by the graph of figure 6.

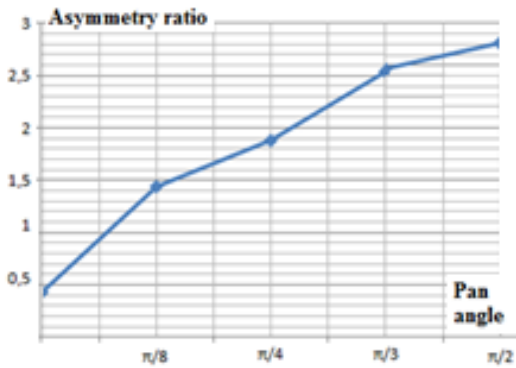


Fig. 6. Asymmetry variation for pan rotation.

The value of  $f_2$  characterizes different motions. For frontal poses (pan angle is equal to zero), whatever the roll angle this ratio is very low (see figure 5). For non-frontal poses, this value increases and is related to pan angle.

The **third feature** concerns the position of nearest region relatively to the axis (W). This feature characterizes the tilt motion. For frontal poses, it is easy to infer the motion tilt (see figure 4). However, for non-frontal poses, the orientation of the axis (L) and the position of nearest region are used conjointly to estimate the tilt motion (see figure 6).

The **fourth feature** concerns the orientation of the axis (L). For frontal poses, the orientation of this axis indicates the roll motion. Figure 5 illustrates this case where symmetry is verified relatively to the axis (L).

### B. Algorithm

The following algorithm gives the main steps for head pose estimation.

#### Algorithm

##### Begin

##### 1- Step learning

For a subset of poses of the data set chosen randomly, where all poses are acquired at least one time, estimate the relationship  $f_1(\text{Pan angle})$

2- Locate the depth map of face in image and compute the bounding box encompassing it and the two axes of symmetry (L), (W).

-The pan and tilt angles are estimated by analysing the values of computed features  $f_1$  and  $f_2$ . Depending on the value of  $f_1$ , the frontal pose or non-frontal pose is distinguished.

If the pose is frontal, depending on the value of position of nearest region relatively to the axis (L), the tilt motion is estimated.

-For non-frontal poses, the asymmetry ratio  $f_2$  allows estimating the pan motion. The position of nearest region relatively to the axis (W) and the orientation of (L) allow the estimation of tilt motion.

-The roll angle is computed as the orientation of the symmetry axis (L) in case of frontal pose. In case of non-frontal poses, the orientation of the axis (L) and the position of nearest region are used conjointly to estimate the tilt motion

End.

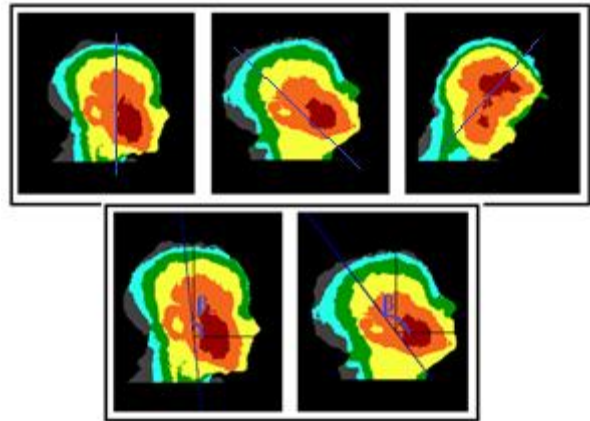


Fig. 6. Head performing pan and tilt rotation.

## IV. EXPERIMENTAL RESULTS

### A. Data set used

In order to evaluate the efficiency of the proposed method, we applied it on acquired data with a Kinect sensor. People were recorded while turning their heads, sitting in front of the sensor, at roughly one meter of distance. The ground truth data corresponding to pose of head are determined using the nose direction and orientation.

Once the face is located, we apply the masque to the depth map. The map of depths associated to heads are extracted and the pixels are colored with 6 colors depending on their depths (see figure 7).

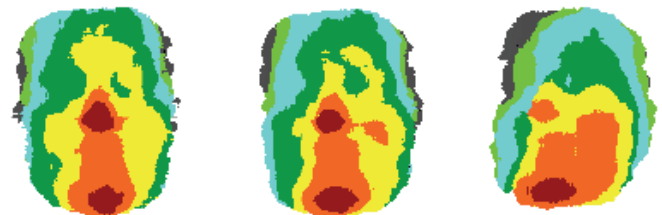


Fig. 7. The depth maps associated to the heads partitioned into 6 plans.

**B. Conducted tests**

The first tests concern some acquired poses of different persons that have been used in order to determine the correspondence between the different values of the two features  $f_1$ ,  $f_2$  and the pan angles.

The symmetry axes are computed and the values of  $f_1$  and  $f_2$  are determined. Figure 8 illustrates the located plans and the axes of symmetry. The values of the feature  $f_1$  are given by figures 9, 10 and 11.

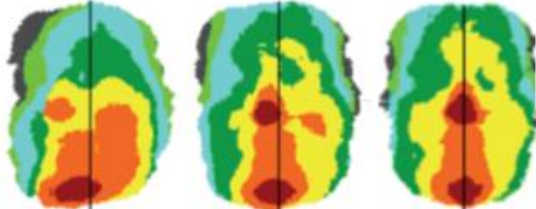


Fig. 8. The computed axes of symmetry (L).

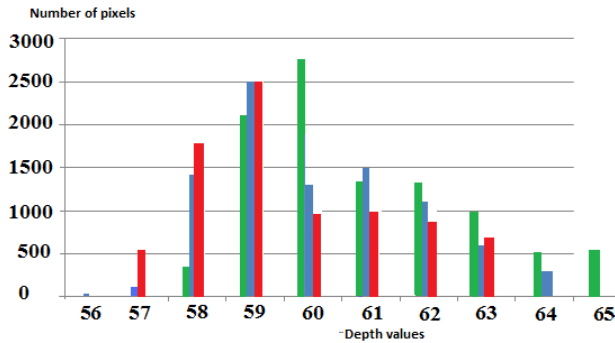


Fig.9. Histogram of depth values for different poses where the colour red, blue, green correspond respectively to pose face, profil 45° and profile 90°

The histogram of figure 10 illustrates the area all depths when head is in face, profile poses. We can see that the area of nearest region grows when head is rotating from face pose toward profile pose 90°.

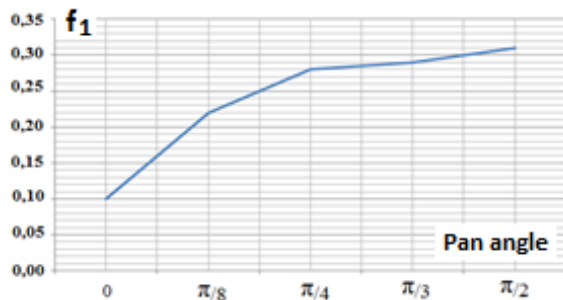


Fig.10. Values of  $f_1$  for different values of pan angle

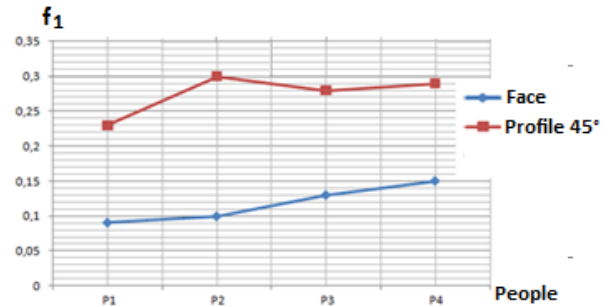


Fig.11. Values of  $f_1$  for different persons and for two values of pan angle

After this, some images were acquired and the poses were estimated using the four features. Figures 12, 13 illustrate the images of the person performing combined motion and the located regions in the depth maps. The estimated poses coincide with the ground truth data. The estimated error does not exceed 2 degrees.

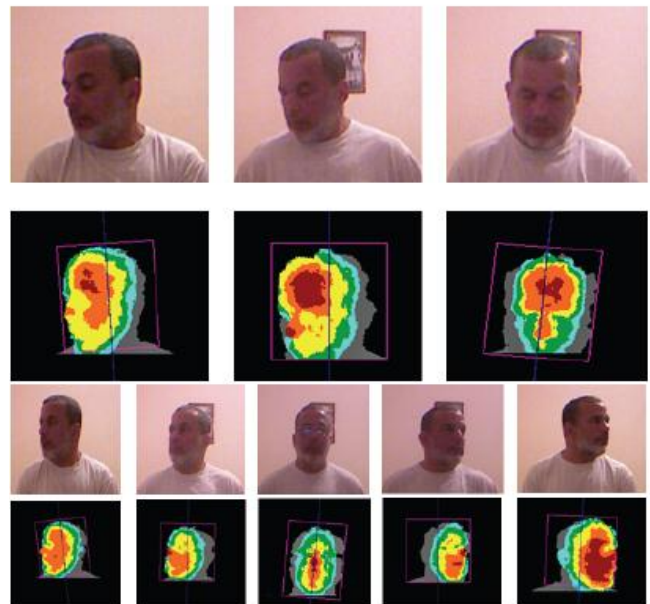


Fig. 12. Some obtained intermediate results (depth maps).

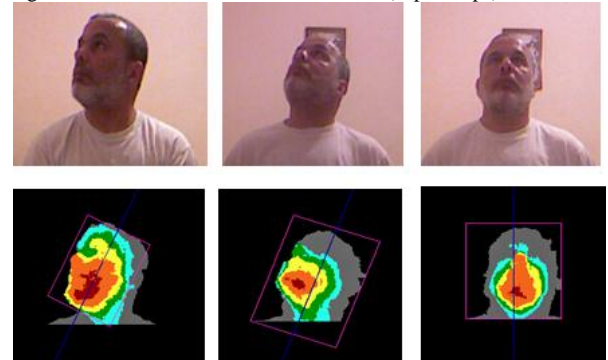


Fig. 13. Some obtained intermediate results (depth maps).

The obtained results demonstrated that symmetry is a good feature for head pose estimation. The next step of this work is the comparison with the ground truth data of BiwiKinect Head Pose Database [1] in order to study its accuracy.



## IV CONCLUSION

We presented a new method for head pose estimation based on the depth map. We exploited the symmetry/asymmetry on the depth map to deal with pan and tilt motions. The orientation of the symmetry axis of the bounding box indicates the roll angle of the head. Two main features are extracted: the relative area of nearest region and the asymmetry ratio. These features are used to infer the motion angle (pan and tilt). The advantage of this method is the low cost of computation and the absence of constraints concerning the limits of head motion or the presence of special cues on face.

We applied our approach of data set acquired in our laboratory. The obtained results are promising.

## REFERENCES

- [1] G. Fanelli, M. Dantone, J. Gall, A. Fossati, L. Gool, Random Forests for Real Time 3D Face Analysis, *Int. J. Comput. Vision*, 101(3), 2013, pp. 437-458.
- [2] E. Chutorian, M. Trivedi, Head pose estimation in computer vision: A survey, *PAMI*, 31(4), 2009, pp. 607-626.
- [3] M. D. Breitenstein, D. Kuettel, T. Weise, L. Van Gool, H. Pfister, Real-time face pose estimation from single range images, *CVPR*, pp. 1-8, 2008.
- [4] S. Malassiotis, M. Srinivas, Robust real-time 3D head pose estimation from range data, *Pattern Recognition*, vol. 38, 2005, pp. 1153-1165.
- [5] R. Yang and Z. Zhang, Model-based head pose tracking with stereovision, *Aut. Face and Gesture Rec.*, pp. 255-260, 2002
- [6] N. Pears, T. Heseltine, M. Romero, From 3D point clouds to pose-normalised depth maps, *Int. J. Comput. Vision*, 89 (2-3), 2010, pp. 152-176.
- [7] F. Kondori, S. Yousefi, H. Li, S. Sonning, S. Sonning, 3D head pose estimation using the kinect, *WCSP*, pp. 1-4, 2011.
- [8] G. Fanelli, J. Gall, L. Van Gool, Real time head pose estimation with random regression forests, *CVPR*, pp. 617-624, 2010.
- [9] Q. Cai, D. Gallup, C. Zhang, Z. Zhang, 3d deformable face tracking with a commodity depth camera, *ECCV 2010*, pp. 229-242.
- [11] P. Paderelis, X. Zabulis, A. A. Argyros, Head pose estimation on depth data based on Particle Swarm Optimization, *Workshop on Human Activity Understanding from 3D Data (HAU3D'2012)*, pp. 1-4, 2012.
- [10] E. Seemann, K. Nickel, R. Stiefelhagen, Head pose estimation using stereo vision for human-robot interaction, *Aut. Face and Gesture Rec.*, pp. 626-631, 2004.
- [12] P. Viola, M. Jones, Robust real-time face detection}, *Int. J. Comput. Vision*, 57(2), 2004, pp. 137-154.

# A Novel Face Feature Extraction Method Based on WPD & LBP & Bi-dimensional Subspace Analysis

Benouis Mohamed<sup>1</sup>, Benkaddour Mohamed Kamel<sup>2</sup>, Senouci Mohamed<sup>1</sup>,

<sup>1</sup> University of Oran, Algeria

<sup>2</sup>University Sidi Bel Abbes, Algeria

<sup>2</sup> INTTIC – Oran, Algeria

<sup>1</sup> mhbenouis@yahoo.com.

<sup>2</sup> kamelcom@live.com

<sup>1</sup> msenouci@yahoo.com

**Abstract**— Face recognition is considered as one of the best biometric methods used for human identification and verification, this is because of its unique features that differ from one person to another, and its importance in the security field. This paper proposes an algorithm for face recognition and classification using a system based on WPD, Local Binary Pattern and two-dimensional subspace for feature extraction, and Combined Learning Vector Quantization and PNN Classifier as Neural Network approach for classification. We propose a new feature extraction algorithm based on WPD with LBP to speed up the feature extraction step. In order to capture the more important information that is contained in a face with a few LBP features, we use a bi-dimensional principal component analysis. The proposed approach is tested on ORL and FEI face databases. Experimental results on these databases demonstrated the effectiveness of the proposed approach for face recognition with high accuracy compared with previous methods.

**Keywords**—biometric, face recognition, 2DPCA, 2DLDA, DWT, PNN, WPD, LBP, LVQ

## I. INTRODUCTION

The security of persons, goods or information is one of the major concerns of the modern societies. Face recognition is one of the most commonly used solutions to perform automatic identification of persons. However, automatic face recognition should consider several factors that contribute to the complexity of this task such as the occultation, changes in lighting, pose, expression and structural components (hair, beard, glasses, etc.) [1].

Several techniques have been proposed in the past in order to solve face recognition problems. Each of them evidently has their strengths and weaknesses, which, in most of the cases, depend on the conditions of acquiring information. Recently, several efforts and research in this domain have been done in order to increase the performance of the recognition, such as support vector machine (SVM), Markov hidden model (HMM), probabilistic methods (Bayesian networks) and artificial neural networks. This latter has attracted researchers because of its

effectiveness in detection and classification of shapes, which has been adopted in new face recognition systems [2].

## II. FACE RECOGNITION SYSTEM

A face recognition system is a system used for the identification and verification of individuals, which checks if a person belongs to the system's database, and identifies him/her if this is the case.

The methods used in face recognition based on 2D images are divided into three categories: global, local and hybrid methods.

- Local or analytical facial features approaches. This type consists on applying transformations in specific locations of the image, most frequently around the features points (corners of the eyes, mouth, nose,). They therefore require a prior knowledge of the images...
- Global approaches use the entire surface of the face as a source of information without considering the local characteristics such as eyes, mouth, etc.
- Hybrid methods associate the advantages of global and local methods by combining the detection of geometrical characteristics (or structural) with the extraction of local appearance characteristics.

This article is organized as follows: Basic notions concerning Two-dimensional subspace, wavelet transform theory are provided in Section 2. LBP features are presented in section 3. Feature vectors results from two-dimensional subspaces is applied to a Combined LVQ and PNN classifier are described in Section 4. Section 5 provides face recognition system based on PNN, LVQ, the experimental results and Comparison between the several types of features obtained using WPD, DWT, LBP, 2DPCA and 2DLDA. A comparison with other approaches is also done in section 6. Conclusion and future works are presented in Section 7.

### A. Two-dimensional principal component approach analysis (2DPCA)

Proposed by Yang in 2004 [4], 2DPCA is a method of feature extraction and dimensionality reduction based on Principal Component Analysis (PCA) that deals directly with face images as matrices without having to turn them into vectors like as the traditional global approach.

#### 1) The steps of face recognition by 2DPCA

Considering a training set  $S$  of  $N$  face images, the idea of this technique is to project a matrix  $X$  of size  $(n \times m)$  via a linear transformation like that:

$$Y_i = X \cdot R_i \quad (1)$$

Where  $Y_i$  is the principal component vector of size  $(n \times 1)$ , and  $R_i$  is the base projection vector of size  $(m \times 1)$ . The optimal vector  $R_i$  of the projection is obtained by maximizing the total generalized variance criterion

$$J(R) = R^T \cdot G_t \cdot R \quad (2)$$

Where  $G_t$  is the covariance matrix of size  $(m \times m)$  given by:

$$G_t = \frac{1}{M} \sum_{j=1}^M (X_j - \bar{X})^T (X_j - \bar{X}) \quad (3)$$

With  $X_j$ : the  $j^{\text{th}}$  image of the training set

$\bar{X}$ : The average image of all the images in the training set.

$$\bar{X} = \frac{1}{M} \sum_{j=1}^M X_j \quad (4)$$

In general, one optimal projection axis is not enough. We must select a set of projection axes like:

$$\{R_1, R_2, \dots, R_d\} = \arg \max J(R) \quad (5)$$

$$R_i^T \cdot R_j = 0, i \neq j, i, j = 1, \dots, d$$

These axes are the eigenvectors of the covariance matrix corresponding to the largest "d" eigenvalues. The extraction of characteristics of an image using 2DPCA is as follows

$$Y_k = X \cdot R_k ; k=1, \dots, d \quad (6)$$

Where  $[R_1, R_2, \dots, R_d]$  is the projection matrix and  $[Y_1, Y_2, \dots, Y_d]$  is the features matrix of the image  $X$ .

### B. the 2DLDA approach

In 2004, Li and Yuan [5] have proposed a new two-dimensional LDA approach. The main difference between 2DLDA and the classic LDA is in the data representation model. Classic LDA is based on the analysis of vectors, while the 2DLDA algorithm is based on the analysis of matrices.

#### 1) Face recognition using 2D LDA

Let  $X$  be a vector of the  $n$ -dimensional unitary columns. The main idea of this approach is to project the random image matrix of size  $(m \times n)$  on  $X$  by the following linear transformation:

$$Y_i = A_j X \quad (7)$$

$Y$ : the  $m$ -dimensional feature vector of the projected image  $A$ .

Let us suppose  $L$ : class numbers.

$M$ : The total number of training images

The training image is represented by a matrix  $m \times n$

$$A_j (j = 1, \dots, M)$$

$\bar{A}_i (i=1, \dots, L)$ : the mean of all classes

$N_i$ : Number of samples in each class

The optimal vector projection is selected as a matrix with orthonormal columns that maximizes the ratio of the determinant of the dispersion matrix of the projected inter-class images to the determinant of the dispersion matrix of the projected intra-class images;

$$J_{\text{FLD}}(W_{\text{opt}}) = \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|} \quad (8)$$

$$P_b = \text{trace}(S_b)$$

$$P_w = \text{trace}(S_w)$$

Where,  $S_b$ : The inter-class dispersion matrix,

$S_w$ : The intra-class dispersion matrix.

The criterion can be expressed by:

$$J(X) = \frac{X^T S_w X}{X^T S_b X} \quad (9)$$

Where  $X$ : unitary column vector.

The unitary vector  $X$  maximizing  $J(X)$  is called the optimal projection axis. The optimal projection is chosen when  $X_{\text{OPT}}$  maximizes the criterion, as the following equation:

$$X_{\text{OPT}} = \arg \max_X J(X) \quad (10)$$

If  $S_w$  is invertible, the solution of optimization is to solve the generalized eigenvalue problem.

$$S_b X_{\text{opt}} = \lambda S_w X_{\text{opt}} \quad (11)$$

Like that  $\lambda$  is the maximum Eigenvalue of  $S_w^{-1} S_b$

In general, it is not enough to have only one optimal projection axis. We need to select a set of projection axes,  $x_1, x_2, \dots, x_d$  under the following constraints:

$$\{x_1, x_2, \dots, x_d\} = \arg \max_X J(X)$$

$$X_i^T X_j = 0, i \neq j, i, j = 1, 2, \dots, d$$

Indeed, the optimal projection axes  $x_1, x_2, \dots, x_d$  are orthonormal eigenvectors of  $S_w^{-1} S_b$  corresponding to the best first "d" eigenvalues permitting to create a new projection matrix  $X$ , which is a matrix of size  $n \times d$ :  $X = [x_1, x_2, \dots, x_d]$

We will use the 2DLDA optimal projection vectors  $x_1, x_2, \dots, x_d$  to extract the image features; we use the equation (07)

### III. DISCRETE WAVELET TRANSFORM (DWT)

Discrete wavelet transform (DWT) is a well-known signal processing field tool; it is widely used in feature extraction and compression and de-noising applications.

The discrete wavelet transform has been used in various face recognition studies. The main advantage of the wavelet transform over the Fourier transform is the time-scale location. Mallat [8] shows that the DWT may be implemented using a filters bank including a low-pass filter (PB) and a high-pass filter (PH).

Discrete Wavelet Package Decomposition (D-WPD) is a wavelet transform where signal is passed through more filters than the Discrete Wavelet Transform (DWT). In the DWT, each level is calculated by passing only the previous approximation coefficients through low and high pass filters. However in the D-WPD, both the detail and approximation coefficients are decomposed [9].

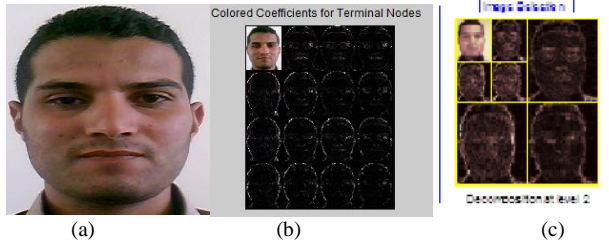


Fig 1. Wavelet decomposition at different levels  
(a) Original image  
(b) 2-level wavelet decomposition using WPD  
(c) 2-levels wavelet decomposition using DWT

#### IV. LOCAL BINARY PATTERN (LBP)

LBP descriptor computed using LBP operator introduced by Ojala et al. [9] is one of the widely used texture descriptors that have shown promising results in many applications [10],[11], J. Meng et al. [13] used it for face recognition, Lian and Lu and Sun et al. [10] employed it for gender recognition. The initial LBP operator associates a label with each pixel of an image; the label is obtained by converting each pixel value in the 3x3-neighbourhood of a pixel into a binary digit (0 or 1) using the center value as a threshold and concatenating the bits, as shown in Figure 1. Later the operator was extended to general neighborhood sizes, and its rotation invariant and uniform versions were introduced [12].

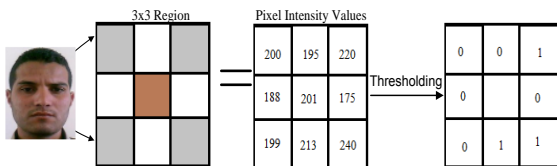


Fig 2: LBP Operator

The general LBP operator is denoted by  $LBP_{P,R}$  and is defined as follows:

$$LBP_{P,R} = \sum_{i=1}^{P-1} 2^i S(p_i - p_c) \quad (12)$$

where P is the total number of pixels in the neighborhood and R is its radius,  $p_c$  is the center pixel and the thresholding operation is defined as follows:

$$S(p_i - p_c) = \begin{cases} 1 & p_i - p_c \geq 0 \\ 0 & p_i - p_c < 0. \end{cases} \quad (13)$$

Commonly used neighborhoods are (8, 1), (8, 2), and (16, 2). The histogram of the labels is used as a texture descriptor. The histogram of labeled image  $f_l(x, y)$  is defined as:

$$H(i) = \sum_{x,y} I\{f_l(x, y) = i\}, \quad i = 0, \dots, n - 1 \quad (14)$$

Where  $n$  is the number of different labels produced by the LBP operator and

$$I\{x\} = \begin{cases} 1, & x \text{ is true} \\ 0, & x \text{ is false.} \end{cases} \quad (15)$$

Figure .3 shows the histogram extracted from an image with LBP operator.



Fig 3: LBP Histogram Calculation for Full image

General LBP operator has three parameters: circular neighborhood (P, R), rotation invariance (ri) and uniformity (u2). For a particular application, it is necessary to explore this parameter space to come up with the best combination of these parameters.

The most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused by illumination variations. Another important property is its computational simplicity [11], which makes it possible to analyze images in challenging real-time settings. The drawbacks in this system are it is highly sensitive to glasses and it is time-consuming process.

In this paper, LBP is introduced in order to extract the face features from the normalized face image based WPD . The output of LBP is the feature matrix with 2D-dimension used as a database of face which is applied two-dimensional subspace for reduction, discrimination and speed time.

#### V. FACE CLASSIFICATION USING NEURAL NETWORKS

Several studies have shown improved face recognition systems using a neural classification compared to classification based on Euclidean distance measure [15].

##### A. probabilistic neural networks:

The probability neural network is proposed by D. F. Specht for solving the problem of classification in 1988 [16]. The theoretical foundation is developed based on Bayes decision theory, and implemented in a feed-forward network architecture.

PNN represent mathematically by the following expression

$$a = \text{radbas}(\|IW - x\|b) \quad (16)$$

$$y = \text{compet}(LW\alpha) \quad (17)$$

The structure PNN :The PNN architecture consists of two layers[17][18]:

**The first layer** computes distances from input vector to the input weights (IW) and produces a vector whose elements indicate how close the input is to the IW .

**The second layer** sums these contributions for each class of inputs to produce as its net output a vector of probabilities .Finally a compete transfer function on the output of the second layer picks up the maximum of these probabilities and produces a 1 for that class and a 0 for the other classes. The architecture for this system is shown above.

The probability of neural network with backs propagation networks in each hidden unit can approximate any continuous nonlinear function. In this paper, we use the Gaussian function as the activation function:

$$radbas = \exp\left[-n^2\right] \quad (18)$$

Finally, one or many larger values are chosen as the output unit that indicates these data points are in the same class via a competition transfer function from the output of summation unit [11], i.e.

$$compet(n) = e_i = [000 \ 0_1 0 \ \dots \ 0_i], n(I) = MAX(n)$$

#### B. LVQ:

The vector quantization technique was originally evoked by Tuevo Kohonen in the mid 80's [17][18]. Both Vector quanzation network and selforganizing maps are based on the Kohonen layer, which is capable of sorting items into appropriate categories of similar objects. Such kind of networks find their application in classification and segmentation problems.

LVQ network comprises of three layers: Input layer, Competitive layer and Output layer [18]. The number of neurons in each layer depends on the input data and the class of the system. The input neurons are as many as the input matrix features of the training pattern, and the number of the output neurons is equal to the number of person's to which face patterns are classified. The number of hidden neurons is heuristic.

In order to implement a face recognition system by our approach, we follow this methodology:

- stage pre-processing using technique WPD
- coding image using LBP
- feature extraction using 2DPCA/2DLDA
- classification using LVQ and PNN network

For face recognition task, we applied Two-dimensional Subspace Analysis (2DPCA, 2DLDA) which are capable of creating the face subspace, defined as orthogonal basis of vectors that contain the most relevant information about a face [12]. These vectors are the eigenvectors of the covariance matrix of the distribution. The recognition process is achieved by forwarding feature matrix, which must be transformed into vectors before providing them to the PNN and LVQ classifier to performed classification and decision.

## V. RESULTS AND DISCUSSION

In order to evaluate and test our approach described for face recognition system, we chose three databases: ORL,

FEI [13][14] and our database of our laboratory. All experiences were performed in Matlab installed on a laptop with a dual core processor T5870 with 2.03 GHz and 2 GB of RAM.

#### A) Our database:

We have collected face images at different moments using a capture device (camera) to form our own database. The database includes 100 face images in a JPG taken on 10 different subjects (N = 10), each been registered under 10 different views.



Fig 4: five images from our database

All experiences have been performed using the ORL own database, and the FEI databases with 5 images for training and 5 images for the test per person for a total of 200 images for each phase (training and test).

To evaluate the performance of our proposed approach, we chose two test databases: ORL and FEI.

1) **Database:** The global performance of algorithms tested on the FEI database is not as good as that of the ORL database. There are two main reasons:

- The image quality of the ORL database is better than that of the FEI database.
- The FEI database is more complex due to variations in the face details and head orientations.
- After a series of experiences, we chose the best values of parameters in order to fix the choice of Eigenvalues, which give a better recognition rate.

2) **Adding some effects:** it is wanted to test our system with and without added noisy in the two data base in order to evaluate robustness of these approaches namely 2DPCA, 2DLDA, DWT, WPD, LBP combined by using two classifier LVQ and PNN.

Noise: Two types of noise are used in this simulation: the Salt and Peppertype noise with a noise density  $a=0.06$  (Figure 5 (a)) and Gaussian noise with mean  $m=0$ , variance  $v=0.04$

Figure 5 illustrates these effects which are obtained as follows.



(a) Salt & pepper Noise (b) Gaussian Noise (c) Gaussian Noise  $m=0, v=0.01$   
 $m=0, v=0.04$



(a) Salt&pepper Noise (b) Gaussian Noise (c) Gaussian Noise  $m=0, v=0.01$   
 $m=0, v=0.04$

Fig5: Adding Noise(database face ORL & FEI)

3) *The pre-processing stage*: we proposed to add a preprocessing stage in order to improve our system's performance in speed by reducing the size and eliminating redundant information from the face images by the means of the DWT and WPD technique, and in other hand reduce the memory and compute of our neural network-training algorithm (PNN) and LVQ

We performed face recognition analysis through WPD and DWT with various wavelet series: Daubechies, Gabor, Coiflets, Symlets and Gauss. in order to select a best wavelet to enhance a rate recognition.

4) *The LBP* is used on WPD and DWT coefficients, derived from WPD to generate detailed high frequency features of animation which forms Feature set one.

In this work, several LBP based algorithms are tested in order to demonstrate the superiority of the proposed approach. Conventionally, selecting the image division strategy for LBP based approaches is heuristical or empirical. Results in previous publications are often obtained by division strategies, which maximize the recognition rates. Here we tested various algorithms with different division strategies. In our test, an image is divided into  $4 \times 4$ ,  $7 \times 7$  and  $16 \times 16$  sub-blocks respectively.

The LBP based face recognition approach consists of extracting the LBP features histograms from the whole face (59-bin histogram), divide the face image into  $7 \times 7$  blocks and concatenate the blocks histograms into a unique vector ( $7 \times 7$  blocks  $\times$  59-bin/block = 2891 features)

5) *Feature extraction using 2DPCA/2DLDA*: After reducing the dimensional of the face images using LBP descriptor. We used the 2DPCA and 2DLDA feature extraction approaches in order to extract the weight images (Features images in the new space) which must be converted into vectors before implementing the classifier network (LVQ and PNN).

6) *Choice of the number of eigenvalues*: Two dimensional methods do not escape this problem, and the choice of the appropriate number depends on the used method and faces database. In our experiences, we have selected the best eigenvalues corresponding to the best variance values (eigenvectors)

7) *Selection parameters and architecture system classifier*

*PNN*: our neural network training algorithm used in system face recognition is not require many parameters compared other neural networks (MLP, BP, LVQ. etc), that only parameter that is needed for performance of the network is the smoothing

parameter  $\sigma$ . Usually, the researchers need to try different  $\sigma$  in a certain range to obtain one that can reach the optimum accuracy [16][18].

To get a higher recognition rate, we have made a series of experiments to choose the best smoothing parameter  $\sigma$  used in PNN.

The probabilistic Neural Network used in our system is composed of two layers:

**Input layer**: The first layer is the input layer and the number of hidden unit is the number of independent variables and receives the input data (number of feature extraction for each approach used in this paper)

**Output layer**: gives the number of faces used in the Database training (ex: ORL 200 person's)

*LVQ*: The changes of LVQ classifier parameters have a high effect on the classification results. In this paper, we found that the best learning rate increases the recognition rate of the system whereas the learning rate is a critical parameter that affected in the recognition process. We use a different number of learning rate (0.1, 0.2, 0.3, and 0.6) with 500 epochs and 80 hidden Neurons experiments.

Table 1. The recognition rate obtained by different methods on the database ORL

Type of classifier	DWT-2DPCA	DWT-2DLDA	WPD-2DPCA	WPD-2DLDA
LVQ	96%	96.8%	94%	95%
PNN	95.8	96%	97%	98%

Table 2. The recognition rate obtained by different methods on the database ORL with added noisy

Type classifier	DWT-2DPCA	DWT-2DLDA	WPD-2DPCA	WPD-2DLDA
LVQ	93%	94%	94%	96%
PNN	92.8%	93%	96%	98%

Table 3. The recognition rate obtained by different methods on the database FEI

Type of classifier	WPD-2DPCA	WPD-2DLDA	WPD/LBP-2DPCA	WPD/LBP-2DLDA
LVQ	92%	94.8%	96%	98%
PNN	94%	95%	98%	99%

Table 4. The recognition rate obtained by different methods on the database FEI with added noisy

Type of classifier	WPD-2DPCA	WPD-2DLDA	WPD/LBP-2DPCA	WPD/LBP-2DLDA
LVQ	90%	92.8%	<b>94%</b>	<b>95%</b>
PNN	94%	95%	<b>97%</b>	<b>98%</b>

Table 5. The running time (s) obtained by different methods on the database FEI with added noisy

	DWT-2DPCA	WPD-2DLDA	WPD&LBP-2DPCA	WPD&LBP-2DLDA
PNN	1.20	1.25	<b>1.10</b>	<b>1.05</b>
LVQ	2.25	2.45	<b>1.98</b>	<b>1.75</b>

*Discussion:*

After these series of experiments, we clearly see the superiority of the two-dimensional methods combined with a probabilistic neural classifier combining those of a LVQ classifier.

We also note that the choice of optimal component and the choice smoothing parameter, which represents a better recognition, rate for both methods, 2DPCA and 2DLDA and accuracy of classification PNN and LVQ.

Table.3 and 4,shows the results obtained with different face images, when using (P=24, R=3), (P=16, R=2), and (P=8, R=1) and the resulting histogram runtime. The results show that the properties of LBP is more reliable and effective for face coding description.

The results obtained after using our approach shows that it reduces the computation time (table .5) of training and improves the recognition rate

VI. CONCLUSION

In this paper, we propose an approach for face recognition based on the combination of two approaches, one used for the reduction of space and feature extractions in two dimensions and the other for classification and decision.

After our experience we have found clear improvement of our system, based on fusion algorithms namely DWT, WPD LBP and bi-dimensional combined by an approach neural classifier, achieved a better recognition rate compared to other techniques described in this article.

Our choice of using WPD with LBP techniques as a preprocessing and coding imageare demonstrated by improved performance of our system in terms of recognition rate, speed of calculation and reduce memory computation of system classifier.

As a perspective, we propose to use this approach in an uncontrolled environment (video surveillance) based on video sequences (dynamic images) in order to make the task of face recognition more robust.

REFERENCES

- [1] Handbook of Biometrics edited by Anil K. Jain Michigan State University, USA Patrick Flynn University of Notre Dame, USA Arun A. Ross West Virginia University, USA © 2008 Springer Science+Business Media, LLC.
- [2] biometric recognition challenges and opportunities Joseph N. Pato and Lynette I. Millett, Editors Whither Biometrics Committee Computer Science and Telecommunications Board Division on Engineering and Physical Sciences Copyright 2010 by the National Academy of Sciences.
- [3] D. Zhang and Z. -H. Zhou. (2D) 2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. Neuro computing, Vol. 69, pp. 224-23 1, 2005.
- [4] N. Nguyen, Wanquan Liu and S. Venkatesh. Random Subspace Two-Dimensional PCA for Face Recognition. Department of Computing, Curtin University of Technology, WA 6845, Australia
- [5] Two-Dimensional PCA:A New Approach to Appearance-Based Face Representation and Recognition Jian Yang, David Zhang,. 26, NO. 1, JANUARY 2004
- [6] S. Noushath, G.H. Kumar, and P. Shivakumara. (2D)2LDA : An efficient approach for face recognition. Pattern Recognition, 39(7) :1396-1400, 2006.
- [7] S. Mallat. A theory of multiresolution signal decomposition : the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7) :674-693, 1989.
- [8] G.C. Feng, P.C. Yuen, D.Q. Dai. Human face recognition using PCA on wavelet subband. SPIE Journal of Electronic Imaging, Vol. 9, No.2, pp. 226-23 3, 2000.
- [9] T. Ojala, M. Pietkainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, July 2002.
- [10] Z. Yang, M. Li and H. Ai, "An experimental study on automatic face gender classification", Proc. IEEE Int. Conf. on Pattern Recognition, (2006) 1099-1102
- [11] Y. Sun, S. Todorovic, and S. Goodison, Local Learning Based Feature Selection for High Dimensional Data Analysis, Vol. 32, No. 9, IEEE Trans. on Pattern Analysis and Machine Intelligence, (2010) 1610-1626
- [12] Computer Vision Using Local Binary Patterns edited by Matti Pietikäinen Abdenour Hadid Guoying Zhao Timo Ahonen ISSN 1381-6446ISBN 978-0-85729-747-1 e-ISBN 978-0-85729-748-8DOI 10.1007/978-0-85729-748-8 Springer London Dordrecht Heidelberg New York 2011
- [13] G. Zhang, et al, "Boosting local binary pattern (LBP)-based face recognition," Advances In Biometric Person Authentication, Proceedings, Vol.3338, Lecture Notes In Computer Science, (2004)179-186
- [14] J. Meng, Y. Gao, X.Wang, T. Lin, J. Zhang, "Face Recognition based on Local Binary Patterns with Threshold", 2010 IEEE, DOI 10.1109/GrC.2010.72[15]Nazish, 2001. Face recognition using neural networks. Proc.IEEE INMIC 2001, pp: 277-281 (2007)
- [15] D. F. Specht, 1990. "Probabilistic neural network and the polynomial adaline as complementary techniques for classification" IEEE Trans. Neural Networks, 1(1): 111-121.
- [16] Neural network toolbox matlabUser's GuideCOPYRIGHT 1992 - 2002 by The MathWorks, Inc.
- [17] Computational intelligence paradigms: theory & applications using MATLAB / S. Sumathi andSurekhaPaneerselvam.© 2010 by Taylor and Francis Group, LLCRC Press is an imprint of Taylor & Francis Group, an Informa businessInternational Standard Book Number: 978-1-4398-0902-0
- [18] ORL. The ORL face database at the AT&T (Olivetti) Research Laboratory (1992), <http://www.uk.research.att.com/facedatabase.html>
- [19] FEI. The FEI face database at the Artificial Intelligence Laboratory of FEI in São Bernardo do Campo, São Paulo, Brazil (June 2005 and March 2006) , <http://fei.edu.br/~cet/facedatabase.ht>

# Combinaison du Filtrage Linéaire et le Seuillage Adaptatif d'Ondelettes pour le Dé-bruitage d'Images Satellitaires

Abdelhak BOUHALI<sup>1</sup> et Daoud BERKANI<sup>2</sup>

Laboratoire Signal & Communications, (ENP), Alger, Algérie

<sup>1</sup>abdelhak.bouhali@g.enp.edu.dz

<sup>2</sup>daoud.berkani@enp.edu.dz

**Résumé**— Le progrès technologique enregistré dans le domaine des engins spatiaux a rendu possible le perfectionnement du rendement des constellations des satellites permettant ainsi d'exécuter la télédétection à une fraction du coût et du temps des missions traditionnelles de celles-ci. Quoique ces progrès ne cessent de s'avancer, les dispositifs de formation d'images satellitaires ne s'épargnent pas de provoquer des artefacts le long de celles-ci qui, sont ainsi souillées par différents types de bruits dont, la plupart du temps modélisés par une distribution Gaussienne. Dans cet article, et en vue de surmonter à ses déficiences, on opte dans le cadre de notre travail pour une des intéressantes approches qui a marqué de notables progrès tout récemment ces trois dernières décennies : la théorie des ondelettes. L'image contaminée est décomposée d'abord en coefficients d'ondelettes via une 2D-DWT à 2 niveaux. Les coefficients résultants sont les détails: LH1, HL1 et HH1 sur lesquels est appliqué un seuillage adaptatif, LH2, HL2 et HH2 auxquels est appliqué un bloc de filtrage stratégique formé d'une combinaison intelligente des avantages des filtres linéaires moyen et wiener. Et finalement l'approximation LL2 est soumise à un filtrage simple de wiener. Les résultats expérimentaux montrent une amélioration notable en termes du PSNR et du coefficient de corrélation par rapport aux résultats du récent état de l'art.

**Mots-clés**— Filtrage Linéaire, Ondelettes, 2D-DWT, Seuillage Adaptatif, Dé-bruitage, Image Satellitaire

## I. INTRODUCTION

Les applications du traitement du signal en technologie de l'information et de communication sont en accrue permanente et jouent un rôle central dans le développement des systèmes numériques de télécommunication et d'automatisation, y incluent les communications mobiles, le Radar, le traitement des signaux médicaux, le traitement d'images médicales, satellitaires, ...etc. De nos jours, l'un des plus importants signaux adoptés dans plusieurs et sensibles applications qu'étant civiles que militaires est l'imagerie satellitaire dont l'employabilité est répandue dans diverses applications telles que les études géo-scientifiques, l'astronomie, les systèmes d'informations géographiques, ... etc. Les facteurs de qualité les plus importants dans de telles images viennent de leur netteté, résolution et leur compressibilité le fait que l'amélioration de l'un d'eux ces critères augmentera intuitivement la qualité des applications appropriées. Dans cet article, on se limite à l'opération du dé-bruitage pour ouvrir

prochainement le champ aux autres applications dans de futurs travaux. Et en vertu du théorème de la limite centrale (TCL), le bruit additif contaminant l'image originale est considéré comme étant Gaussien.

Ainsi, la naissance de l'analyse multi-résolution via le développement des transformées d'ondelettes dès le début des années 1980s par GROSSMAN et MORLET (1984) [1]-[4] a ouvert un champ très important d'applications dans les différentes disciplines du traitement du signal grâce à un certain nombre d'avantages par rapport aux approches classiques. On se consacre dans le cadre de ce travail à l'utilisation de la théorie des ondelettes en combinaison avec les avantages des filtres linéaires moyen et wiener pour bien traiter les signaux satellitaires bidimensionnels ; ou l'imagerie satellitaire. Et à cet effet, l'image acquise est décomposée en sous-bandes à deux niveaux de décompositions via une 2D-DWT, et puis appliquer un seuillage adaptatif aux détails les plus fins (LH1, HL1, HH1), un filtrage stratégique aux détails (LH2, HL2 et HH2) et un filtrage simple de wiener pour l'approximation grossière (LL2).

La suite de cet article est organisée comme suit. La prochaine section est consacrée principalement pour un survol sur l'analyse multi-résolution et le seuillage d'ondelettes. La section III décrit l'approche proposée constituée d'un mélange hybride du domaine d'ondelettes et du filtrage linéaire. Les résultats de simulation ainsi que les commentaires appropriés sont présentés dans la section IV. Et finalement une conclusion et des perspectives sont données dans la section V.

## II. L'ANALYSE EN ONDELETTES

Souvent une composante spectrale particulière survenant à un instant quelconque peut revêtir un intérêt certain. Dans un tel cas, il serait très bénéfique de connaître l'intervalle de temps où cela se produit. L'analyse d'ondelettes peut fournir simultanément l'information temporelle et fréquentielle lors d'une représentation d'un signal donné [4]. Elle a été développée par Grossmann et Morlet [1]-[4] au début des années 1980s comme une approche alternative à la STFT pour surmonter au problème de résolution de cette dernière.

### A. Ondelettes et CWT

Une ondelette est une forme d'une petite onde irrégulière et souvent asymétrique, de durée limitée (a un début et une



fin) et de moyenne nulle. Des caractéristiques qui rendent son usage très intéressant pour décrire des anomalies, des impulsions et d'autres événements apparaissant dans un signal [5].

La fonction  $\Psi(t)$  est dite d'ondelette si elle vérifie la condition d'admissibilité suivante dans le domaine fréquentiel [3], [6]:

$$\int_0^{\infty} \frac{|\hat{\Psi}(\omega)|^2}{\omega} d\omega = \int_{-\infty}^0 \frac{|\hat{\Psi}(\omega)|^2}{|\omega|} d\omega = C_{\Psi} < +\infty, \text{ avec :}$$

$\hat{\Psi}(f)$  est la transformée de Fourier de  $\Psi(t)$ .

Cette condition implique que:  $\int_{-\infty}^{\infty} \Psi(t) \cdot dt = 0$  [3], [4], [6] et,

elle est souvent renforcée par le fait que l'ondelette ait  $m$  moments nuls [6], c.à.d.:  $\int_R t^k \Psi(t) \cdot dt = 0$ , pour  $k = 0, \dots, m$ .

De ce prototype, sont souvent tirées d'autres fonctions de base utilisées dans les processus de transformation tout en jouant sur les paramètres de translation 's' et/ou d'échelle 'u' [2], [6]. Elles sont décrites par :

$$\Psi_{u,s}(t) = \frac{1}{\sqrt{s}} \cdot \Psi\left(\frac{t-u}{s}\right)$$

La transformée en ondelettes peut être vue comme étant un outil découpant des données ou des fonctions dans différentes composantes de fréquence, et puis étudiant chacune d'elles avec la résolution correspondant à son échelle [4]. Ainsi, la transformée continue d'une fonction 'f' est formulée par [2], [4], [6]:

$$Wf(u,s) = \int_R f(t) \cdot \Psi_{u,s}^*(t) \cdot dt = \int_{-\infty}^{\infty} f(t) \cdot \frac{1}{\sqrt{s}} \cdot \Psi^*\left(\frac{t-u}{s}\right) \cdot dt$$

### B. Discrétisation de la CWT et Naissance de la DWT

Avec l'évolution importante constatée dans le monde des mathématiques appliquées ainsi que le progrès technologique immense enregistré dans le domaine des dispositifs numériques, les ordinateurs et les calculateurs numériques sont de plus en plus adoptés pour mettre en œuvre la plupart des calculs et de manipulations mathématiques. Ainsi, l'une des plus importantes et intéressantes manipulations qui peut bénéficier de telles technologies est la transformée en ondelettes ; le fait qu'il est plus intéressant (en termes du coût et du temps de calcul) de la discrétiser et de la calculer numériquement que de l'accomplir analytiquement.

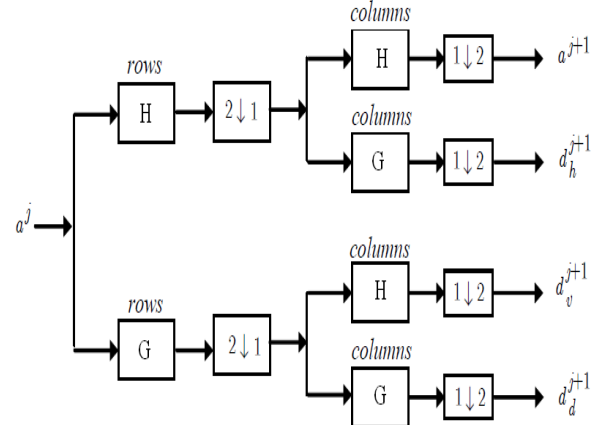
Ainsi, est-il nécessaire de connaître  $Wf(u,s)$  le long de tout son domaine de définition pour pouvoir reconstruire  $f(t)$  [4], [6]? Si la réponse est négative, l'utilisation d'un sous-ensemble discret semble un objectif assez-raisonnable [6]. L'idée est la suivante [2], [4], [6]: considérons des sous-ensembles discrets de  $\mathbf{R}^{+*}$  et de  $\mathbf{R}$ . En suite, posons :  $a > 1$ ,  $u_0 > 0$ ,  $s \in \{a^j\}_{j \in \mathbf{Z}}$  et  $u \in \{n \cdot a^j \cdot u_0\}_{j,n \in \mathbf{Z}}$ . Alors, la forme des atomes (fonctions) de base devient :

$$\Psi_{j,n}(t) = \frac{1}{\sqrt{a^j}} \cdot \Psi\left(\frac{t - n u_0 a^j}{a^j}\right),$$

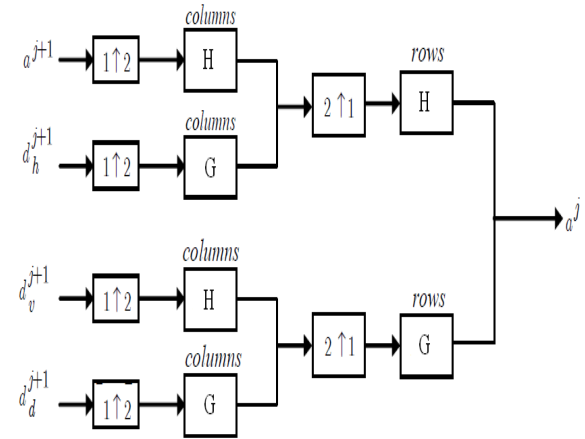
Une forme d'échantillonnage dite *dyadique* est la plus adoptée pour de telles situations en raison de sa convenance au théorème d'échantillonnage de Shannon [6] et sa conformité en termes de programmation et de réalisation pratique. En effet, la valeur la plus commode pour cette approche s'avère '2' pour "a" et '1' pour "u<sub>0</sub>" permettant ainsi d'avoir des bases ortho-normales  $\{\Psi_{j,n}\}_{(j,n) \in \mathbf{Z}^2}$  [2], [4], [6]. Dans ce cas, l'équation ci-dessus devient:

$$\left\{ \Psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \cdot \Psi\left(\frac{t - 2^j n}{2^j}\right) \right\}_{(j,n) \in \mathbf{Z}^2}$$

Quoique la transformée en ondelettes discrétisée puisse être calculée sur un ordinateur numérique, ce calcul peut prendre quelques secondes à quelques heures ou plus selon la taille du signal et la résolution voulue. Un algorithme suffisamment rapide développé par S. Mallat à la fin des années 1980s [1], [6] est réellement disponible pour calculer la transformée d'ondelettes d'un signal; c'est la naissance de transformée discrète d'ondelettes (DWT) [1], [2], [6].



Basic step for the decomposition in wavelets for images



Basic step for wavelet reconstruction for images

Fig. 1 2D-Discrete Wavelet Transform (2D-DWT)

La DWT analyse le signal à différentes bandes de fréquence avec différentes résolutions en lui décomposant en une approximation globale (basses fréquences) et une information détaillée (hautes fréquences). Cette décomposition est simplement obtenue par des filtrages successifs passe haut  $g(n)$  et passe bas  $h(n)$ . Le signal résultant est soumis ainsi à un sous-échantillonnage par 2. Ceci peut être exprimé comme suit [1], [2]:

$$a_{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n-2p] \cdot a_j[n]$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{+\infty} g[n-2p] \cdot a_j[n]$$

Ainsi, la procédure précédente est suivie dans un ordre renversé pour la reconstruction. Par conséquent, la formule de reconstruction devient [1], [2]:

$$a_j[p] = \sum_{n=-\infty}^{+\infty} h[p-2n] \cdot a_{j+1}[n] + \sum_{n=-\infty}^{+\infty} g[p-2n] \cdot d_{j+1}[n]$$

Toute la procédure de la DWT ainsi que sa transformée inverse sont bien illustrées dans la Fig. 1 [1], [2], [6], tout en l'appliquant sur un signal bidimensionnel pour bien éclaircir l'idée générale de cette approche.

### C. Seuillage d'ondelettes

L'évaluation des fonctions basées sur des représentations d'ondelettes mène à des algorithmes simples et puissants qui sont souvent plus faciles d'accorder que les méthodes traditionnelles. Quoiqu'elles soient souvent considérées comme étant des problèmes particulièrement difficiles, la discrimination d'un signal noyé dans un fond de bruit (dé-bruitage) ainsi que sa compression constituent deux grands succès liés à la théorie des ondelettes [6]. Ceux-ci sont principalement fondés sur le mécanisme de seuillage des coefficients d'ondelettes. Les méthodes de seuillage les plus simples sont basées sur l'exécution d'une transformation d'ondelettes des données transmises et puis employer le même seuil pour traiter tous les coefficients dans l'expansion (ou au moins tous les coefficients au-dessus d'un niveau de résolution primaire au-dessous duquel aucun seuillage n'est effectué) [7]. L'algorithme de base de telles approches est très simple et consiste des trois étapes suivantes: décomposition, seuillage et reconstruction [6], [8]-[11]. Alors, commencer d'abord par décomposer le signal à traiter via une base orthogonale d'ondelettes utilisant une transformée discrète, puis choisir une partie des coefficients utiles par seuillage des coefficients bruts grâce à des algorithmes bien choisis pour finalement clôturer cette approche par l'utilisation des coefficients seuillés pour la reconstruction du signal voulu.

Ainsi, l'estimation d'un signal à base du seuillage des coefficients d'ondelettes est principalement régie par une multitude de variantes [6], à savoir :

- Stratégie du seuillage : global, niveau par niveau, local en temps, par blocs ;
- Choix de la valeur du seuil ;

➤ Type du seuillage : deux principales approches sont généralement adoptées ; le seuillage dur (hard) et le seuillage mou (soft) [2], [6], [7], [12]-[16]. Dans le premier cas, on ne conserve que les coefficients supérieurs au seuil, tandis que le seuillage mou ne conserve que les coefficients supérieurs au seuil avec une atténuation des amplitudes par la valeur de celui-ci afin de s'assurer d'avoir enlevé l'effet du bruit même pour les forts coefficients. Pour des coefficients ' $d$ ', ils sont formulés par :

$$\begin{cases} \text{hard}(d) = \text{soft}(d) = 0 & \text{si } |d| < \text{seuil} \\ \text{hard}(d) = d & \text{si } |d| \geq \text{seuil} \\ \text{soft}(d) = d - [\text{sign}(d) \cdot \text{seuil}] & \text{si } |d| \geq \text{seuil} \end{cases}$$

➤ **Stratégie de construction de l'estimateur** : Une estimation est dérivée via un opérateur qui atténue les petits coefficients présentant en général le bruit additif tout en préservant le signal utile. L'objectif fondamental de cette opération est de minimiser l'erreur de l'estimation qui est mesurée par une fonction dite de *perte*. Dépendamment de cette fonction de perte, y découle une variété de stratégies de construction d'un estimateur. L'optimisation d'un estimateur requiert généralement de tirer profit des informations antérieures. En effet, la théorie de *Bayes* utilise un modèle probabiliste du signal pour dériver les estimateurs qui réduisent au minimum le risque moyen. Le plus grand défi de cette approche est d'acquérir suffisamment d'informations pour définir la distribution préalable du signal, ce qui est souvent impossible pour les signaux complexes tels que les images naturelles. Une alternative est offerte par l'approche de *minimax* qui exige seulement de savoir l'ensemble antérieur où le signal est garanti d'être. Cette procédure consiste à minimiser le risque maximum à travers cet ensemble [2].

Tandis que l'idée du seuillage est simple et efficace, l'acte de trouver un bon seuillage est la tâche la plus difficile dans de tels processus. Ainsi, de récents états de l'art ont bien montré que tout seuil nécessite la connaissance préalable du niveau ( $\sigma$ ) du bruit. En effet, DONOHO et JOHNSTONE [9] ont affirmé qu'il est plus pratique d'estimer le niveau ( $\hat{\sigma}$ ) du bruit à partir des données à analyser que de supposer ce même niveau est connu. Ils ont stipulé qu'il est très important d'utiliser un estimateur assez robuste à savoir la médiane en cas où les coefficients d'ondelettes de la fine échelle contiennent une petite proportion du signal utile mélangé dedans avec du bruit. Donc, ils ont dérivé pratiquement un estimateur à partir des coefficients pratiques d'ondelettes de l'échelle la plus fine comme le montre la formule suivante :

$$\hat{\sigma} = \frac{\text{median} \left( \left\{ \left| \hat{d}_{L-1,k} \right| : k = \overline{0, 2^{L-1} - 1} \right\} \right)}{0.6745}$$

Le seuillage des coefficients d'ondelettes est donc un bon moyen d'ignorer les détails les plus faibles que l'on peut assimiler à du bruit et ne conserver que les coefficients les plus importants. On reconstruit alors le signal à partir de ces

seuls coefficients, les autres étant mis à zéro. L'opération du seuillage pour le dé-bruitage/compression est donc accomplie suivant l'algorithme de la Fig. 2 :

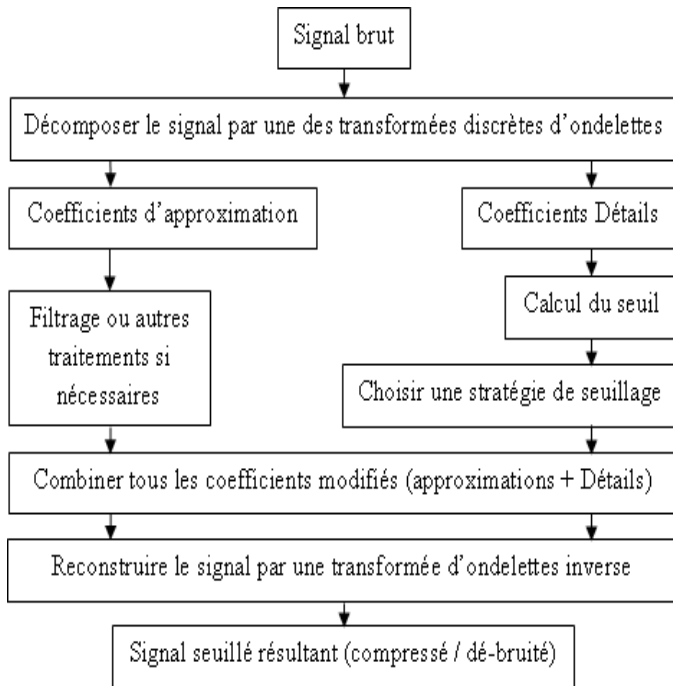


Fig. 2 Algorithme du Seuillage d'Ondelettes

### III. L'APPROCHE PROPOSEE

L'idée principale derrière la réalisation de ce travail peut être résumée comme suit. Quoique l'idée du seuillage des coefficients d'ondelettes ait notablement réussi de résoudre beaucoup de problèmes du dé-bruitage d'images, l'opération du rétrécissement des coefficients à travers toutes les sous

bandes détails via ce mécanisme conduit toutefois à de pertes considérables dans les détails de l'image traitée faute au phénomène de lissage accompagnant généralement l'opération du seuillage.

Ainsi, et voir les performances des filtres linéaires de wiener et moyen en termes de la qualité des résultats obtenus et du temps de réponse, ainsi que leurs rendements complémentaires en fonction du niveau du bruit, c.à.d. chacun est favorisé dans une gamme de bruits complémentaire à celle de l'autre, une combinaison intelligente de ceux-ci (filtres de wiener et moyen) s'avère d'un intérêt majeur dans les processus de réduction du bruit permettant ainsi d'améliorer considérablement leurs performances.

Dans le cadre de ce travail, et afin de bénéficier des avantages de la combinaison des filtres wiener et moyen, ainsi que de surmonter aux déficiences du seuillage d'ondelettes, une alternative à ce dernier est proposée en lui apportant quelques modifications à son mode d'emploi tout en lui joignant les avantages de la combinaison des filtres de wiener et moyen.

L'algorithme de base de cette approche (voir Fig. 3) peut être exprimé via le processus suivant :

1. Décomposer l'image à traiter via une 2D-DWT à deux niveaux de décomposition pour obtenir les sous-bandes détails du premier niveau (LH1, HL1, HH1), du second niveau (LH2, HL2, HH2) et l'approximation globale LL2.
2. Étant donné que le bruit a plus de probabilité de présence dans les sous-bandes du premier niveau de décomposition (détails les plus fins), celles-ci (LH1, HL1 et HH1) sont soumises au seuillage adaptatif d'ondelettes afin de bien supprimer la majeure partie du bruit.

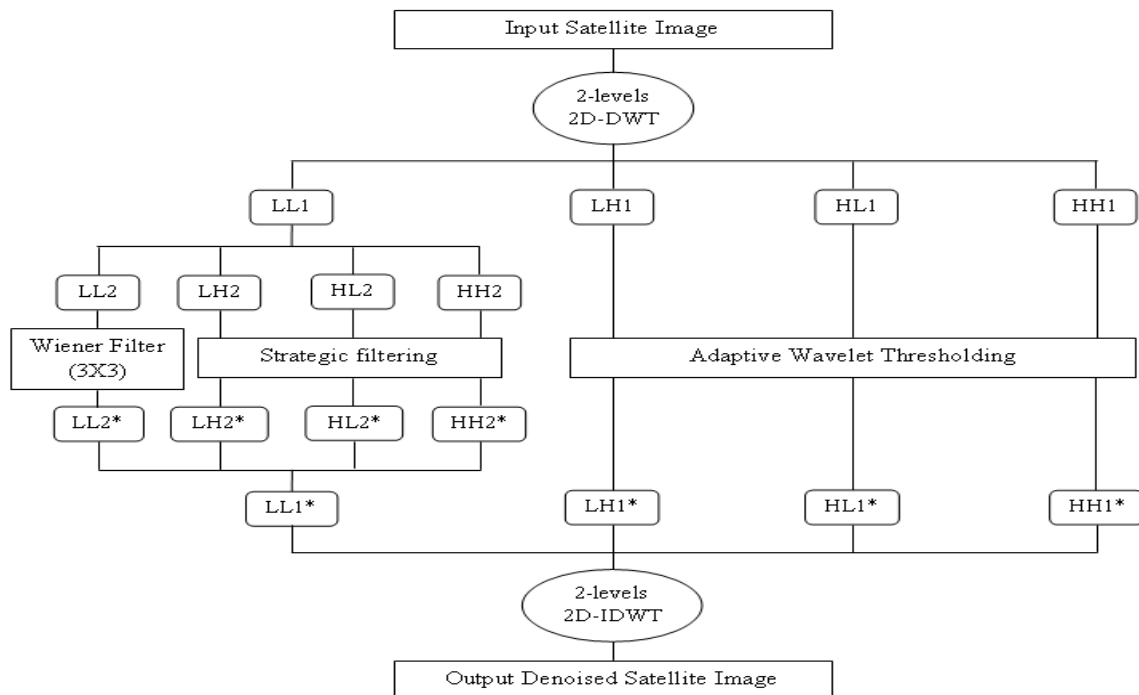


Fig. 3 L'Approche Proposée

3. Contrairement aux détails les plus fins, les sous-bandes LH2, HL2 et HH2 ont plus de probabilité de présenter les détails de l'image que le bruit additif. Et afin de surmonter au problème de lissage pouvant se manifester à cause de l'opération de seuillage, une bonne combinaison des avantages des filtres wiener et moyen peut être introduite et appliquée aux sous-bandes LH2, HL2 et HH2 permettant ainsi d'enlever le bruit résidu de l'étape précédente tout en préservant les détails perdus via les approches classiques du seuillage.
4. En tant que représentative de l'approximation globale de l'image, un simple filtre de wiener (de fenêtre 3x3) est appliqué à la sous-bande LL2 afin de la nettoyer de la faible proportion du bruit dont est probablement infectée.
5. Et finalement, tous les coefficients résultants du seuillage et/ou du filtrage des sous-bandes détails du premier et second niveau de décomposition, ainsi que ceux issus du filtrage de l'approximation globale LL2 sont recombinaés de nouveau pour être soumis ensuite à la transformée d'ondelettes inverse (2D-IDWT à deux niveaux de décompositions). L'image acquise aura donc moins de bruit avec plus de détails préservés.

#### IV. RESULTATS ET COMMENTAIRES

Les performances de l'approche proposée dans cet article concernant le dé-bruitage d'images satellitaires d'une zone urbaine sont évaluées en termes du PSNR et du coefficient de corrélation. Le premier étant une mesure quantitative, il renseigne sur la quantité de bruit subsistant dans l'image dé-bruitée, et il est formulé par :

$$PSNR (dB) = 10 \cdot \log_{10} \left\{ \frac{(255)^2}{\sum_{x,y} [X(x,y) - \hat{X}(x,y)]^2} \right\}$$

La seconde mesure représentée par le coefficient de corrélation  $\rho$  peut être vue comme étant une mesure qualitative le fait qu'elle renseigne sur la similitude et la corrélation entre l'image originale et l'image dé-bruitée. Elle est donnée par l'expression suivante:

$$\rho(\%) = \frac{COV(X, \hat{X})}{\sqrt{VAR(X) \cdot VAR(\hat{X})}} \times 100 \%$$

Avec :  $X(x, y)$  et  $\hat{X}(x, y)$  présentent l'image originale et dé-bruitée respectivement.

Les expériences effectuées lors de ce travail sont accomplies sur une image test prise dans le toolbox Matlab « westconcordorthophoto.png », et qui est une image satellitaire d'une zone urbaine, prise orthogonalement et en niveaux de gris de « West-Concord City », et dont la taille est ajustée pour convenance à (512 x 512) pixels. Un bruit additif *iid* de distribution Gaussienne et à différents niveaux est généré via la commande « randn » de Matlab. Pour la transformée d'ondelettes, l'ondelette de « symlet » à huit moments nuls est utilisée.

Pour évaluer les performances du travail accompli, une comparaison avec les meilleurs travaux du récent état de l'art relatif au traitement d'ondelettes à savoir BayesShrink (BS) et NormalShrink (NS) ainsi qu'au filtrage linéaire tel que les filtres de wiener et moyen est accomplie.

TABLE I  
COMPARAISON DES DIFFERENTES METHODES DE REDUCTION DU BRUIT EN TERMES DU PSNR (DECIBEL)

Noise Level	Noisy Signal	Average Filter (3X3)	Wiener Filter (5X5)	NormalShrink	BayesShrink	Our Proposed Approach
05	34.1515	28.5376	27.7141	35.1144	<b>35.653</b>	34.4203
10	28.1432	28.1613	27.2844	31.0858	<b>31.3708</b>	30.7767
15	24.6163	27.5748	26.6876	28.9193	<b>29.0935</b>	28.9487
20	22.0829	26.8967	26.0887	27.3865	27.5551	<b>27.8518</b>
30	18.5747	25.4061	24.906	25.2892	25.5188	<b>26.002</b>
40	16.0759	23.8879	23.7959	23.78	24.0098	<b>24.6173</b>
50	14.1587	22.5326	22.7835	22.626	22.8633	<b>23.5064</b>
60	12.5651	21.2541	21.8128	21.5878	21.8283	<b>22.4144</b>
70	11.2474	20.1575	20.9614	20.7471	21.0487	<b>21.4888</b>

TABLE II  
COMPARAISON DES DIFFERENTES METHODES DE REDUCTION DU BRUIT EN TERMES DU COEFFICIENT DE CORRELATION (%)

Noise Level	Noisy Signal	Average Filter (3X3)	Wiener Filter (5X5)	NormalShrink	BayesShrink	Our Proposed Approach
05	99.586	98.519	98.248	99.666	<b>99.705</b>	99.618
10	98.38	98.371	98.044	99.153	<b>99.208</b>	99.094
15	96.456	98.115	97.724	98.602	<b>98.66</b>	98.629
20	93.908	97.776	97.354	98.006	98.088	<b>98.217</b>
30	87.687	96.835	96.461	96.763	96.94	<b>97.246</b>
40	80.787	95.53	95.377	95.438	95.68	<b>96.196</b>
50	73.85	93.964	94.141	94.079	94.382	<b>95.096</b>
60	67.323	92.068	92.692	92.565	92.92	<b>93.749</b>
70	61.627	90.074	91.188	91.121	91.622	<b>92.371</b>

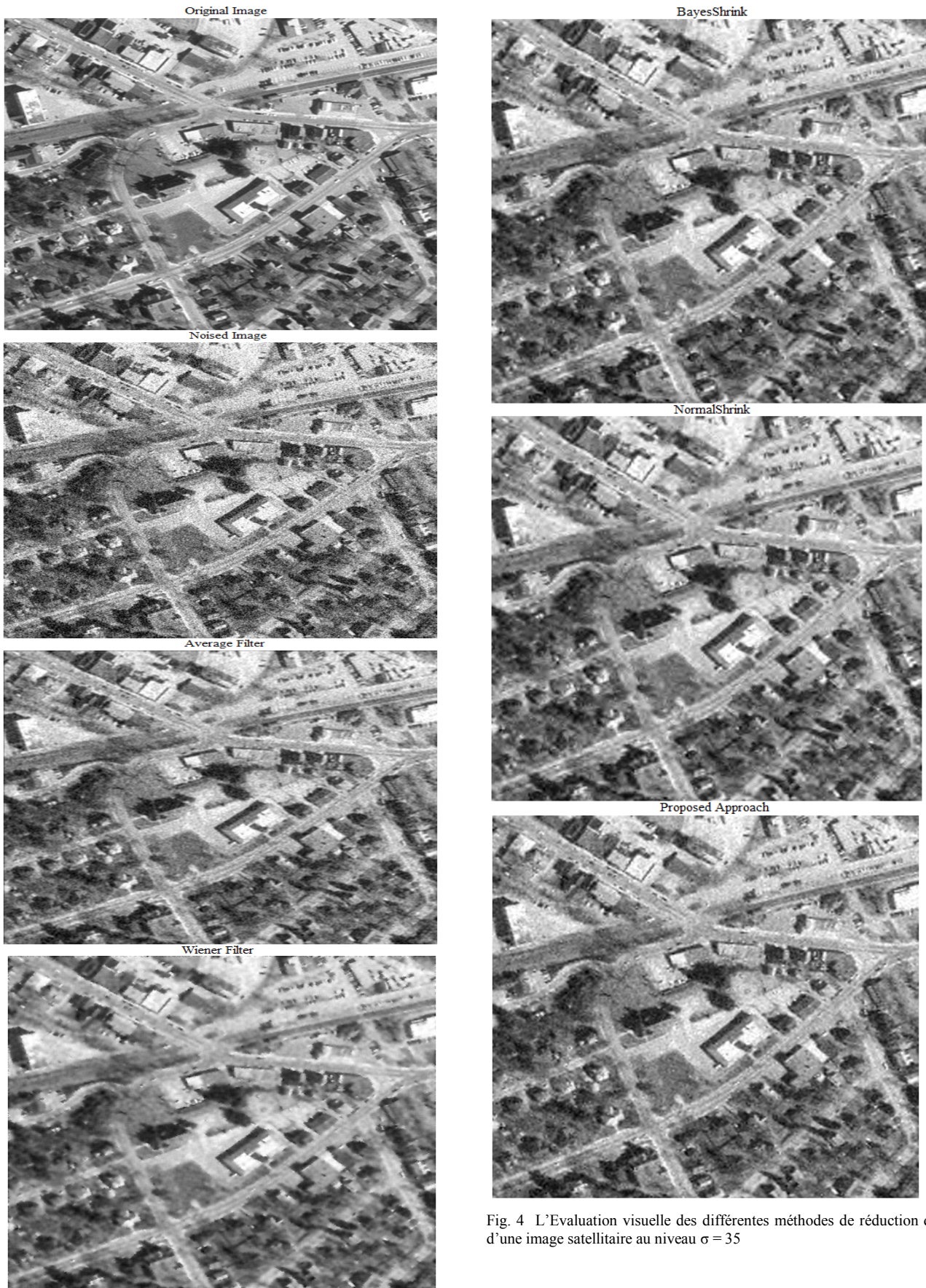


Fig. 4 L'Evaluation visuelle des différentes méthodes de réduction du bruit d'une image satellitaire au niveau  $\sigma = 35$

Une évaluation objective de notre approche est assurée quantitativement via sa comparaison en termes du PSNR avec les meilleures approches existantes dans la littérature, à savoir les filtres moyen et wiener ainsi que les rétrécissements d'ondelettes BayesShrink (BS) et NormalShrink (NS) ; comme est montrée dans la TABLE I. Ainsi, une évaluation qualitative est illustrée dans la TABLE II par une étude de similitude de l'image dé-bruitée à l'image originale via une comparaison accomplie en termes du coefficient de corrélation. En plus aux critères numériques cités ci-avant, une alternative visuelle est optée toutefois dans de telles situations comme le présente la Fig. 4, tout en se basant sur la clarté et la netteté de l'image dé-bruitée avec préservation du maximum possible des détails descriptifs de cette dernière.

Les résultats présentés dans la TABLE I et II montrent que les méthodes de seuillage d'ondelettes demeurent très intéressantes dans la majorité des opérations du dé-bruitage d'images y compris l'imagerie satellitaire dont est l'objet de cet article. Les résultats du PSNR et du coefficient de corrélation montrent que les rétrécissements d'ondelettes BayesShrink et NormalShrink sont plus appropriés aux niveaux bas du bruit le fait que le calcul de la valeur du seuil et par conséquent le taux de rétrécissement des coefficients dépend souvent du niveau du bruit additif, d'où l'effet du rétrécissement sera sensiblement faible à ces niveaux. Par contre, aux niveaux hauts du bruit, les réponses des quatre méthodes comparées auparavant sont presque pareilles avec quelques légères et contradictoires différences, c.à.d. malgré que BayesShrink (BS) a un meilleur PSNR et coefficient de corrélation ( $\rho$ ), le filtre de wiener a la meilleur perception visuelle et préservation de bords.

Comme est prévu théoriquement, l'approche proposée dans cet article a clairement marqué de très bons résultats en comparaison aux autres travaux du récent état de l'art. En effet, et selon les résultats présentés dans les TABLES I et II, on remarque bien que notre approche surpasse généralement les méthodes de BayesShrink et NormalShrink ainsi que les filtres moyen et wiener tant qu'en termes du PSNR par une amélioration atteignant l'ordre de 0.6dB ainsi qu'en termes du coefficient de corrélation. Et quoique l'effet du dé-bruitage soit plus important pour les méthodes de seuillage adaptatif d'ondelettes aux niveaux bas du bruit, la perception visuelle de l'approche proposée demeure très semblable à celle de BayesShrink et NormalShrink à de tels niveaux, comme est illustré dans la Fig. 4. Cette amélioration est atteinte grâce à l'exploitation efficace des avantages du filtrage linéaire en collaboration avec le seuillage adaptatif d'ondelettes via une utilisation intelligente de ceux-ci en fonction des sous-bandes de la décomposition en cours, ce qui l'attribue le caractère d'efficacité en préservant les bords et les différents détails importants de l'image de la meilleure façon lors de la suppression du bruit. Ainsi, cette approche hybride est testée avec succès sur plusieurs images satellitaires des zones urbaines dont la préservation des bords est le plus grand défi lors de leurs traitements, et par conséquent, elle est bien recommandée comme un modèle bien compétent et efficace pour le dé-bruitage de tels genres d'images. Également, le graphe de la Fig. 5 illustre la représentation graphique de la comparaison des différentes méthodes du dé-bruitage de l'image satellitaire en fonction du niveau du bruit.

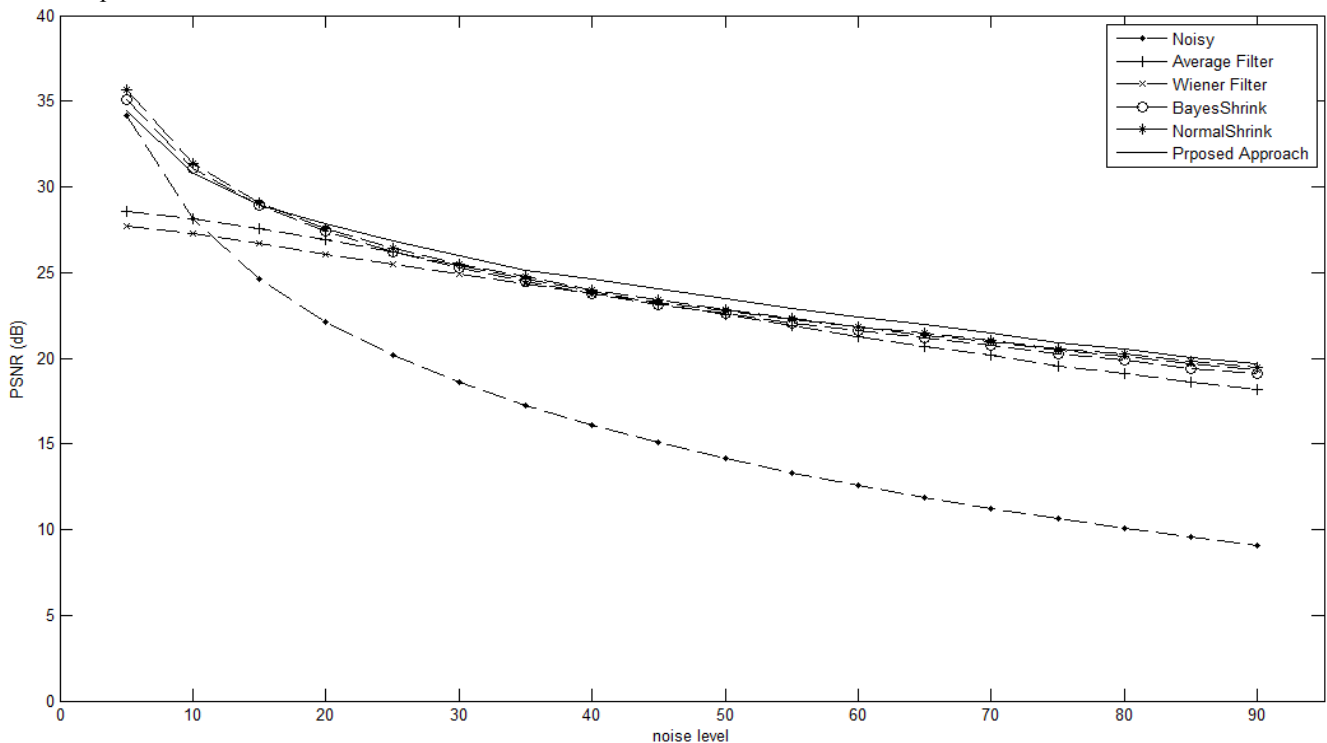


Fig. 5 Graphes de comparaison des différentes méthodes de réduction du bruit d'une image satellitaire

## V. CONCLUSION

Dans cet article, une simple et efficace approche pour le dé-bruitage d'images satellitaires des zones urbaines - combinant les avantages du filtrage linéaire et du seuillage adaptatif d'ondelettes - est présentée. Cette alternative permet de bien accomplir l'opération du dé-bruitage de telle sorte que l'image traitée soit plus proche possible de l'image originale tout en préservant la plupart de ses propriétés importantes. Le facteur clé d'efficacité de cette approche peut être récapitulé en deux points principaux. Le premier étant la restriction de l'application du seuillage adaptatif d'ondelettes aux détails les plus fins représentant la majeure partie du bruit additif, permettant ainsi de minimiser l'effet du lissage accompagnant généralement le mécanisme du seuillage. Tandis que le second point de cette approche est l'application du filtrage linéaire approprié pour le reste des sous-bandes de la décomposition, servant par conséquent d'enlever et d'extraire le reste du bruit résidu de la première partie, et donc une préservation efficace et maximale des bords et les détails importants de l'image traitée. Il est possible aussi d'améliorer les performances de cette approche en la généralisant sur d'autres types d'images tout en jouant sur le mécanisme à appliquer aux sous-bandes autres que les détails les plus fins, ainsi que sur la nature de la transformée à utiliser.

## RÉFÉRENCES

- [1] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.
- [2] S. Mallat, "A Wavelet Tour of Signal Processing," 3rd edition, *The Sparse Way*. Academic Press, December 2008.
- [3] S. Mallat and W. L. Hwang, "Singularity Detection and Processing with Wavelets," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617–643, March 1992.
- [4] I. Daubechies, *Ten Lectures on Wavelets, Vol. 61 of Proc. CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia, PA: SIAM, 1992.
- [5] D. L. Fugal, "Conceptual Wavelets In Digital Signal Processing," *Space & Signals Technical Publishing, San Diego, California*, 2009.
- [6] M. Misiti, Y. Misiti, G. Oppenheim and J. M. Poggi, "Wavelets and Their Applications," *ISTE Ltd*, 2007.
- [7] I. M. Johnstone and B. W. Silverman, "Wavelet Threshold Estimators for Data with Correlated Noise," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 59, no. 2, pp. 319–351, 1997.
- [8] D. L. Donoho and I. M. Johnstone, "Ideal Spatial Adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Aug. 1994.
- [9] D. L. Donoho and I. M. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, Dec. 1995.
- [10] D. L. Donoho, "De-Noising by Soft-Thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [11] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian and D. Picard, "Wavelet Shrinkage: Asymptopia?" *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 2, pp. 301–369, 1995.
- [12] D. Zhigang, Z. Jingxuan and J. Chunrong, "An Improved Wavelet Threshold Denoising Algorithm," *2013 Third International Conference on Intelligent System Design and Engineering Applications*, pp. 297–299, 2013.
- [13] Y. Norouzzadeh and M. Rashidi, "Image Denoising in Wavelet Domain Using a New Thresholding function," *International Conference on Information Science and Technology*, Nanjing, Jiangsu, China, March 26–28, 2011, pp. 721–724.
- [14] S. G. Chang, B. Yu and M. Vetterli, "Adaptive Wavelet Thresholding for Image Denoising and Compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, September 2000.
- [15] B. J. Yoon and P. P. Vaidyanathan, "Wavelet Based Denoising by Customized Thresholding," *ICASSP*, pp. II-925–II-928, 2004.
- [16] A. Fathi and A. R. N. Nilchi, "Efficient Image Denoising Method Based on a New Adaptive Wavelet Packet Thresholding Function," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 3981–3990, September 2012.

# Urban Roads Extraction from Very High Resolution Satellite Images

GACEM.A, BERRACHED.N, MERAD BOUDIA.S

Laboratory of Intelligent Systems Research (LARESI) Electronics Department - FGE – USTO  
Oran, ALGERIA Address

amel\_telecom@yahoo.ca

samy\_merad@yahoo.fr

Laboratory of Intelligent Systems Research (LARESI) Electronics Department - FGE – USTO

nasrlberrached@yahoo.fr

**Abstract**— Currently, Remote Sensing images with high spatial and spectral resolution are available in urban areas. It is potentially a very important contribution to the mapping of cities. The high spatial resolution of these new sources allows a true representation of streets on a map, but generates a significant increase in artifacts. We present in this paper, a method based on fast marching algorithm for the extraction of roads. Results of the extraction are vectorized, and can be integrated in GIS. The algorithm is fast, has a few parameters, and can be affected by the resolution, or the noise in the image. This algorithm is available from Orfeo ToolBox.

**Keywords**— Image Processing, Remote sensing, Road extraction, Fast marching

## I. INTRODUCTION

Today satellite remote sensing systems provide large volumes of data, that are invaluable in monitoring Earth resources and the effects of human activities. Road feature extraction from remotely sensed images has been a long-term topic of research, and because of its complexity is still a challenging topic. Urban road mapping from high spatial resolution images are an asset, because they allow the discrimination of urban features, but they also bring new problems because, a lot of the urban features can be considered as noise, for example, cars, shadows, and trees, hinder the discrimination of roads. However, with high resolution images, the road-width can vary considerably, and additional variations are present such as lane markings, vehicles, shadows projected by buildings and trees. These variations make the extraction of road networks a much more difficult problem. Since 1990, different techniques have been proposed to extract road from remote sensing images in automatic or semi-automatic procedure. In [1], an exhaustive bibliography on automated road extraction is provided. Among these methods: active contours [3], snakes [2], probabilistic approaches [5] or dynamic programming [4]. However, to our best knowledge, none of these algorithms satisfies the operators, mainly because of computation time. In most situations, road extraction is still manual. The purpose of this project is to experiment the Fast marching algorithm in

road extraction, from high resolution (1-meter) panchromatic and (0.5-meter) multispectral satellite remote sensing images. An operator provides an initial seed point on the road of interest, then the region is extracted using a Fast Marching method.

## II. METHODOLOGY

In this paper, a semi-automatic road extraction method from high-resolution satellite images is performed, in which the extracted roads are converted into road vectors. The processing steps are shown in Fig. 1.

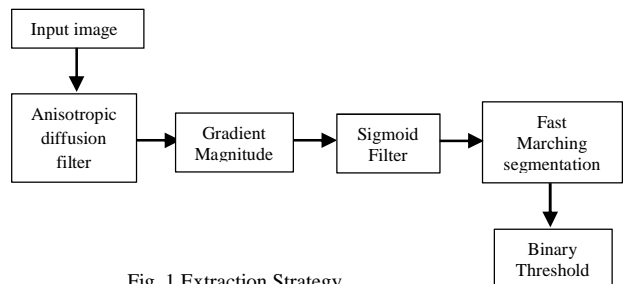


Fig. 1 Extraction Strategy.

### A. Anisotropic diffusion filtering

To preserve relevant discontinuities, while continuing to reduce noise, it became necessary to adjust the intensity of diffusion according to zone. It is for this purpose that Perona and Malik [6] proposed in the 90s, a model of nonlinear diffusion. They were the first to propose a model of non-linear diffusion in the two-dimensional case, to overcome the shortcomings of the isotropic diffusion. The authors postulated that the transitions belong to areas of higher gradient than those corresponding to noise. From there, they have sought to limit the diffusion when the gradient is high, as follows:

$$\frac{\partial u}{\partial t} = \text{div}(g(|\nabla u|)\nabla u) \quad (1)$$

The diffusivity is a decreasing function of the modulus of the gradient, called diffusion function. The  $g$  functions proposed by Perona and Malik are as follows:

$$g(|\nabla u|) = \exp\left(-\left(\frac{|\nabla u|}{k}\right)^2\right) \quad (\text{Gaussian function}) \quad (2)$$



And

$$g(|\nabla u|) = \frac{1}{1 + \left(\frac{|\nabla u|}{k}\right)^2} \quad (\text{Lorentzian function}) \quad (3)$$

The parameter K is called diffusion threshold. As its name suggests, it is used to fix the boundary between the high gradients corresponding to transitions to maintain, and low gradients corresponding to the noise.

Fig. 2(b) shows the result of Perona and Malik filter. The edges are retained while the noise is largely removed.

### B. Gradient magnitude filter

Differentiation is an ill-defined operation against digital data. Practically it is appropriate to define a scale in which the differentiation should be performed; it is done by preprocessing the data with a smoothing filter. It has been demonstrated that a Gaussian kernel is the most convenient choice for performing such smoothing operator [8]. By choosing a particular value for the standard deviation ( $\sigma$ ) of the Gaussian, The filter computes the magnitude of the image gradient at each pixel location. The computational process is equivalent to first smoothing the image by convolving it with a Gaussian kernel and then applying a differential operator [7] [10].

Fig. 2(c) presents the gradient magnitude of the smoothed image.

### C. Sigmoid filter

The Sigmoid filter is generally used as intensity transform. It maps a specific range of intensity values into a new intensity range by making a very smooth and continuous transition in the borders of the range. Sigmoids are widely used as a mechanism for focusing attention on a particular set of values and progressively attenuating the values outside that range [7] [8]. The following equation represents the sigmoid intensity transformation, applied pixel-wise:

$$I' = (A1 - A2) \cdot \frac{1}{1 + e^{-\left(\frac{I - \beta}{\alpha}\right)}} + A2 \quad (4)$$

The parameters are used to intensify the contrast between regions of low and high values in the speed image. The speed value should be 1.0 in the homogeneous regions (Ideal case) and the value should decrease rapidly to 0.0 around the edges of roads. From the gradient magnitude image, let's call A1 the minimum value along the contour of the road to be segmented, and A2 an average value of the gradient magnitude in the middle of the road. Those values indicate the dynamic range that we want to map to the interval [0,1] in the speed image. Our objective is that the sigmoid mapping A1 to 0.0 and A2 to 1.0. In knowledge that A1 should be higher than A2 and we want to map those values to 0.0 and 1.0 respectively, and we want to select a negative value for  $\alpha$  so that the sigmoid function will also do an inverse intensity mapping. This mapping will realize a speed image such that the level set will march rapidly on the homogeneous region and will definitely stop on the contour of road. The value for  $\beta$  is  $(A1+A2)/2$ , while the suggested value for  $\alpha$  is  $(A2-A1)/6$ , which must be a

negative number. The user may estimate these values by observing the gradient magnitude image [8] [10].

### D. Fast Marching Segmentation

Fast marching process is a numerical method for solving boundary value problems of the Eikonal equation:

$$F|\nabla T| = 1 \quad (5)$$

This problem describes the evolution of a closed curve as a function of time T, with speed F in the normal direction at a point on the curve. The speed function is defined, and the time at which the contour crosses a point is obtained by solving this equation. This problem is a particular case of level set methods. Other algorithms exist but are usually slower. [9]

Judicious approximation of the gradient is needed to solve the problem. The schema used by Sethian [9] is:

$$\text{Laying: } D_{ij}^{+x}T \equiv \frac{T_{i+1j} - T_{ij}}{\Delta x} \quad (6)$$

Taking:

$$\max(D_{ij}^{-x}T, -D_{ij}^{+x}T, 0)^2 + \max(D_{ij}^{-y}T, -D_{ij}^{+y}T, 0)^2 = \frac{1}{F_{ij}^2} \quad (7)$$

To reduce the cost of computations, a classification at the level of the pixels is made. Indeed, only the pixels in a neighborhood of the contour are interested. There is concern that pixels located in a narrow band around the interface. This is called the "Narrow Band". Classification of pixels is as follows:

1. The pixel type "frozen", pixels which  $T_{ij}$  value is known.
2. Neighboring pixels, are pixels of the "Narrow Band" which we have an estimation of the  $T_{ij}$  value.
3. "Far Away" pixels, for which  $T_{ij}$  value is not estimated.

It is limited in the following in the case were F is equal to 1, which is to build the field of distances of the points of the grid points of the initial condition. The principle of the method is as follows:

Starting from an initial condition, determine the distances field of the point. The initialization pixel is called "frozen" as previously established. We then calculate the distances to neighbors. The pixels which have been calculated the distance but are not "frozen" are pixels of the "Narrow Band".

For each iteration of the main loop, the pixel of the "Narrow Band" with the smallest distance becomes "frozen". Distances of its neighbors are again calculated: two cases to consider at this level of the algorithm.

1. Among neighbors, some were already affected by a distance value. It is therefore necessary to update these values.
2. Some neighbors were not included in the "Narrow Band". In this case, they should be added.

"Frozen" pixels are used to determine the values of the distances of the other pixels but never recalculated. We can see the "Fast Marching" method as a front of pixels of the "Narrow Band" propagating from the initial condition, which transforms, by progressing, the pixels of the "Narrow Band" in "frozen" pixels [9].

The contour will grow from a seed point provided by the user. The probability of segmenting a road where there is cars, shadows, without missing a large part of the road, will increase, if a large number of seeds can be passed. To reduce the time needed by the front to visit an entire road, the user must to provide a multiple seeds, and that contributes to reduce the risk of leaks on the edges of regions visited earlier. when segmenting an elongated road, the best strategy is to place several seeds along the axis of the road, to ensure that will be segmented by the expansion of the front, it will take shortly time to propagate to the other end of it. Level sets techniques have an important property to fuse implicitly several fronts in expansion. Using multiple seeds takes good advantage of this property [7] [10].

The mapping should be done in such a way that the propagation speed of the front will be very low close to high image gradients while it will move rather fast in low gradient areas. This agreement will make the contour propagate until it arrives at the edges of roads, and then slow down in front of those edges. The result of this Filter is a map indicating, how much time it would take for each pixel of the front to arrive at the pixel location.

Thresholding operation of the output image is equivalent to taking a snapshot of the contour at a particular time during its evolution [7] [10].

### III. EXPERIMENTAL RESULTS

We have tested the algorithm on several samples of IKONOS (Multispectral image 0.5-meter, and Panchromatic image 0.5-meter) images in urban, and Radar image in rural area. We must provide a primer point from which the contour will spread. We can actually provide a set of points, like shown in Fig. 2(d) and Fig. 3(d), in order to increase the probability of segmenting several separate roads or streets obstructed by objects such as cars, shadows etc. without losing parts of this road. The use of multiple seeds, can also reduce the time required for the fully visit a road curve and thus reduces the risk of leakage around the edges of the regions previously visited. The results obtained from the road extraction process is illustrated in Fig. 2(f), regarding a highly urbanized city zone, in which one can notice that the task is quite difficult, due to the high urbanization of the area, shadows and car occlusions.

Nevertheless, we can see that a large part of roads was extracted. Another result is illustrated in Fig. 3, in a rural zone; the proposed approach is able to correctly extract the roads in the image.

The same algorithm can be applied to radar images. In radar images, as roads usually present specular reflections, they also appear in darker color (Fig. 4(a)). These results are good despite of the noise, but radar images contain a multiplicative noise so the gradient filter is not the most adapted.

### IV. CONCLUSION

In this paper, we presented a semi-automatic road extraction algorithm based on the level set method, which incorporates the geometric characteristics of the road. The operator initializes the algorithm by providing a starting point. This extraction algorithm offers several advantages such as the ability to extract roads in different topology, and requires minimal operator intervention. The results showed that the proposed method can extract multiple roads with high efficiency, and the computational cost is generally a few milliseconds or less. We encountered a mediocre result in some parts, as the road is not well contrasted with its surroundings in urban scenes (visible change in intensity, presence of shadows of buildings, trees, cars, etc.) and it comes to the acquisition system, the age of the road and its coating.

### V. REFERENCES

- [1] J. Mena, "State of the art on automatic road extraction for GIS update: a novel classification," *Pattern Recognition Letters*, vol. 24, pp. 3037–3058, 2003.
- [2] I. Laptev, H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner, "Automatic extraction of roads from aerial images based on scale space and snakes," *Machine Vision and Applications*, vol. 12, pp. 23–31, 2000.
- [3] M. Rochery, I. Jermyn, and J. Zerubia, "Higher order active contours," *International Journal of Computer Vision*, vol. 69, pp. 27–42, Aug. 2006.
- [4] M. Barzohar and D. B. Cooper, "Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 707–721, July 1996.
- [5] M. Bicego, S. Dalfini, G. Vernazza, and V. Murino, "Automatic road extraction from aerial images by probabilistic contour tracking," in *IEEE International Conference on Image Processing, ICIP'03*, vol. 3, pp. 585–588, 2003.
- [6] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 12:629–639, 1990.
- [7] "The ITK software guide." <http://itk.org>, 2013.
- [8] J. Inglada and E. Christophe, "The Orfeo toolbox remote sensing image processing software," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS'09*, Cape Town, South Africa, jul 2009.
- [9] Sethian, J.A., A Fast Marching Level Set Method for Monotonically Advancing Fronts, *Proc. Nat. Acad. Sci.*, 93, 4, pp.1591-1595, 1996.
- [10] E. Christophe and J. Inglada, "Open Source Remote Sensing: Increasing the Usability of Cutting-Edge Algorithms," *IEEE Geoscience and Remote Sensing Newsletter*, pp. 9-15, mar. 2009

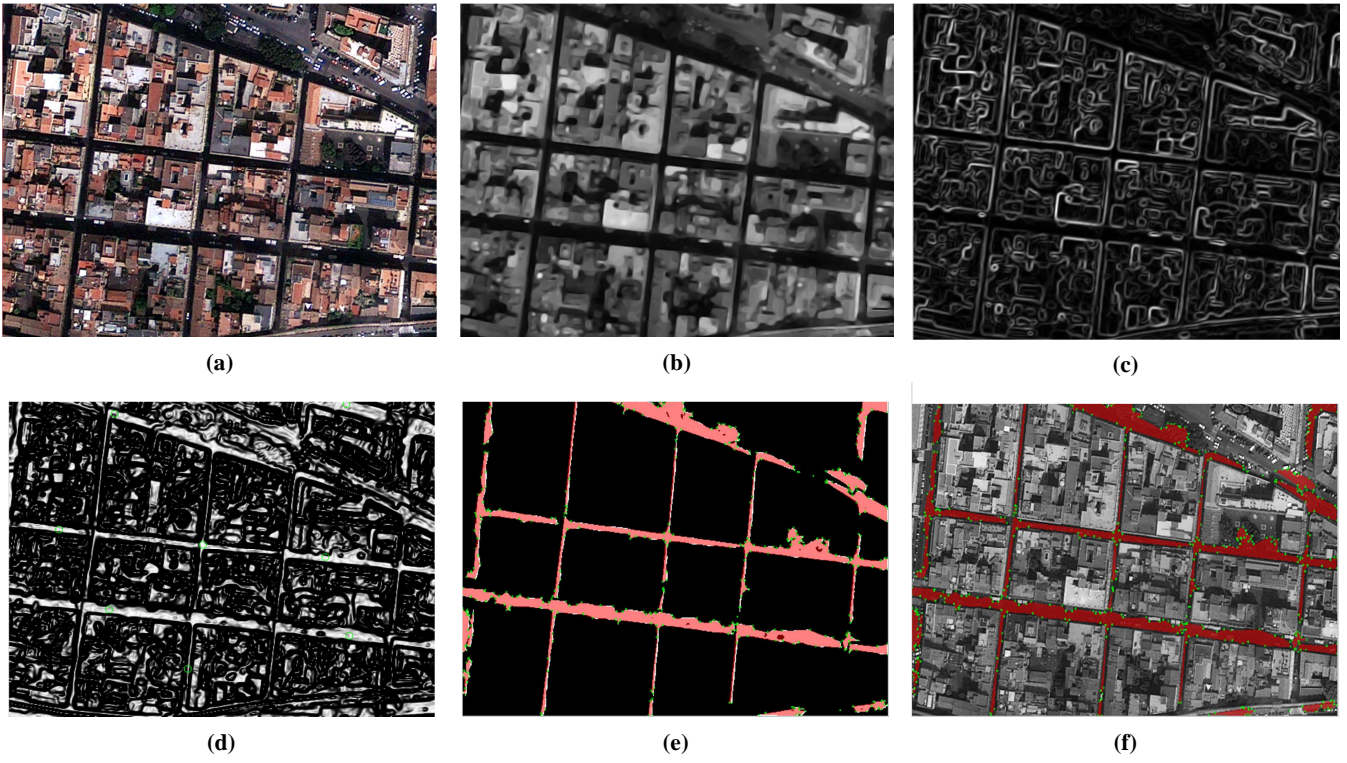


Fig. 2 Urban area: (a) Original image (0.5-meter), (b) Anisotropic diffusion filtering result, (c) Gradient magnitude, (d) Sigmoid of the gradient magnitude, (e) Result of the fast marching segmentation after thresholding, (f) Road extraction result.

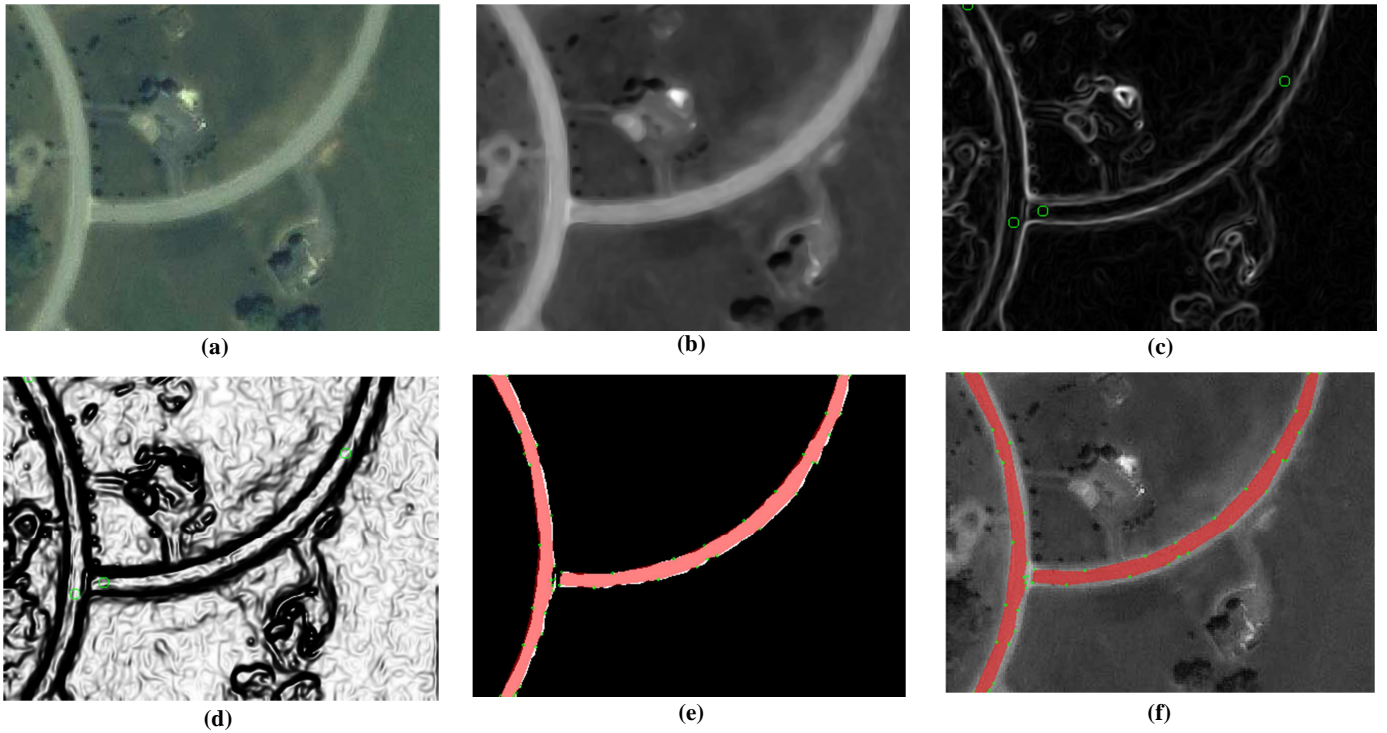


Fig. 3 Rural area: (a) Original image (1-meter), (b) Anisotropic diffusion filtering result, (c) Gradient magnitude, (d) Sigmoid of the gradient magnitude, (e) Result of the fast marching segmentation after thresholding, (f) Road extraction result.

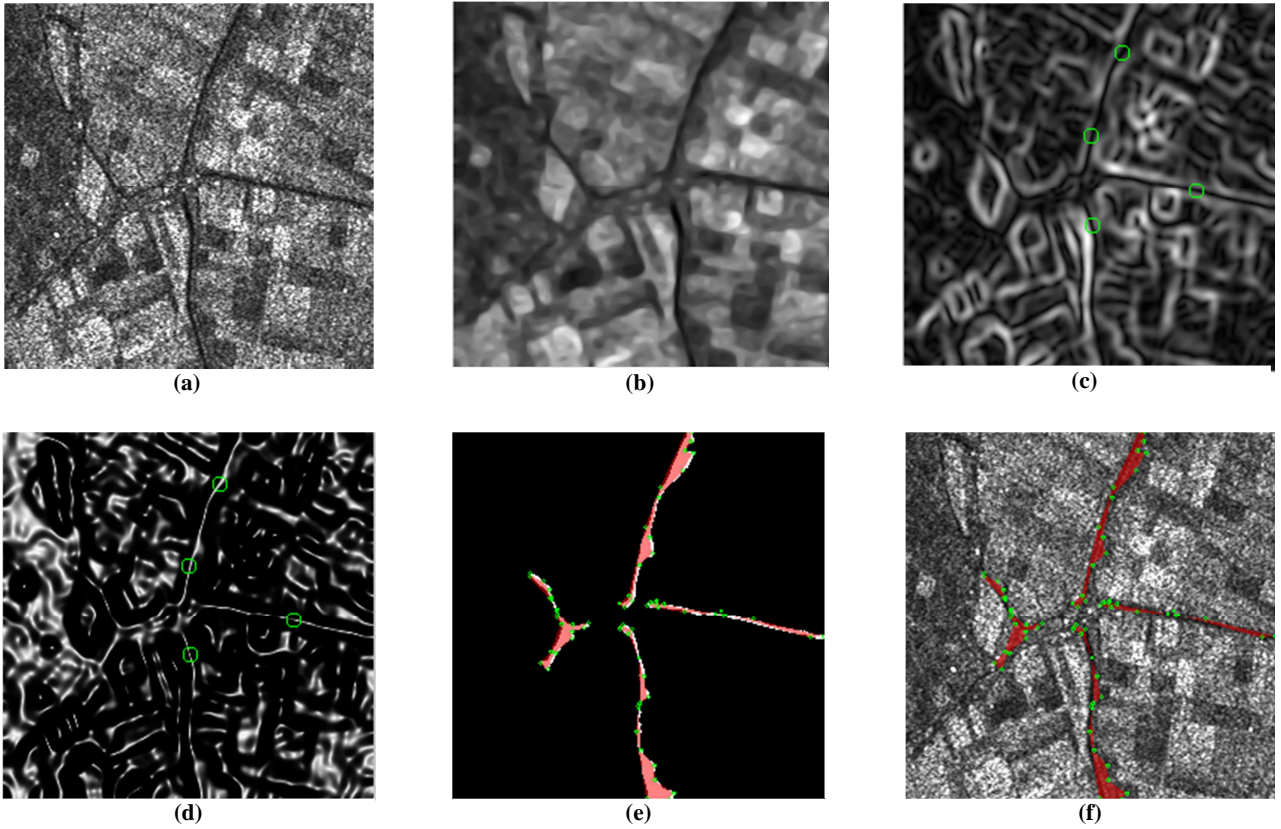


Fig. 4 Radar image: (a) Original image, (b) Anisotropic diffusion filtering result, (c) Gradient magnitude, (d) Sigmoid of the gradient magnitude, (e) Result of the fast marching segmentation after thresholding, (f) Road extraction result.

# Multi-scale approach for urban road extraction using remote sensing image

<sup>1</sup>Bendouda Malika <sup>2</sup>Berrached Nasreddine  
Intelligent Systems research Laboratory  
Electronic Dept, Electronic Engineering faculty  
University of sciences and technology Mohamed Boudiaf  
Oran, Algeria  
<sup>1</sup>vrf06bm@gmail.com  
<sup>2</sup>nasr1berrached@yahoo.fr

**Abstract**— The mapping of a city is essential for various applications such as planning, transport management, vehicle navigation and intervention in natural disasters. The rapidly changing urban environments increase the frequently updates and revisions of the road network databases. However, due to the extreme complexity of the urban environment, automating the extraction of the urban road network is one of the most challenging research topics in the field of photogrammetric and computer vision. Many methods exist for the extraction of road network in rural and suburban areas; but in urban environment the majority of these methods failed, due to the complexity of this environment.

In this article we propose a multiscale approach for the urban road network extraction; the principle is to switch from low-resolution satellite imagery to a high resolution satellite image to extract the entire urban road network.

**Keywords**—cartography; urban environment; multiscale; urban road network; low-resolution; high resolution; satellite image.

## I. INTRODUCTION

We use remote sensing images in order to extract the urban road network, this extraction is a very complex task because of the complexity of this environment, such as the presence of cars, trees and shadows of buildings [3], [5]. So that, many research works have been published in both geospatial and computer vision domains, and especially with the advent of remote sensing images of high resolution.

On low-resolution images, the work focused on the extraction of linear urban road network, Wang and his team have applied directional masks on remote sensing images to 10 m spatial resolution to detect and extract the streets represented by thin or thick lines [03]. Haverkamp proposed a semi automatic method to extract way of grid networks from IKONOS satellite images; he use both the information contained in the multispectral image and the panchromatic image [03]; the search of optimal path algorithms has been widely used for the extraction of roads, the optimization can be realized by dynamic programming [2], [5], active contours or by Kalman filtering [8].

On high-resolution images, Price use images of resolution from 0.20 to 0.80 m, the road network is modeled by a grid, after manual initialization of the first stitch, the grid is extended by success-sive iterations adding the other mesh. Couloigner extracted the two edges of street the image approximations at different spatial resolutions, the algorithm is applied on images of resolution between 0.85 and 10 m [3].

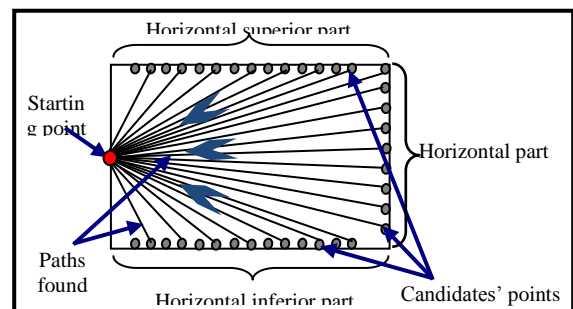
Other work head to the technical image fusion from multiple sensors from different sources such as satellite / radar [1], or injection of information in a low-resolution Multispectral image into a panchromatic high resolution image, such as the work of Y. Zhang, which merged the panchromatic image with the multispectral image to get a picture with a high spatial and spectral resolution[6]. Lee Sung Yiau used multispectral images MSI to properly highlight the urban road network, considering color as the original dis-nant for its shape recognition system [7].

## II. PROPOSED APPROACH

### A. The algorithm $F^*$ modified

The complexity of urban and sub-urban scenes (presence of trees along the road, a building occludes a part of a road, roofs might look similar to roads ...) makes the task of extracting the road network very difficult. Extraction throughout the global picture is not preferable because the result may be completely incoherent. Additionally to their robustness, road tracking methods, such as dynamic programming algorithms extract local segments belong to the road [2].

The algorithm  $F^*$  is a dynamic programming algorithm, from an initial point, detect the segments that belong to the road by optimizing certain criteria [5]. Given the complexity of urban areas we have improved the classical algorithm  $F^*$  to detect roads segments in each direction of the search window, as shown in figure.



The different steps of the algorithm F \* are described in the following chart:

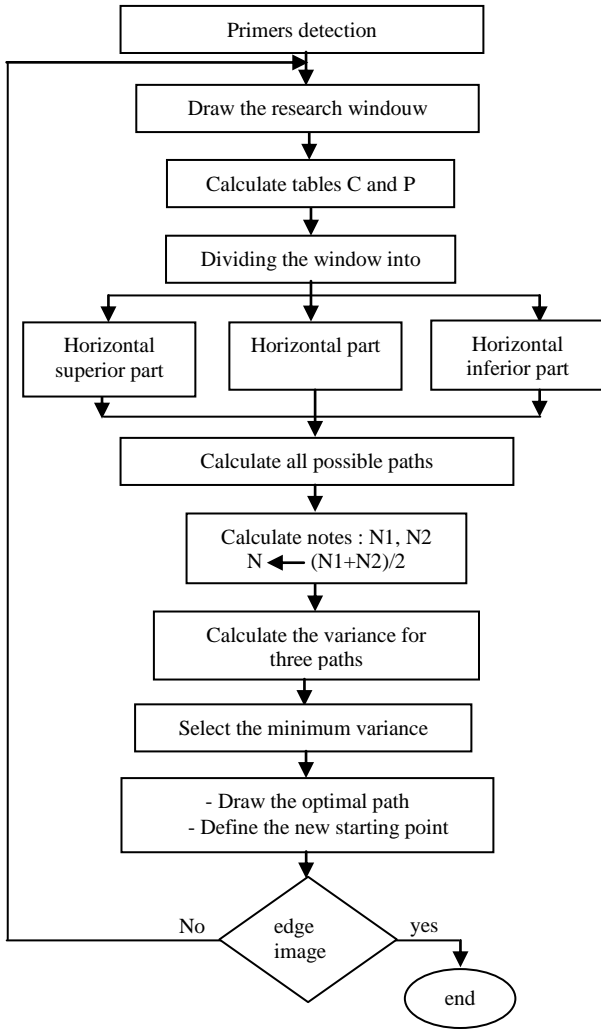


Fig. 2. →General organization of the F \* algorithm..

We set the number of lines in the search window to 21, and the columns to 50. The number of lines is odd because the starting point is in the middle of the search window to obtain ten candidates points above this line and ten other candidate points below this line [5]. These dimensions are large enough that the algorithm F \* Enhanced gives significant results, while being small enough to properly assess their validity.



Fig 3 →(a) Original image Riyadh SPOT4 10m resolution, (b) road network extracted.

We tested the modified algorithm F \* on low resolution satellite images, we were able to extract up to 85.58% of main roads (highway) to their visibility, and get a rate of 40, 08% for secondary roads as shown in Fig.3.

This confirms that the low-resolution images cannot extract the fine structure of the urban road network; hence high resolution satellite images must be utilized.

### B. Multispectral extraction

With the emergence of satellite images with high resolution ; surface extraction of the road network becomes approachable. The aim of this method is to classify the input image into two main classes, the first is the class of the object of interest 'roads' and the second one is the non-road class [7].

The principle of this method is illustrated by the following chart:

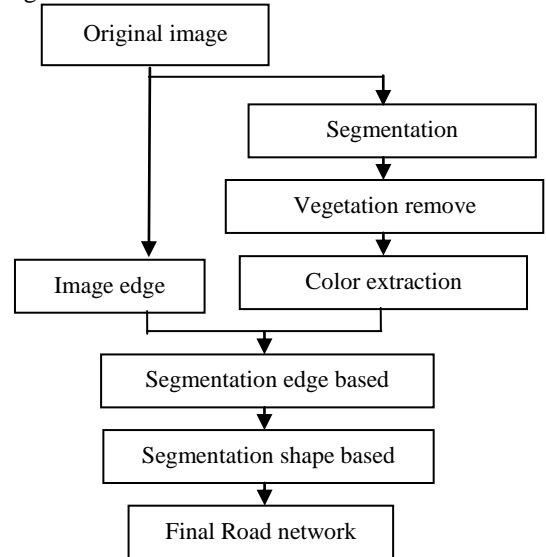


Fig. 1. General organization.

- *Segmentation*

The color image is converted to grayscale. Each pixel of that image is compared with the value of the three components so that the absolute difference of the pixel with these three components is less than a certain threshold, as described in the following equations [7].

$$|Ave\_Pixel\_value - Red\_pixel\_value| < Max\_colour\_diff \quad (1)$$

$$|Ave\_Pixel\_value - Green\_pixel\_value| < Max\_colour\_diff \quad (2)$$

$$|Ave\_Pixel\_value - Blue\_pixel\_value| < Max\_colour\_diff \quad (3)$$

- *Removal of vegetation*

In this step, we proceed to remove vegetation zones; such as trees planted along the road, and this through the use of color information in the spectral components, by applying equations (4) and (5):

$$Red\_pixel\_value < Green\_pixel\_value \quad (4)$$

$$Blue\_pixel\_value < Green\_pixel\_value \quad (5)$$

$$Output(i, j) = \begin{cases} 0 & \text{if equations 4, 5 verified} \\ Input(i, j) & \text{Otherwise} \end{cases} \quad (6)$$

we obtain image  $Output(i, j)$ , if a pixel has a value of  $Output(i, j) = 0$  it is not part of a road network.

- *Color extraction*

In this step, we will perform the result extraction, a high threshold and other low are calculated automatically.

The two thresholds are calculated using both the RGB values of the original color image and values of the image result of the Removal vegetation step.

As shown in the following equations. This choice allows us to have some measure of security.

$$S_h = Mean + StdDeviation \quad (7)$$

$$S_b = Mean - StdDeviation \quad (8)$$

Where :  $NG(I, j)$  is a grayscale image.

$$Mean = \frac{\sum_{i=1}^N \sum_{j=1}^M NG(i, j)}{Total\ No\ of\ pixel\ Greater\ than\ 0} \quad (9)$$

$$StdDeviation = \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M (input(i, j) - Mean)^2}{Total\ No\ of\ pixel\ Greater\ than\ 0}} \quad (10)$$

This method produces a mark improvement in the accuracy of road network extracted; because it use the color information present in the satellite images; to weed out an object that does not correspond to a road network. This method is quite successful in improving the accuracy of the extraction process as compared with the normal grey level

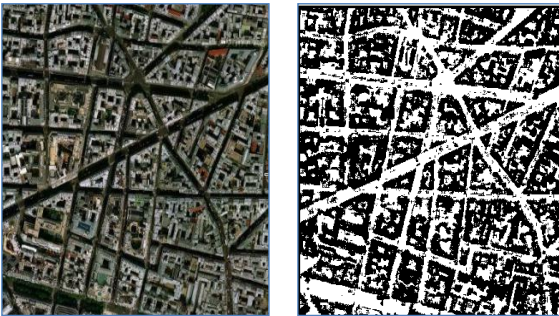


Fig. 2. Extracted road urban road network of Paris town

It is clear that almost all the road networks are correctly extracted. However, the rate of misclassification is high. For example, there are small family driveways connected to road networks, and many house roofs are classified into the road networks. These make it impossible to obtain an accurate road network without elimination noise processing.

- *Segmentation edge based*

The first type of noise to be eliminated is that of pepper and salt; filters are used to remove isolated pixels. We used a 5x5 filter to detect isolated pixels and remove them later.

After that, in order to eliminate object non road connected to road network, we proceed to segmentation based on the contour.

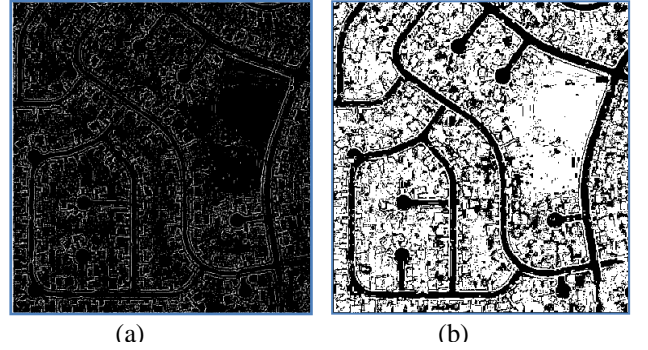


Fig. 3. (a) contour image detection, (b) sementation based edge

- *Segmentation shape based*

In this step, removes the small non-road groups of pixels, it is possible assuming that the roads are generally connected with a number of very important pixels compared to the other groups. We applied connected components analysis algorithm CCA ; the connectivity feature of roads that allows us to say that the roads components are usually connected with a large number of pixels compared to other components not road.

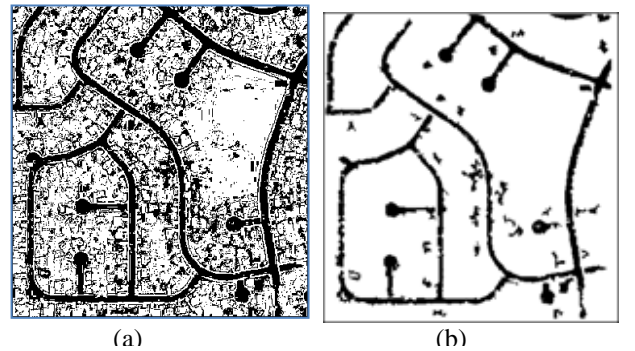


Fig. 4. (a) sementation based edge, (b) sementation based shape.

### III. ACCURACY ASSESSMENT

We calculate four factors to determine the accuracy of the extraction [09]:

**1. Correctness:** is the ratio between the length of extraction correctly extracted road network and the total length of roads extract.

**2. Completeness:** This measure gives us the precision of the extraction; it represents the ratio of the correct length of roads extract with the total length of the road network in the reference image.

**3. Omission:** the failure rate of pixels, which represents the pixels belonging to the road that we could not extract them.

**4. Commission:** the rate of pixels, which represents the pixels not belonging to the road that we extract as belonging to it .

These measures are calculated by the following equations:

$$Completeness = Nb\_correct / Nb\_ref \quad (11)$$

$$Correctness = Nb\_correct / Nb\_extract \quad (12)$$

$$Omission = Nb\_VN / Nb\_ref \quad (13)$$

$$Commission = Nb\_FP / Nb\_ref \quad (14)$$

Where :

Nb\_correct: number of pixels correctly extracted as road pixels.

Nb\_ref : number of pixels in the reference route.

Nb\_extract: total number of pixels extracted.

Nb\_VN: number of non-road pixels detected or missed

Nb\_FP: number of non-road pixels detected as road pixels.

The following table contain the results for the evaluation of the multispectral extraction road algorithm; we have achieved an average rate of extraction of up to 84.216% for sub-urban area and up to 66.275% urban area.

Image	Completeness	Correctness	Omission	Commission
Image1	66,05	61,05	05,06	07,47
Image2	94,84	53,12	07,14	10,36
Image3	84,11	50,56	04,75	14,69
Image4	90,10	22,19	08,55	33,55
Image5	85,98	85,50	04,76	02,59
Image6	61,19	17,42	08,32	32,46
Image7	57,32	33,88	05,31	19,00
Image8	52,15	36,33	06,00	14,00
Image9	94,44	46,65	02,62	31,84
Average	<b>76,24</b>	<b>45,18</b>	<b>05,83</b>	<b>18,44</b>

TABLE 1 → Evaluation of multispectral extraction results.

## IV. CONCLUSION

We propose a multi-scale approach to extract the urban road network, first we applied F\* modified algorithm on low resolution satellite image with road extraction up to 85.58% of main roads and rate of 40, 08% for secondary roads then we use high resolution satellite image to extract the fine structure of road network ; we implemented a method which integrate the spectral information to improve the extraction accuracy. We have provided in this method a noise elimination module in order to reduce the number of objects belonging to the road class and are not part of the road.

The experiments in road extraction from high resolution image demonstrate that the proposed approach is effective, the completeness and correctness of road extraction reaches 90,17% .

## References

- [1] Y. Hu, "Automated extraction of digital terrain models, roads and buildings using airborne lidar data", PhD dissertation, Department of Geomatics Engineering, University of Calgary, 2003.
- [2] R. Bonnefon & P. Dhérété, J. Desachy, "Automatic tracking of linear features on SPOT images using dynamic programming", Dans un *Proceedings d' Image and Signal Processing for Remote Sensing V*, Italie, 1999.
- [3] R.Péteri, "Extraction de réseaux de rues en milieu urbain à partir d'images satellites à très haute résolution spatiale", *Ecole des Mines de Paris*, 2003.
- [4] C. Lacoste, "Extraction de réseaux linéiques à partir d'images satellitaires et aériennes par processus ponctuels marqués", *INRIA Sophia Antipolis*, 2004.
- [5] R. Bonnefon, « Extraction d'objets cartographiques à partir d'images de télédétection : possibilité d'application à la mise à jour de Systèmes d'Information Géographique », *Université Paul Sabatier*.
- [6] Y. Zhang & R. Wang, "Multi-resolution and multi-spectral image fusion for urban object extraction", *Department of Geodesy and Geomatics Engineering, University of New Brunswick*, 2003
- [7] L.S Yiau, "Extracting Road Networks from High-Resolution Satellite Images", *School of Computer Engineering, Nanyang technological university*, 2003/04.
- [8] M.Rochery, "Contours actifs d'ordre supérieur et leur application à la détection de linéiques dans des images de télédétection", *université de Nice - Sophia Antipolis*, 2005.
- [9] C. Heipke, H. Mayer & C. Wiedemann, "Evaluation of automatic road extraction", *Technische Universit at Munchen*.1997.



# Méthodologie de mise en place de règles de connaissance pour la classification des objets sur des images à THRS

Habib MAHI<sup>#</sup>, Mounia KAOUADJI<sup>#</sup>

<sup>#</sup>*Division Observation de la Terre, Centre des Techniques Spatiales  
01, avenue de la Palestine, BP. 13, 31200 Arzew Algérie*

<sup>1</sup>hmahi@cts.asal.dz, mahihabib@yahoo.fr

<sup>2</sup>kaouadjimounia@yahoo.fr

**Résumé** — Avec l'avènement des images satellitaire à Très Haute Résolution Spatiale (THRS), l'analyse basée objets s'est imposée comme une méthode valide pour l'analyse de ce type de données. Bien que cette approche offre des outils efficaces pour représenter la connaissance d'une scène, les tâches de construction d'un réseau sémantique en sélectionnant les attributs sont coûteuses en temps. Ces procédés exigent au préalable l'intervention d'un expert pour caractériser les objets d'intérêt. Pour pallier à ce problème, nous proposons dans cet article, d'utiliser l'algorithme See5 pour la construction de règles de connaissance en sélectionnant les meilleurs attributs. L'objectif étant de montrer que l'introduction de ces règles peut aider à la classification des objets en un temps optimal.

**Mots clés** — images satellitaires à THRS, segmentation multi-résolution, règles de connaissances, classification basée objets.

## I. INTRODUCTION

La disponibilité des images à Très Haute Résolution Spatiale (THRS) a permis l'émergence de nouvelles applications en télédétection, telles que la cartographie à grande échelle, l'extraction des bâtiments, etc. Les résolutions sub-métriques offertes par les satellites civils comme Quickbird, Ikonos, Geoeye, Worldview et Pléiades ont permis une perception d'objet comparable à celle obtenue par les photographies aériennes. En effet, il est devenu possible d'identifier et d'analyser des espaces de petite superficie telles les parcelles agricoles, les bâtiments, les routes, les carrefours, etc., qui n'étaient pas détectables par les anciens capteurs. Cette résolution spatiale avantageuse permet d'extraire des caractéristiques géométriques des objets existants sur une scène à THRS et offre ainsi une information complémentaire à l'information spectrale et texturale. Néanmoins, l'augmentation de la résolution spatiale est synonyme d'augmentation de la variabilité spectrale des pixels formant les objets d'intérêt, ce qui entraîne une diminution de la précision de la classification spectrale basée pixels [1], [2]. La caractérisation des objets par l'utilisation de l'information spectrale est remplacée par l'analyse des formes et de l'arrangement des pixels internes : la texture. D'un autre côté, la fine résolution spatiale des données à THRS est accompagnée d'une résolution spectrale souvent inférieure à celle des autres capteurs, ce qui rend les méthodes de traitement purement spectrales inappropriées pour l'analyse de

ce nouveau type de données. L'intégration d'autres informations, telle l'information de texture, par exemple, s'avère donc indispensable. Ainsi, l'utilisation de la donnée à THRS nous amène à adopter des approches de traitement différentes de celles utilisées pour le traitement des données à moyenne et haute résolution spatiale. L'analyse basée objets a montré ses preuves concernant l'analyse de ces images, principalement dans les zones urbaines [3], [4], [5]. Parmi les logiciels commerciaux qui proposent ce type d'approche, on trouve eCognition® [6]. Il est largement utilisé par la communauté de télédétection vu la performance de ses résultats [7], [8], [9]. Néanmoins, la principale difficulté réside dans la construction d'une base de règles de connaissances qui entre en compte pour l'identification des objets géographiques due aux innombrables caractéristiques et rend ainsi la tâche de classification longue et complexe. Pour pallier ce problème, l'utilisation de techniques d'extraction de données sont nécessaires car elles sélectionnent rapidement les attributs les plus représentatifs de chaque classe et génèrent ainsi des règles de classification simples en un temps réduit. Dans le cadre de ce travail, nous avons généré une base de connaissance à l'aide de l'algorithme See5<sup>1</sup> [10].

Le reste de l'article est structuré de la manière suivante : dans la section 2, nous présentons les aspects théoriques de cette étude. Les résultats expérimentaux obtenus sur notre jeu de données à THRS sont abordés en section 3. Enfin, en section 4, nous terminons par une brève conclusion.

## II. METHODOLOGIE

La méthodologie adoptée dans le cadre de ce travail consiste en quatre étapes :

- Segmentation des données à THRS,
- Sélection des objets de référence et calcul des attributs,
- Extraction des règles en utilisant See5,
- Intégration des règles dans eCognition® et classification des données à THRS.

La Fig. 1 illustre l'organisation générale des étapes de traitements adoptées en vue d'une classification basée objets.

---

<sup>1</sup> Connu sous le nom de C5.0 pour Unix.

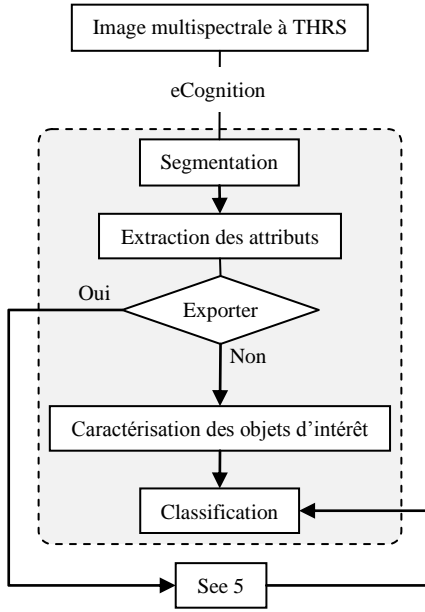


Fig. 1 schéma méthodologique de l'approche proposée.

### A. Segmentation Multi-résolution

Cette section décrit brièvement la segmentation multi-résolution proposée par Baatz et Schäpe [11] et utilisée par le logiciel eCognition®.

Le procédé consiste à fusionner les régions en appliquant la règle de minimum d'hétérogénéité. Dans un premier temps, chaque segment représente un seul pixel de l'image et les pixels sont associés à un certain segment. Les segments se développent dans la mesure où ils sont unis avec leurs voisins, et la plus faible augmentation de l'hétérogénéité est utilisée comme critère de sélection du voisin avec lequel un segment sera fixé. Pour simuler une croissance parallèle, chaque segment est sélectionné qu'une seule fois pour chaque itération. Le critère de fusion ( $f$ ) exprime l'augmentation de l'hétérogénéité résultant de l'union de deux segments. Avant une opération d'union, le critère de fusion est calculé pour chacun des voisins du segment sélectionné. Le voisin qui a le critère de fusion minimum est choisi pour la fusion. Toutefois, l'union ne se produit que si le critère de fusion est sous un certain seuil défini comme étant le paramètre d'échelle, noté  $\lambda$ . Cette procédure continue à fusionner les segments jusqu'à ce qu'il n'y ait plus aucune union possible.

Le critère de fusion  $f$  contient une composante d'hétérogénéité spectrale ( $h_{couleur}$ ) et une composante d'hétérogénéité spatiale ( $h_{forme}$ ), et est donnée par :

$$f = w_{couleur} \cdot \Delta h_{couleur} + w_{forme} \cdot \Delta h_{forme} \quad (1)$$

où  $w_{couleur}$  et  $w_{forme}$  sont les poids des propriétés spectrale et spatiale.

L'équation 2 montre la formulation de la différence d'hétérogénéité spectrale  $\Delta h_{couleur}$ ; où le segment sélectionné est noté  $obj1$ ,  $obj2$  est le voisin analysé et  $obj3$  est le résultat de la fusion entre  $obj2$  et  $obj1$ . Dans cette

équation,  $c$  est l'indice de la bande spectrale et  $w_c$  est un poids arbitraire associé à la bande  $c$ ;  $\sigma$  est l'écart type de l'objet dans la bande spectrale  $c$ ; et  $n$  est le nombre de pixels contenus dans un objet.

$$h_{couleur} = \sum_{c=1}^N w_c [n_{obj3} \sigma_{c,obj3} (n_{obj1} \sigma_{c,obj1} - n_{obj2} \sigma_{c,obj2})] \quad (2)$$

Sachant que  $N$  est le nombre de bandes spectrales utilisées.

La mesure de la différence d'hétérogénéité spatiale est influencée par deux composantes : la composante relative à la compacité  $h_{cmpct}$  et la composante de lissage  $h_{lissage}$ .

$$\Delta h_{forme} = w_{cmpct} \cdot \Delta h_{cmpct} + w_{lissage} \cdot \Delta h_{lissage} \quad (3)$$

avec

$$\Delta h_{cmpct} = n_{obj3} \frac{l_{obj3}}{\sqrt{n_{obj3}}} - \left( n_{obj1} \frac{l_{obj1}}{\sqrt{n_{obj1}}} + n_{obj2} \frac{l_{obj2}}{\sqrt{n_{obj2}}} \right) \quad (4)$$

$$\Delta h_{lissage} = n_{obj3} \frac{b_{obj3}}{\sqrt{b_{obj3}}} - \left( n_{obj1} \frac{b_{obj1}}{\sqrt{b_{obj1}}} + n_{obj2} \frac{b_{obj2}}{\sqrt{b_{obj2}}} \right) \quad (5)$$

où  $l$  et  $b$  représentent le périmètre de l'objet et le périmètre de sa boîte englobante.

### B. Attributs de Textures

La matrice de co-occurrence (GLCM pour Gray Level Co-occurrence Matrix en anglais) a été introduite par Haralick [12]. Elle est largement utilisée par la communauté de télédétection dans l'extraction des attributs texturaux [13], [14]. Elle s'intéresse aux relations qui existent entre les niveaux de gris des pixels de l'image pour un déplacement  $\vec{d}$  donné. Pour une image  $I$  de dimension  $N_x \times N_y$  ayant  $N_G$  niveaux de gris, la composante  $P(i, j, d, \theta)$  de la matrice de co-occurrence des niveaux de gris  $P$  avec un vecteur de séparation  $d = (d_x, d_y)$  et une direction  $\theta$  est donnée par :

$$P(i, j) = \text{Card}\{(x, y) : I(x, y) = i, I(x + d_x, y + d_y) = j\} \quad (6)$$

où  $\text{Card}$  représente la fonction cardinale,  $I(x, y)$  le niveau de gris de l'image  $I$  aux coordonnées  $(x, y)$  et  $i, j \in \{1, 2, \dots, N_G\}$ . Généralement, la valeur  $d$  est fixée à 1 et  $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ .

À partir de cette matrice, on extrait différents attributs appelés indices de texture du second ordre ou caractéristiques d'Haralick du nom de leur concepteur. Au total 14 indices ont été définis par Haralick.

### C. Attributs de Formes

Les descripteurs de forme doivent être invariants par rapport aux transformations géométriques usuelles telles la translation, la rotation et le changement d'échelle. Deux grandes familles de représentation de la forme existent : approche basée sur les contours et approche basée sur les régions. Pour la première approche, seule la frontière externe entre objets ou régions est exploitée, tandis que pour la seconde, la forme de toute la région est prise en considération. Quatre paramètres ont été sélectionnés, à savoir :

**La surface** : correspond au nombre de pixels qui composent l'objet en question.

**Le périmètre :** correspond au nombre de pixels qui composent la frontière de l'objet.

**L'indice de Miller (circularité):** Varie de 0 (forme linéaire) à 1 (forme circulaire). Il est défini par l'équation suivante :

$$I_{Miller} = \frac{4\pi S}{P^2} \quad (7)$$

où  $S$  et  $P$  représentent la surface et le périmètre respectivement.

**L'indice de solidité :** décrit la mesure dans laquelle la forme est convexe ( $I_{Solidité} = 1$ ) ou concave ( $I_{Solidité} = 0$ ) et est défini par :

$$I_{Solidité} = \frac{S}{H} \quad (8)$$

où  $S$  représente la surface de l'objet et  $H$  est l'enveloppe convexe de la forme.

### III. RÉSULTATS ET DISCUSSION

Dans le cadre de nos expérimentations, nous disposons d'un jeu de données à THRS de la ville d'Arzew acquis par le satellite Quickbird. L'image originale est constituée de quatre bandes spectrales (2.44 m de résolution spatiale) : bleu (0.45-0.52  $\mu\text{m}$ ), vert (0.52-0.60  $\mu\text{m}$ ), rouge (0.63-0.69  $\mu\text{m}$ ) et proche infrarouge (0.76-0.90  $\mu\text{m}$ ). Seules les trois premières bandes ont été utilisées. Ces dernières ont été fusionnées avec la bande panchromatique (0.61 m de résolution spatiale) afin d'obtenir une meilleure résolution spatiale (e.g. Fig 2).



Fig. 2 Composition colorée RVB de la zone d'étude

La première étape consiste à segmenter l'image en régions homogènes et connexes. La méthode de segmentation multi-résolution proposée par le logiciel eCognition® a été appliquée. Plusieurs valeurs du niveau d'échelle ( $\lambda$ ) ont été testées. La valeur du paramètre de forme ( $h_{forme}$ ) a été fixée à 0.1 afin de donner moins de poids à la caractéristique de forme et accorder plus d'attention à la caractéristique spectrale. La valeur du paramètre de compacité ( $h_{compct}$ ) a été elle aussi fixée à 0.5 pour donner le même poids à la fois au paramètre de compacité et celui du lissage des objets. Au final la segmentation multi-résolution avec un niveau d'échelle de 55 a été retenue.

La Fig. 3 illustre deux images de segments obtenues en fonction du niveau d'échelle. La première avec une échelle de 10 et la seconde avec une échelle de 55.



Fig. 3 Segmentation multi-résolution. Haut : Echelle 10. Bas : Echelle 55.

La seconde étape consiste à la sélection d'exemples d'apprentissage. Ceux-ci sont déterminés en choisissant des régions représentatives des classes (au nombre de sept) d'intérêts (e.g. Fig 4). Elles sont ensuite caractérisées selon leurs propriétés intrinsèques (spectrale, texturale, forme). L'ensemble d'apprentissage est formé de 253 objets.

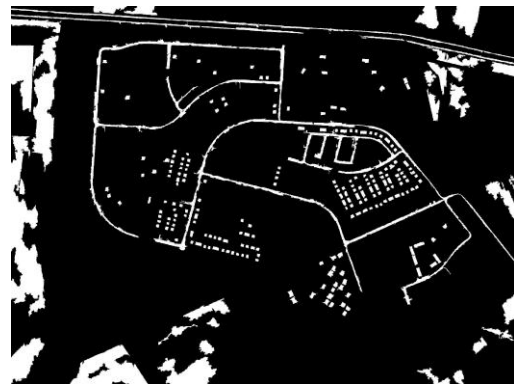


Fig. 4 Objets cartographiques retenus pour la phase d'apprentissage

Pour ce qui est de la caractérisation texturale des objets, nous avons retenu quatre indices texturaux de Haralick les moins corrélés, à savoir, l'homogénéité, la dissimilarité, l'entropie et le second moment angulaire. Chaque indice est extrait selon quatre directions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$ . Rappelons, que seul le canal vert a été utilisé dans les expérimentations, du fait qu'il est le moins corrélé des trois canaux. La Fig. 5 illustre quelques indices de texture calculés sur notre jeu de données satellitaires.

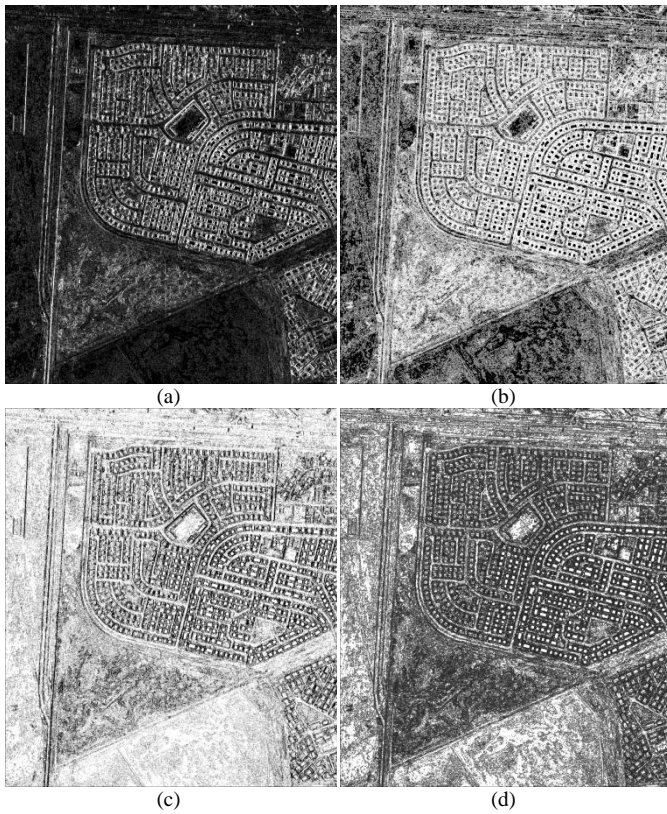


Fig. 5 Images de texture obtenues à partir du canal vert selon la direction  $\theta = 0^\circ$  (a) Dissimilarité, (b) Entropie, (c) Homogénéité et (d) 2<sup>nd</sup> Mom. Ang.

Le tableau I résume le nombre de caractéristiques utilisées pour l'extraction des règles de connaissances.

TABLE I  
CARACTERISTIQUES UTILISEES POUR L'EXTRACTION DES REGLES DE CONNAISSANCES

ID	Caractéristiques
1	Moyenne canal rouge (B1)
2	Moyenne canal vert (B2)
3	Moyenne canal bleu (B3)
4	Surface
5	Périmètre
6	Indice de Miller
7	Solidité
8	GLCM Homogénéité (0°) B2
9	GLCM Entropie (0°) B2
10	GLCM 2 <sup>nd</sup> Mom. Ang. (0°) B2
11	GLCM Dissimilarité (0°) B2
12	GLCM Homogénéité (45°) B2
13	GLCM Entropie (45°) B2
14	GLCM 2 <sup>nd</sup> Mom. Ang. (45°) B2
15	GLCM Dissimilarité (45°) B2
16	GLCM Homogénéité (90°) B2
17	GLCM Entropie (90°) B2
18	GLCM 2 <sup>nd</sup> Mom. Ang. (90°) B2
19	GLCM Dissimilarité (90°) B2
20	GLCM Homogénéité (135°) B2
21	GLCM Entropie (135°) B2
22	GLCM 2 <sup>nd</sup> Mom. Ang. (135°) B2
23	GLCM Dissimilarité (135°) B2

La troisième étape consiste à apprendre des règles de classification automatiquement à partir de ces régions caractérisées. Ces règles sont obtenues par l'application d'un algorithme d'apprentissage supervisé See5. Cet algorithme permet d'obtenir un modèle de classification représenté sous forme de règles de décision. Il associe à chaque classe un ou plusieurs attributs les plus représentatifs et pour chaque attribut est défini une valeur seuil. Les règles obtenues par l'algorithme See5 sont listées dans le tableau II.

TABLE II  
RESULTAT DES REGLES D'APPRENTISSAGE OBTENUES PAR L'ALGORITHME SEE5

Classes	Caractéristiques	Seuil
B. Blanc	Moyenne canal rouge (B1)	> 180.78
	GLCM 2 <sup>nd</sup> Mom. Ang. (135°) B2	> 0.06
B. Marron	Moyenne canal rouge (B1)	<= 180.78
	GLCM Entropie (0°) B2	> 0.17
	Surface	<= 602.00
B. Gris 1	Moyenne canal rouge (B1)	<= 180.78
	GLCM Entropie (0°) B2	<= 0.17
	GLCM Dissimilarité (45°) B2	> 0.08
B. Gris 2	GLCM Entropie (0°) B2	> 0.17
	GLCM Dissimilarité (45°) B2	> 0.08
	Surface	> 602.00
B. Beige	Moyenne canal rouge (B1)	> 180.78
	GLCM 2 <sup>nd</sup> Mom. Ang. (135°) B2	<= 0.06
Route	GLCM Dissimilarité (0°) B2	<= 0.37
	GLCM Dissimilarité (45°) B2	<= 0.08
	GLCM Dissimilarité (90°) B2	> 0.45
Sol nu	GLCM Dissimilarité (45°) B2	<= 0.08

La dernière étape consiste à importer ces règles sous eCognition® pour aider à la classification des objets en utilisant l'approche d'appartenance qui s'avère nettement meilleur que celle basée sur l'algorithme supervisé du Plus Proche Voisin (PPV) [15].



Fig. 6 Résultat de classification obtenue pour la classe bâtiment blanc

Les résultats montrent que la classification à base de règles de connaissances permet un bon classement de la classe bâtiments blanc (**98.59%** de bien classées). Notons aussi que les caractéristiques issues de l'algorithme See5 sont principalement des attributs de texture et la moyenne spectrale dans la bande rouge. Néanmoins, en ce qui concerne les autres classes, les résultats sont peu probants, cela est du principalement à la phase de segmentation.

#### IV. CONCLUSIONS

Ce travail vise à évaluer la façon dont l'algorithme See5 pourrait aider à la mise en place de règles de connaissance. En effet, les résultats de classification obtenue ont montré que l'algorithme See5 a été très efficace en ce qui concerne la sélection des attributs les plus représentatifs de la classe recherchée. Finalement, on peut dire que l'algorithme See5 intégré aux autres outils de l'analyse d'image comme la segmentation multi-résolution et les méthodes de classification a un grand potentiel en analyse des images satellitaire à THRS.

#### REFERENCES

- [1] A. P. Carleer, O. Debeir and E. Wolff, *Assessment of very high resolution satellite image segmentations*, Photogrammetric Engineering and Remote Sensing, vol. 71, issue 11, pp. 1284 – 1294, 2005.
- [2] P. Aplin, P. M. Atkinson and P. J. Curran, *Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom*, Remote Sensing of Environment, vol. 68, issue 3, pp. 206 – 216, 1999.
- [3] K. Navulur, *Multispectral Image Analysis Using the Object-Oriented Paradigm*, CRC Press Taylor & Francis Group, 204 pages, 2007.
- [4] S. Bhaskaran, S. Paramananda and M. Ramnarayan, *Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data*, Applied Geography, vol. 30, pp. 650 – 665, 2010.
- [5] N. Thomas, C. Hendrix and R. G. Congalton, *A comparison of urban mapping methods using high-resolution digital imagery*, Photogrammetric Engineering & Remote Sensing, vol. 69, issue 9, pp. 963 – 972, 2003.
- [6] eCognition Developer, *eCognition Developer 8.7 User Guide*, Munich, Germany: Trimble Germany GmbH, Trappentreustr, 2011b.
- [7] W. Wei, X. Chen and A. Ma, *Object-oriented Information Extraction and Application in High-resolution Remote Sensing Image*, IEEE, 2005.
- [8] J. Schiewe, *Segmentation of high-resolution remotely sensed data-concepts, applications and problems*, Symposium on geospatial Theory, Processing and Applications, Ottawa, 2002.
- [9] Th. Bauer and P. Strauss, *A rule-based image analysis approach for calculating residues and vegetation cover under field conditions*, CATENA, vol. 113, pp. 363 – 369, 2014.
- [10] J. R. Quinlan. (1997) See5. Site: <https://www.rulequest.com/see5-info.html>.
- [11] M. Baatz and A. Schäpe, *Multi-resolution segmentation – an optimization approach for high quality multi-scale segmentation*, (eds.): Angewandte Geographische Informationsverarbeitung XII. Wichmann, Heidelberg, pp. 12 – 23, 2000.
- [12] R. M. Haralick, K. Shanmugan and I. Dinstein, *Textural features for image classification*, IEEE Transactions System, Man and Cybernetics, vol. SMC-3, pp. 610 – 621, 1973.
- [13] J. Longhao, Z. Zhongfa and L. Bo, *Study of SAR Image Texture Feature Extraction Based on GLCM in Guizhou Karst Mountainous Region*, 2nd International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE), pp. 1 – 4, 2012.
- [14] G. R. Shahriari, A. Gharibi and A. R. Motlagh, *Classification of Satellite Images by Texture-Based Models Modulation Using MLP, SVM Neural Networks and Nero Fuzzy*, International Journal of Electronics and Electrical Engineering, vol 1, issue 4, pp. 245 – 251, 2013.
- [15] R. V. Platt and L. Rapoza, *An evaluation of an object-oriented paradigm for land use/land cover classification*, Professionnal Geographer, vol. 60, pp.87 – 100, 2008.

# A Classifier Ensemble Model based on Random Subspace and Static Classifiers Selection : Application on classification problems

Soraya Cheriguene, Nabihaz Azizi

Labged laboratory: Computer science department, Badji Mokhtar University of Annaba, PO box 12 , Annaba, 23000, Algeria  
[soraya\\_cheriguene@yahoo.fr](mailto:soraya_cheriguene@yahoo.fr), [azizi@labged.net](mailto:azizi@labged.net)

**Abstract**— In the modeling of Classifier ensembles systems, diversity is considered as one of the main aspects to be taken into account, since there is no gain in combining identical classification methods. Thus classifier selection became a crucial problem in ensemble learning. To select the best classifier set from a pool of classifiers, the classifier diversity is the most important property to be considered. One way of increasing diversity is to provide different data sets (patterns and/or attributes) for the individual classifiers such as Random Subspace. Therefore, in this paper, we propose an approach for ensemble construction based on Random Subspace method and static selection strategy. The experiments have been conducted using 6 benchmark data sets from the UCI machine learning repository. The experimental results confirm that our approach leads to better performance in many aspects.

**Keywords**— Machine learning; Ensemble classifier construction; Random subspace; static classifier selection; diversity measures

## I. INTRODUCTION

Classification, one of the most popular and significant tasks in Pattern Recognition, which is the main reason why the past few decades have seen a vast number of research projects devoted to classification methods applied to different fields of the human activity. The latest researches have been proposing the combination of multiple classifiers forming Ensembles as an attempt to improve the accuracy and efficiency of a classification system. [1-3]. During learning, the base classifiers are created separately on the data set, then the next step is to choose an effective way of combining the outputs of the classifiers to provide a decision on a test pattern. There are a great number of combination methods in the literature (also called fusion functions) including majority voting [4,5], BKS [6], Naive Bayes [7]...Etc.

The philosophy of the ensemble classifier is that different classifiers usually make different errors on different parts of the input space, which means that by combining classifiers we can create an ensemble that makes more accurate decisions. In order to have classifiers with different errors advisable to create diverse classifiers and group them into what is known as an Ensemble of Classifiers (EoC) [4].

There are several groups of methods to achieve diversity in an EoC. The largest set of ensembles generates ensemble classifiers by data resampling to obtain different samples from the training set to derive each individual classifier. Bagging and Boosting [8, 9] are two of the most popular ensemble methods based on data resampling. In another group of ensemble classifiers diversity among the base classifiers is achieved by ma-

nipulating the input feature space where the base classifiers are trained on different attribute subsets, i.e. Random Sub [10].

One of the most important issues surrounding EoC creation is classifier selection. Classifier selection has been proposed by Dasarathy and Sheela first [11]. The selection of a classifier can be either static or dynamic. In Static Selection, the classifiers that will form the ensemble are chosen at the optimization phase, once and for all and used to classify the test instances. In Dynamic Selection, the classifiers are chosen at the classification phase. That is, different testing instances can be classified by different ensemble configurations. In this study, a new approach is proposed based random subspaceMethod and static selection strategy forclassifier ensemble construction using homogenous learning algorithms. The subset of classifiers selected in the selection static procedure is combined using majority voting.

The paper is organized in the following sections. Section 2 explains the most popular methods of classifier ensemble constructions. In Section 3, details about the proposedmethodare presented.Then, the experiments and the results are presented in section 4. Section 5 closes the paper with some final remarks.

## II. ENSEMBLE CLASSIFIER SELECTION

As already mentioned, there is no accuracy gain in the ensemble composed of a set of identical classifiers. The ideal situation is a set of base classifiers with uncorrelated errors. In other words, the base classifiers should be diverse among themselves in order to improve the classification performances. In literature, there are different methods for creating diverse classifiers

### A. Bagging

Bagging was introduced by Breiman in 1996 [8] the word Bagging is a contraction of the words bootstrapping [12] and aggregating. Bootstrapping is based on random sampling with replacement; each classifier  $D_i$ ,  $i = 1; \dots; L$  in a bagging ensemble is trained using the same learning algorithm on different bootstrap samples of the original training set. Bootstrapping is less effective when the training sample size is large compared with the data dimensionality. Hence, in bagging, classifiers in the ensemble become more diverse when the number of training objects decreases. The final classification decision for an unlabeled pattern  $x$  is obtained by taking the majority vote of the ensemble. Figure 1 illustrates this method. Further study on the technique "bagging" was set up by ([13-

16]), these studies showed that, generally bagging is useful for the unstable classifiers such as neural network and classification trees. Unstable classifiers are sensitive to small alterations in the data set. Thus, training the same classifier model on two slightly different training sets might result in substantially different classifiers.

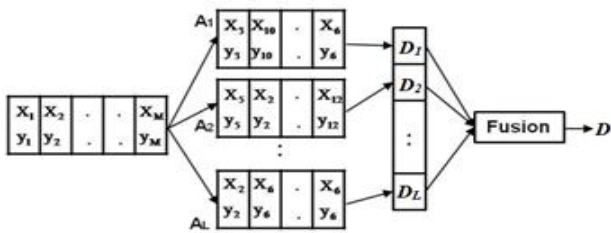


Figure 1. Schema of bagging method

### B. Boosting

Boosting, proposed by Freund and Schapire [9], is a procedure for combining the performance of weak classifiers in order to get a strong classifier with a better performance. Boosting is a deterministic procedure, it differs from bagging in those classifiers and training sets are generated sequentially, based on the results of the previous iteration (Figure 2). In contrast, bagging obtained the training sets and the classifiers in parallel and randomly. Boosting assigns a weight to each pattern in the training set, reflecting its importance, and constructs a classifier using the training set and the set of weights. Therefore, it requires a classifier that can handle weights on the training samples. Some classifiers may be unable to support weighted patterns. In this case, a subset of the training examples can be sampled according to the distribution of the weights and these examples used to train the classifier in the next stage of the iteration [17].

In the literature, there are many versions of Boosting. AdaBoost (Adaptive Boosting) is one of the popular algorithms; it is an iterative algorithm where the probability of samples selection is not the same for all. At each iteration, we increase the probability of being fired incorrectly classified instances. In this way, the difficult instances to classify are better represented and have more weight when learning.

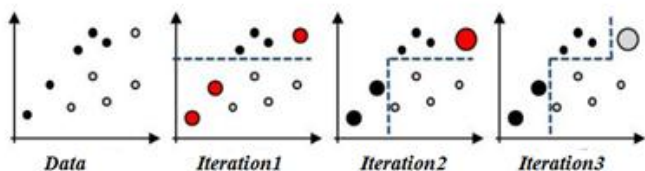


Figure 2. Learning and aggregation of classifiers with boosting method.

### C. Random Subspace

In problems with a large number of features, the number of training objects is small compared with the data dimensionality. To solve this problem Ho [10] developed new combining technique called "Random SubSpace Method" (RSS). This method has taken the lead in many application domains characterized by high-dimensional data. Examples include, but are not limited to, phenotype recognition [18], cancer diagnosis [19, 20], face recognition [21, 22] and credit scoring [23].

The random subspace method is quite similar to Bagging; the randomly selected for the used examples in the "Bagging" to build a set of classifiers is replaced by a random selection of the original attributes. The basic idea is to train each classifier on a subspace of attributes drawn randomly from the original base using all training examples. In [11], the author has shown that the best results are obtained when we using half of the feature components set. And she has also shown that the RSM is applicable to any type of classifier.

### D. Classifier Selection

The classifier selection mechanism is assigned to select best classifiers from a pool of different classifiers, so that the selected group of classifiers can achieve the optimum recognition rates. Currently, most of the existing classifier selection methods used for designing ensemble employ differ criterion. The diversity of the base classifiers plays an important role to choose the ensemble members [3]. Classifier selection can be divided into static classifier selection selects an EoC for all test patterns, and dynamic classifier selection selects different classifiers for different test patterns.

1) *Static Classifier Selection*: In static selection the best performing classifier ensemble is chosen during a training phase and used for the classification of unseen patterns. Another definition of static classifier selection can be found in the Neural Network literature. It is called either the Overproduce-and-Choose Strategy (OCS) [24] (also called "test and select" approach [25]) in order to design the MCS most appropriate for the task at hand. Methods based on OCS are divided into two phases: (1) overproduction; and (2) selection. In the first phase, an initial large number of base classifiers are constructed using the training data set. The second phase is devoted to generating and test different combinations of the initial classifiers in order to identify the best subset of classifiers that can be combined to achieve optimal accuracy.

2) *Dynamic Classifier Selection*: In dynamic classifier selection, the selection is done online, during classification, based on training performances and also various parameters of the actual unlabeled pattern to be classified. In [26], Woods et al proposed to select the single classifier that shows the best performance in the closest neighborhood defined by an arbitrarily set number of neighboring training samples. Giacinto and Roli [27], [28] enriched the selection criterion by incorporating classifier outputs produced during classification.

3) *Diversity measures*:As diversity are an important criteria used in classifier ensemble selection, researchers have proposed different indices for measuring it in an objective way. According to the taxonomy given by Kuncheva and Whitaker the diversity measures can be categorized into two types [29], pairwise measures and non-pairwise measures. Pairwise diversity measures calculate diversities between each pair of base classifiers, then the total diversity of the ensemble is computed with the average of these pairwise measures while the non-pairwise measures are calculated on the whole ensemble. In our approach, we have adopted correlation diversity measure

a) *Correlation between the errors*:Correlation between the classifiers errors is natural choice to compare the classifiers subsets:

$$\rho_{i,j} = \frac{N^{11}N^{00} - N^{01}N^{10}}{\sqrt{(N^{11} + N^{10})(N^{01} + N^{00})(N^{11} + N^{01})(N^{10} + N^{00})}} \quad (1)$$

Where  $N^{00}$  is the number of patterns that both classifiers wrongly classified; in contrast,  $N^{11}$  stands for the number of patterns that both classifiers correctly classified;  $N^{10}$  is the number of patterns classified correctly by classifier  $D_i$  but not by  $D_j$ ; likewise,  $N^{01}$  is the total of patterns classified correctly by classifier  $D_j$ , but not by  $D_i$ .

### III. PROPOSED APPROACH

The classification method proposed (RSS-SCS) in this paper is based first on the Random Subspace Method (RSS) which creates the initial pool of diverse classifiers by using different subsets of features to train classifiers and based secondly on Static Classifier Selection (SCS) using the diversity measure, in which the testing patterns are classified by the same ensemble configuration. Once a selection method is used to choose the ensemble members, the last step is to choose an effective way of combining the outputs of the ensemble members. Many methods can be used to combine the outputs of classifiers among which the majority voting is a most widely used method [4, 5]. It is in fact the simplest ensemble scheme: given a new pattern  $x$ , each classifier vote for a target class. The class that gets the highest number of votes is selected.

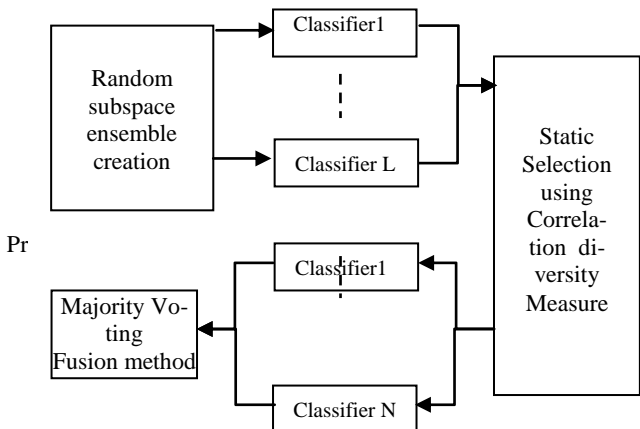


Figure 3. Summarizes the proposed approach.

### IV. EXPERIMENTS

In our experiments, six data sets for classification from the UCI machine learning repository [30] are used, each of which contained at least 150 instances. These six data sets are the Audiology, the Ionosphere, the Glass, the Sonar, the Vote, and the Iris. The details about these data sets can be found in Table 1. The training and testing data sets were extracted from each database using the ten-fold cross validation method. It is basically based on dividing ten equal parts of a data set. Any nine of the ten subsets (or segments) are selected to perform classifier training. The remaining part will be executed for evaluating diversity measure and testing the classifier.

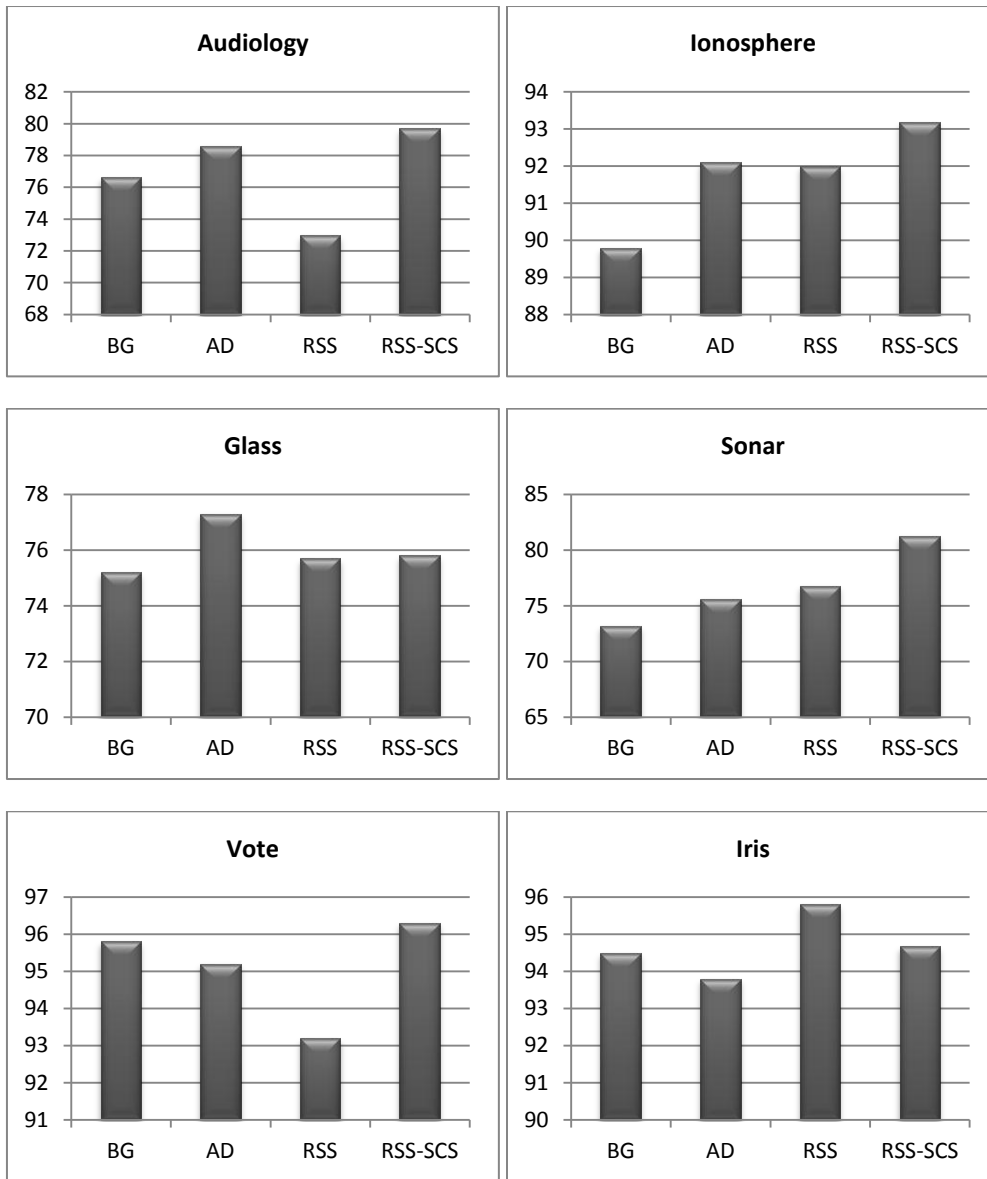
TABLE I. DATA SETS USED FOR CLASSIFICATION

Data Set	#Classes	#Features	#Objects
Audiology	24	69	226
Ionosphere	2	34	351
Glass	7	9	214
Sonar	2	60	208
Vote	2	16	435
Iris	3	4	150

In our approach (RSS-SCS) and for generating individual (base) classifiers, we used ten 10 pruned decision tree classifiers J48 (a Java implementation of C4.5) which are taken from the Waikato Environment for Knowledge Analysis (Weka) version 3.6.11. The parameters used for each algorithm in this empirical study were set at the default settings. To prevent overlearning and obtaining diversity between classifiers, each classifier was trained using randomly selected 70% of features from the training dataset (RSS-SCS).The proposed percentage has been determined experimentally. We have fixed  $M$  at 7 for selected set classifiers size. We also tested three best-known ensemble methods (Bagging (BG), AdaBoost (AD) and Random Subspace (RSS)) in order to compare the results of the classifier RSM-SCS. Each ensemble contained 10 decision tree (J48). Classification accuracies (i.e. the percentage of correctly classified examples) are shown in Table 2. The row "Average" contains results averaged over all datasets. For a better visualization, Fig. 4 depicts all results sorted by classification accuracy.

From the results given in the Table 2, some comments can be drawn. First, The RSS-SCS system outperformed the BG, AD and RSS classifiers by 2.63%, 1.38% and 2.40% respectively; We can also result that The RSS-SCS system achieved the highest classification accuracy for 4 datasets when the number of attributes is relatively large compared with the data dimensionality.





**Figure 4.** The best results using different learning methods in the six data sets.

TABLE II. CLASSIFICATION ACCURACIES OF THE EOCs. THE BEST RESULT FOR EACH DATABASE IS BOLDED.

Data Sets	BG	AD	RSS	RSS-SCS
Audiology	76.60	78.60	73.00	<b>79.68</b>
Ionosphere	89.80	92.10	92.00	<b>93.17</b>
Glass	75.20	<b>77.30</b>	75.70	75.80
Sonar	73.20	75.60	76.80	<b>81.28</b>
Vote	95.80	95.20	93.20	<b>96.31</b>
Iris	94.50	93.80	<b>95.80</b>	94.67
<b>Average</b>	84.18	85.43	84.41	<b>86.81</b>

## V. CONCLUSION

The classifier ensemble is widely considered to be an effective technique for improving accuracy of a classification system with numerous classifier components. In this paper, a new classifier ensemble construction method is proposed. Experiments were carried out on six real world problems from the UCI repository. We observed that RSS-SCS outperforms bagging, AdaBoost, and Random Subspace when the number of attributes is relatively large compared with the data dimensionality. This kind of result can be attributed several reasons: RSS effects based on random attribute selection used; SCS effects due to the selection of the most diverse ensemble using correlation measure effects. That is the proposed method complements the disadvantages of both random subspace and static classifier selection. In the future work, it is possible to use base classifiers other than J48 or using a heterogeneous ensemble of classifiers. The key point in using other types of classifiers is that they should produce vector outputs.

## REFERENCES

- [1] N.C. Oza, K. Tumer, "Classifier ensembles: select real-world applications", *Inf. Fus.*, 2008, pp. 4–20.
- [2] L. Rokach, "Taxonomy for characterizing ensemble methods in classification tasks: a review and annotated bibliography", *Comput. Stat. Data Anal.*, 2009, pp. 4046–4072.
- [3] L. I. Kuncheva, "Combining Pattern Classifiers: Methods and Algorithms". Wiley, 2004.
- [4] J. Kittler, M. Hatef, R. Duin, And J. Matas, "On combining classifiers". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, pp. 226–239.
- [5] D. Ruta And B. Gabrys, "Classifier selection for majority voting", *International Journal of Information Fusion*, 2005, pp. 63–81.
- [6] Y. S. Huang, & Suen, C. Y, "A method of combining multiple experts for the recognition of unconstrained handwritten numerals", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, pp. 90–93.
- [7] C. A. Shipp And L. I. Kuncheva, "Relationships between combination methods and measures of diversity in combining classifiers", *International Journal of Information Fusion*, 3(2), 2002, pp. 135–148.
- [8] L. Breiman, "Bagging predictors," *Machine Learning*, 1996, pp. 123–140.
- [9] Y. Freund , R. Schapire, "Experiments with a new boosting algorithm". *Proceedings 13th International Conference on Machine Learning*, 1996, pp. 148–156.
- [10] T. Ho, "The random subspace method for constructing decision forests," *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, 1998, pp. 832–844.
- [11] B.V. Dasarathy, B.V. Sheela, "A composite classifier system design: concepts and methodology", *Proceedings of the IEEE*, 1978, pp. 708–713.
- [12] B. Efron, R. Tibshirani, "An Introduction to the Bootstrap", Chapman & Hall, New York, 1993.
- [13] M. Skurichina , and R. Duin, Bagging, "Boosting and the Random Subspace Method for Linear Classifiers", *Pattern Analysis & Applications*, 2002, pp. 121–135.
- [14] L.I. Kuncheva, M. Skurichina b, R.P.W. Duin. « An experimental study on diversity for bagging and boosting with linear classifiers", *Information Fusion* 3, 2002, pp. 245–258.
- [15] Y. Grandvalet, "Bagging equalizes influence". *Machine Learning*, 2004, pp. 251–270.
- [16] L. Breiman, "Arcing classifiers", *Ann Statistics*, 1998, pp. 801–849.
- [17] A. R. Webb, "Statistical Pattern Recognition", Wiley, 2002.
- [18] B. Zhang ,and T. D. Pham, "Phenotype recognition with combined features and random subspace classifier ensemble", *BMC Bioinformatics*, 2011.
- [19] G. Armano, C. Chira, and N. Hatami. "A new gene selection method based on random subspace ensemble for microarray cancer classification". In: M. Loog, L. Wessels, M. J. T. Reinders, and D. Ridder, editors, *Pattern Recognition in Bioinformatics*, 2011, pp. 191–201.
- [20] A. Bertoni, R. Folgieri, and G. Valentini, "Random subspace ensembles for the biomolecular diagnosis of tumors". In: *Models and Metaphors from Biology to Bioinformatics Tools*, NETTAB, 2004.
- [21] N. V. Chawla and K. W. Bowyer. Random subspaces and subsampling for 2-d face recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (CVPR), 2005, pp. 582–589.
- [22] X. Wang, and X. Tang. "Random sampling for subspace face recognition". *International Journal of Computer Vision*, pp. 91–104, 2006.
- [23] G. Wang, J. Ma, L. Huang, and K. Xu. "Two credit scoring models based on dual strategy ensemble trees". *Knowledge-Based Systems*, pp. 61–68, 2012.
- [24] F. Roli, G. Giacinto, and G. Vernazza, "Methods for designing multiple classifier systems", In *Proceedings of MCS*, pp. 78–87, 2001.
- [25] A.J.C. Sharkey, and al, "The "test and select" approach to ensemble combination". In: Kittler J., Roli F. (eds.): *Multiple Classifier Systems*, 2000, pp. 30–44.
- [26] A. Bertoni, R. Folgieri, and G. Valentini, "Random subspace ensembles for the biomolecular diagnosis of tumors", In: *Models and Metaphors from Biology to Bioinformatics Tools*, NETTAB, 2004.
- [27] G. Giacinto and F. Roli, "Approach to the automatic design of multiple classifier systems," *Pattern Recognition Letters*, 2001, pp. 25–33.
- [28] L. Didaci, G. Giacinto, "Dynamic Classifier Selection by Adaptive k Nearest-Neighbourhood Rule", *Multiple Classifier Systems*, 2004, pp. 174–183.
- [29] L. I. Kuncheva, "That Elusive Diversity in Classifier Ensembles", In *Proceedings of the 1st Iberian Conference on Pattern Recognition and Image Analysis*, 2003, pp. 1126–1138.
- [30] C.L. Blake, C.J Merz, *UCI Repository of Machine Learning Databases* [<http://www.ics.uci.edu/~mllearn/MLRepository.html>]. Irvine, CA: University of California, Department of Information and Computer Science, 1998

# Discrete Composite Differential Evolution Algorithm: Practical Engineering Problems Optimization

F. Djeddou<sup>1\*</sup>, A. Hammoudi<sup>2\*</sup>,

<sup>\*</sup> *Applied Precision Mechanics Laboratory, Institute of Optics and Precision Mechanics, Setif -1- University, Algeria*

<sup>1</sup>auteur: djeddouferhat@yahoo.fr

<sup>2</sup>auteur: hammoudiabderazek@gmail.com

**Abstract**— Differential evolution (DE) is an efficient and robust evolutionary algorithm, which has been widely applied to solve global optimization problems. As we know, the choice of trial vector generation strategies and the control parameters values such as: crossover probability CR, scaling factor F, population size NP, plays a very important role on the performance of DE. In the present work, we applied an extended variant of DE (DCoDE: discrete composite DE). This method uses three trial vector generation strategies and three control parameter settings. It randomly combines them to generate trial vectors. To investigate the robustness of DCoDE in solving practical engineering problems highly constrained, we have chosen to optimize HCR (High Contact Ratio) gear. The objective function is to maximize the value of the contact ratio of a gear, where we consider tooth resistance to bending and pressure contact as well as issues of gear geometry such as interference and tooth head curvilinear thicknesses. This problem is however subjected to seven constraints nonlinear inequalities. The best values of the objective function was obtained for the low gear pressure angles.

**Keywords**— *Maximization, Discrete composite differential evolution algorithm, Mixed variable, High contact ratio HCR.*

## I. INTRODUCTION

Differential evolution (DE), proposed by Storn and Price [1], [2], is a very popular evolutionary algorithm (EA) and exhibits remarkable performance in a wide variety of problems from diverse fields. Like other EAs, DE is a population-based stochastic search technique. It uses mutation, crossover, and selection operators at each generation to move its population toward the global optimum. However, the DE performance mainly depends on two components. One is its trial vector generation strategy (i.e., mutation and crossover operators), and the other is its control parameters (i.e., population size NP, scaling factor F, and crossover control parameter CR) [3].

To improve the DE performance, several variants have been developed in recent decades. Some researchers are focused on the trial vector generation strategy. Indeed, Fan and Lampinen [4] proposed a trigonometric mutation operator that accelerates the DE convergence. New mutation strategies "current-to-best/1" and "current-to-pbest/1" have been developed respectively by Das et al. [5] and Zhang and Sanderson [6]. Other research has focused on optimizing the DE control parameters (NP, F, CR). Storn and Price [2]

suggested that the NP must be between 5D and 10D where D is the number of problem variables, F should be 0.5 and CR can take a value between 0.1 and 0.9. Ronkkonen et al. [7] suggested that NP is to be taken in the range [2D, 4D], F is to be chosen in the interval [0.4, 0.95] and CR should be between 0.2 and 1.0.

Dass et al. [8] introduced two schemes to adapt the scaling factor F in DE. The first scheme varies F in a random manner and the other one linearly reduces the value of F from a preset maximal value to a minimal one. Yong Wang et al [1] developed a new method called composite DE (CoDE), it combines three trial vector generation strategies "rand /1/ bin, rand/2/bin and current-to-rand /1" with three values of F and CR (F = [1 1 0.8] and CR = [0.1 0.9 0.2]) they are chosen randomly. The performance of the CoDE to solve optimization problems has been tested only by the CEC2005 contest test instances.

This work is focused on mixed variable structural optimization problem as dimensions of a HCR gear using the DCoDE. The three mutation schemes used to generate the trial vectors are: "rand/1/bin, rand/2/bin, and rand/best/2", the parameter values are: F = [0.5 0.3 0.1] and CR = [0.9 0.8 0.7]. The objective function is to maximize the value of the gear contact ratio. Profile shift coefficients ( $x_1, x_2$ ), addendum coefficients ( $h_{a1}^*, h_{a2}^*$ ), and the real pressure angle ( $\alpha$ ) are then decision variables to be optimized. This issue is subjected to seven non-linear inequalities constraints.

This manuscript is organized as follows. The second section will describe the principle and operation of the used DCoDE. In the third section the mathematical model of optimization problem with a gear HRC will be formulated, and the fourth section will be reserved for numerical results. We end this paper with a conclusion on the method used.

## II. DCoDE

The DE is an iterative optimization algorithm that follows the classical pattern of evolutionary algorithms. It can be described in four phases: initialization, mutation, crossover and selection.

The main idea of the DCoDE (Fig.1) used in this work, is to combine the three most popular DE schemes mutation "rand/1/bin, rand/2/bin and best/2/bin" with the three values recommended in previous studies of scaling factor, F = [0.5 0.3 0.1] and CR = [0.9 0.8 0.7], to generate trial

individuals of the population vectors for each generation. This strategy ensures diversification competent solutions to achieve the best values of the objective function. The best individual is selected for the next iteration of the population, i.e. if the value of the corresponding objective function is the best in the three diagrams.

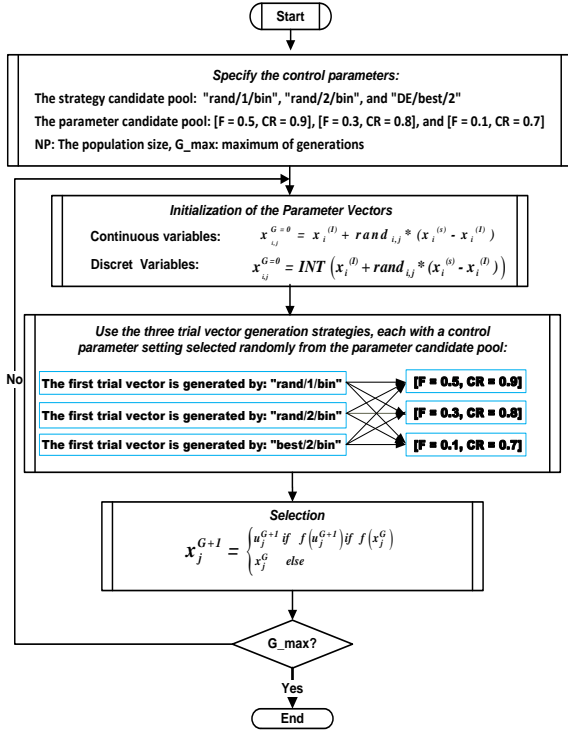


Fig. 1 DCoDE diagram

### A. Handling of Discrete Variables

To treat this type of variable (discrete), the idea is to consider as a variable, the variable index or its position in the array or vector of normalized values. If the index is created for example 2, then this is the second normalized value that will be taken. The indexes are created randomly according to the following expression:

$$x_{i,j}^{G=0} = INT(x_i^{(L)} + rand_{i,j} * (x_i^{(U)} - x_i^{(L)})) \quad (1)$$

Where,  $INT$  is a function that transforms the real into the nearest integer value of  $x_{i,j}^{G=0}$ ,  $x_i^{(L)}$  and  $x_i^{(U)}$  are respectively the lower and upper limits of the  $i^{th}$  variable problem. The operator  $rand_{i,j}[0,1]$  can create a random value uniformly distributed in the interval  $[0,1]$ .

### B. Treatment of Constraints

Any value found unfeasible for the mutant vector, i.e.  $v_{i,j}^{G+1} \notin [x_i^{(L)}, x_i^{(U)}]$ , is replaced by a value generated as:

$$v_{i,j}^{G+1} = \begin{cases} \min\{x_i^{(L)}, 2x_i^{(U)} - u_{i,j}^{G+1}\} & \text{if } v_{i,j}^{G+1} < x_i^{(L)} \\ \max\{x_i^{(L)}, 2x_i^{(U)} - u_{i,j}^{G+1}\} & \text{if } v_{i,j}^{G+1} > x_i^{(U)} \end{cases} \quad (2)$$

Static penalty method was used to solve this problem. It transforms the problem-solving constraints with an unconstrained problem by creating penalized as follows function:

$$F(x) = f(x) + \lambda_j \sum_{j=1}^4 \langle g_j(x)^2 \rangle \quad (3)$$

Where  $f$  is the objective function,  $g_j$  the inequality constraints and  $\lambda_j$  are the penalty coefficients that differ from one to another.

## III. OPTIMIZATION OF HCR GEAR GEOMETRY

The research efforts in recent decades have led to progress in the control of gearing degradation mechanisms. Technical progress resulted in an improvement in their behavior in service and a significant increase in power density [9]. In recent years, the changes in customer demand and legislation sound environment led to the consideration of the noise radiated by the gear drives as a quality criterion, especially in applications of cylindrical spur gears [10]. Several studies have demonstrated that the low value of the contact ratio of a gear is the main source of radiated noise. For this reason, the use of HCR involute gears is an important key to reduce levels of radiated noise and also increase the load capacity of a gear. This type of gears is generally obtained by introducing changes in either the pressure angle, or at addendum teeth level [11].

For correct operation without interrupting the movement of the transmission, the minimal value of the gear contact ratio must be greater than 1. In practice, in order to obtain a stable transmission and reducing vibrations, contact ratio should be greater than 1.3. Increasing the value of this contact ratio can be provided either by reducing the pressure angle, or by increasing addendum height [10].

In this work we tried to combine these two methods to get the best performance among the different gears. In this case, it should be mentioned that we do not limit the study only on a gear at  $20^\circ$  pressure angle.

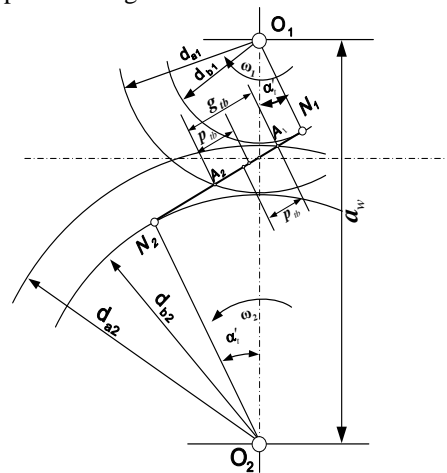


Fig. 2 Geometrical parameters of an external gear

- The optimization variables chosen are:

$$X = (x_1, x_2, h_{a1}^*, h_{a2}^*, \alpha)$$

- The cost function is to maximize the contact ratio. It is expressed as follows:

$$f(X) = \frac{g_{tb}}{P_{tb}} = \frac{\sqrt{r_{a1}^2 - r_{b1}^2} + \sqrt{r_{a2}^2 - r_{b2}^2} - a_w \sin \alpha_w}{\pi m \cos \alpha_w} \quad (4)$$

Where:

- $r_{a1}$  and  $r_{a2}$  are respectively tip diameters of pinion and wheel. They may be calculated as:

$$r_{a1} = r_1 + (h_{a1}^* \cdot m + x_1 \cdot m) \quad (5)$$

$$r_{a2} = r_2 + (h_{a2}^* \cdot m + x_2 \cdot m) \quad (6)$$

- $h_a^*$  is the addendum coefficient.
- $a_w$  and  $\alpha_w$  are respectively the center distance and the operating pressure angle.
- $\alpha$  is the pressure angle.
- $m$  is the gear module.

- In this case the objective function is subjected to the following constraints:

$$g_1(X) = \frac{2 - z_1 \cdot \sin^2 \alpha}{2} \cdot h_{a1}^* - x_1 \leq 0 \quad (7)$$

$$g_2(X) = \frac{2 - z_2 \cdot \sin^2 \alpha}{2} \cdot h_{a2}^* - x_2 \leq 0 \quad (8)$$

$$g_3(X) = \frac{m \cdot (z_2 + 2 + 2x_2)}{2} - \sqrt{h_{b2}^2 + (a_w \cdot \sin \alpha_w)^2} < 0 \quad (9)$$

$$g_4(X) = \frac{m \cdot (z_1 + 2 + 2x_1)}{2} - \sqrt{h_{b1}^2 + (a_w \cdot \sin \alpha_w)^2} < 0 \quad (10)$$

$$g_5(X) = 0.2m - s_{a1} \leq 0 \quad (11)$$

$$g_6(X) = 0.2m - s_{a2} \leq 0 \quad (12)$$

$$g_7(X) = |g_{s1max} - g_{s2max}| \leq 3 \quad (13)$$

Where:  $s_{a1}$  and  $s_{a2}$  are respectively are curved head thickness of the pinion and the wheel.  $g_{s1max}$  and  $g_{s2max}$  are respectively denoting the specific sliding of pinion and wheel at the end contact points.

- The limits of the problem variables are:

$$-0.7 \leq x_1 \leq 1, \quad -0.7 \leq x_2 \leq 1, \quad 1 \leq h_{a1}^* \leq 2, \quad 1 \leq h_{a2}^* \leq 2$$

- And  $\alpha$  may take one of the discrete values listed in Table I.

TABLE I NORMALIZED VALUES OF THE PRESSURE ANGLE

$\alpha$							
14	14°30'	15	16°30'	17	17°30'	18	18°30'
20	20°30'	21	21°30'	22	22°30'	23	23°30'
24°30'	25	25°30'	26	26°30'	27	27°30'	28

## IV. RESULTS AND DISCUSSION

For optimization of HCR geometry, the following data is used: number of pinion teeth  $z_1 = 80$ , number of gear wheel teeth  $z_2 = 121$ , module  $m = 4 \text{ mm}$ , helix angle  $\beta = 0^\circ$ , face width  $b = 175 \text{ mm}$ , rotational wheel speed  $n_1 = 4.1596 \text{ min}^{-1}$  and wheel torque  $T_1 = 1263 \text{ N.m}$ . The gear material was steel 18CrNiMo7-6. In order to find the value of  $\alpha_w$  and  $a_w$  for each value of  $x_1$  and  $x_2$  the Newton-Raphson method was used.

The control parameters of the DCoDE and DE classical considered in this case study are: population size  $NP = 100$  and the maximum number of generations  $G_{max} = 200$ . For DCoDE, the scaling factor  $F$  and the probability  $CR$  randomly take one of the following three values  $F = [0.5 \ 0.3 \ 0.1]$ ,  $CR = [0.9 \ 0.8 \ 0.7]$ , and  $F = 0.5$ ,  $CR = 1$  for the DE classical. The evaluation of the objective function using both algorithms is shown in Fig.3.

Table 3 shows the results for the case considered in this study. The maximum value of the objective function (contact ratio) is  $\varepsilon_\alpha = 3.55$  using the DCoDE (for  $x_1 = 0.7338$ ,  $x_2 = 0.4159$ ,  $h_{a1}^* = 1.4049$ ,  $h_{a2}^* = 1.460$ ,  $\alpha = 14^\circ 30'$ ) and  $\varepsilon_\alpha = 2.79$  using the DE classical (for  $x_1 = 0.4574$ ,  $x_2 = 0.4867$ ,  $h_{a1}^* = 1.2739$ ,  $h_{a2}^* = 1.173$ ,  $\alpha = 19^\circ$ ). It is clear from these results that the DCoDE is more powerful than the DE classical.

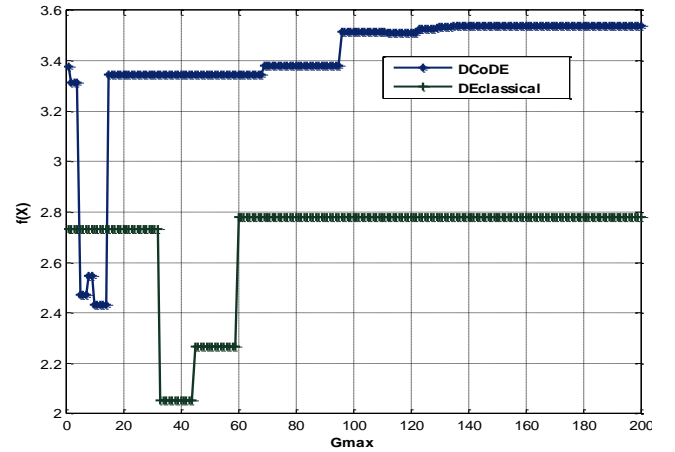


Fig.3 Function objective evaluation by declassical and DCoDE

TABLE II OPTIMIZATION RESULTS

Variables	DCoDE	DE Classical
$x_1$	0.7338	0.4574
$x_2$	0.4159	0.4867
$h_{a1}^*$	1.4049	1.2739
$h_{a2}^*$	1.4600	1.1736
$\alpha(^{\circ})$	14°30'	19
$g_1$	-48.450	-0.3282
$g_2$	-76.16	-0.9081
$g_3$	-9.855	-18.36
$g_4$	-23.84	-38.96
$g_5$	-0.0001	-0.0003
$g_6$	-0.0086	-0.0006
$g_7$	-2.82	-2.946
$\varepsilon_\alpha$	3.5500	2.7990

## V. CONCLUSIONS

In this article, mixed variable structural optimization problem «HCR Gearing Geometry Optimization» using DCoDE has been resolved. A mathematical model was established with five design variables and seven inequality constraints, the objective function to maximize is the contact ratio of a gear transmission. The optimization results indicate that DCoDE is more efficient than the DE classical. It is however worth to mention that best values of the objective function were obtained for low gear pressure angles.

## REFERENCES

- [1] R. Storn and K. Price, "Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces," Int. Comput. Sci. Inst., Berkeley, CA, Tech. Rep. TR-95-012, 1995.
- [2] R. Storn and K. V. Price, "Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.
- [3] Y. Wang, Z. Cai, and Q. Zhang, "Differential Evolution with Composite Trial Vector Generation Strategies and Control Parameters," *IEEE Trans. Evol. Comput.*, VOL. 15, n0. 1, pp.55-66, Feb. 2011.
- [4] H. Y. Fan and J. Lampinen, "A trigonometric mutation operator to differential evolution," *J. Global Optim.*, vol. 27, no. 1, pp. 105–129, 2003.
- [5] S. Das, A. Abraham, U. K. Chakraborty, and A. Konar, "Differential evolution using a neighborhood-based mutation operator," *IEEE Trans. Evol. Comput.*, vol. 13, no. 3, pp. 526–553, Jun. 2009.
- [6] J. Zhang and A. C. Sanderson, "JADE: Adaptive differential evolution with optional external archive," *IEEE Trans. Evol. Comput.*, vol. 13, no. 5, pp. 945–958, Oct. 2009.
- [7] J. Ronkkonen, S. Kukkonen, and K. V. Price, "Real parameter optimization with differential evolution," in Proc. IEEE CEC, vol. 1. 2005, pp. 506–513.
- [8] S. Das, A. Konar, and U. K. Chakraborty, "Two improved differential evolution schemes for faster global search," in Proc. GECCO, Jun. 2005, pp. 991–998.
- [9] J. Bruyère and P.Velex, "A simplified multi-objective analysis of optimum profile modifications in spur and helical gears", *Mechanism and Machine Theory*, vol. 80, pp. 925-935, 2014
- [10] G. Henriot, *Engrenages conception fabrication mise en œuvre*, 8ed, Ed. Dunod, Paris, France, 2008.
- [11] M. Vereš, Z. Kanovic and M. Rackov, "HCR Gearing Geometry Optimization by Using of Generalized Particle Swarm Algorithm", In *Modern Methods of Construction Design*, Springer International Publishing Switzerland 2014, pp. 539–545.

# Utilisation de l'Algorithme de Recherche Coucou pour le Regroupement des Données de Télédétection

Kaouter LABED<sup>#</sup>, Habib Mahi<sup>\*</sup>, Hadria FIZAZI<sup>#</sup>

<sup>#</sup>Faculté d'Informatique

Université Mohamed Boudiaf – USTOMB, Oran, Algérie

<sup>1</sup>kaouter.labeled@univ-usto.dz, klabeled@gmail.com

<sup>3</sup>fizazi@univ-usto.dz

<sup>\*</sup>Division Observation de la Terre

Centre des Techniques Spatiale, Arzew, Oran, Algérie

<sup>2</sup>hmahi@cts.asal.dz, mahihabib@yahoo.fr

**Résumé**—L'algorithme de recherche coucou (ARC) est un algorithme d'optimisation bio-inspiré qui simule la stratégie de reproduction agressive de certaines espèces d'oiseaux appelés Coucou (Cuckoo en anglais). La classification de données est un élément essentiel et indispensable de tout système d'analyse de données. Les algorithmes bio-inspirés sont considérés comme des bons outils de classification et peuvent être utilisés en conjonction avec les indices d'évaluation.

**Mots clés**— Algorithme Coucou, Classification, XB, DB, PBMF.

## 1. INTRODUCTION

La classification de données vise à effectuer des assemblages de points homogènes dans un nuage de données. L'image satellitaire est considérée comme un nuage de pixels, dont le regroupement forme les différentes classes existantes dans l'image satellitaire tel que la Mer, la forêt ou le sol nu. Les regroupements de pixels peuvent être guidés par des informations sur les données fournies au préalable, ce genre d'approche est connue sous le nom de classification supervisée. A l'opposé, lorsqu'il n'y a aucune information à fournir sur l'ensemble des données à classifier l'approche suivie est appelée classification non supervisée connue aussi sous le terme de regroupement (Clustering). Une panoplie de méthodes des deux approches ont été développées et testées dans ce contexte dont les algorithmes Bio-inspirés (ABIs).

Les algorithmes Bio-inspirés (ABIs) sont issus des comportements observés dans les systèmes biologiques, et sont de plus en plus utilisés dans la résolution des problèmes de regroupement [1]. Il existe plusieurs approches dans les ABIs, les exemples les plus connus sont les algorithmes évolutionnaires (AEs) et les réseaux de neurones artificiels (RNAs) [2]. En général, les méthodes utilisées dans les ABIs exploitent l'intelligence collective considérée comme un phénomène qui émerge naturellement de l'interaction, de la collaboration et de la concurrence de nombreux individus d'une population. L'intelligence collective est généralement définie comme la capacité d'un groupe à résoudre les problèmes que chacun de ses membres ne peut pas résoudre individuellement [3]. Les ABIs ont attiré un intérêt de recherche important dans le domaine de traitement d'images

[1] et en particulier classification d'images satellitaire [4] [5]. Les Réseaux de Neurones sont parmi les premiers outils appliqués à la classification d'images satellitaire [6] [7] [8]. Les algorithmes évolutionnaires sont aussi très utilisés, on trouve un nombre considérable de papiers et en particulier les algorithmes génétiques qui sont utilisés en mono-objective [9] [10] [11] et multi-objective [12] [13]. Différents articles ont employés les algorithmes à essaims de particules à titre d'exemples [14] pour les colonies de fourmis, [15] pour les systèmes immunitaires artificiels. Les auteurs dans [4] [5] présentent un état de l'art important concernant les ABI en télédétection. L'algorithme de recherche coucou (ARC) (Cuckoo Search Algorithm) est un ABI développé par Xin en 2010 pour la résolution des problèmes d'optimisation numériques [16]. Depuis, plusieurs papiers ont apparu utilisant l'ARC dans différents domaines, dans [17] et [18] les auteurs ont exposé les différents travaux réalisés avec l'ARC. Dans [19] et [20] l'ARC est utilisé pour le regroupement des données numériques et satellitaires respectivement.

Malgré la multitude des méthodes de regroupement, il n'existe pas une méthode universelle et optimale pour tous les types de données [21]. Des mesures doivent être effectuées afin de vérifier la véracité des résultats et de les comparer. Dans ce contexte différentes mesures ont été développées, elles sont connues sous le terme : Indices de validation de regroupement (Cluster validity indices). Aussi, ces mesures peuvent être utilisées pour trouver le nombre optimal de centres dans un nuage de données [22] [23]. On distingue trois groupes d'indices de validation : Indices internes, indices externes et indices relatives [24]. D'autres auteurs classifient les indices en deux groupes seulement interne et externe [25]. Aussi les indices de validation peuvent être flous ou durs. Les mesures XB [26], PE et PC de Bezdek [27] [28] sont des indices flous tandis que les mesures DB [29], Dunn [30] et Calinski-Harabasz [31] sont des indices durs. Dans [21] [23] [24] on retrouve une étude bibliographique importante des différents indices de validation utilisés dans le regroupement. Aussi, les ABIs utilisent les indices de validation autant que fonction objective afin de résoudre les problèmes de regroupement. Le nombre important des indices de validation dans la littérature rend le choix de la fonction objective fastidieux.

Dans ce papier l'ARC sera appliqué au regroupement des données synthétiques et satellitaires. La fonction objective sera un indice de validation de regroupement à savoir : DB, XB et PBMF. Ce travail constitue la suite de [32] où on a effectué une étude comparative entre les trois indices de validation cités plus haut. Ces indices ont été testés avec trois méthodes de regroupement (KMeans, FCM et K-Harmonic Means) afin de choisir l'indice le plus adéquat pour une utilisation conjointe avec un ABI. Selon [32] DB et PBMF constituent les indices le plus intéressant, cela sera vérifié dans ce travail. Ce papier est organisé comme suite, une étude détaillée de l'ARC est donnée dans la section suivante. L'approche et les paramètres utilisés sont expliqués dans la section 3. Suivie par la section 4 qui survolera les résultats expérimentaux. La section 5 conclura le papier.

## II. METHODOLOGIE

Cette section présente les concepts de base de l'ARC.

### A. Comportement des oiseaux coucous

ARC est une technique récente inspirée de la nature pour résoudre les problèmes d'optimisation non linéaire [33]. Il a été développé par Yang et Deb en 2009 [33] et les études préliminaires montrent qu'il est très prometteur et pourrait dépasser les algorithmes existants tels que les algorithmes génétiques et les réseaux de neurones [33].

Cet algorithme est basé sur la stratégie agressive de reproduction de certaines espèces des oiseaux coucous en combinaison avec le processus aléatoire *Lévy flight* qui s'inspire du système de vol des mouches des fruits [34].

Il existe plusieurs espèces de coucous, certains d'entre eux déposent leurs œufs dans des nids communs, ils peuvent aussi enlever les œufs d'autres oiseaux pour augmenter la probabilité d'éclosion de leurs propres œufs [35]. Autres espèces coucou pondent leurs œufs dans les nids d'autres oiseaux appelés oiseaux hôtes (souvent d'autres espèces). Si un oiseau hôte découvre que les œufs dans son nid ne sont pas les siens, il peut soit les jeter, soit abandonner son nid et en construire un autre ailleurs. En outre, il existe d'autres espèces de coucous qui ont de femelles parasites très spécialisées dans l'imitation des sons de quelques espèces hôtes choisies [35] en couleur et le motif de leurs œufs, cela permis de réduire aux coucou la probabilité de découverte ainsi qu'abandon de leurs œufs, cela augmentera aussi leur capacité de reproduction.

### B. L'algorithme de recherche coucou

Yang et Deb [33] ont utilisé les trois règles suivantes afin d'appliquer et simplifier l'ARC :

- Chaque coucou pond un œuf à la fois, et le dépose dans un nid choisis aléatoirement.
- Les meilleurs nids ayant la qualité des œufs la plus élevées seront gardés pour les générations suivantes.
- Le nombre de nids d'hôte disponibles est fixé, l'oiseau accueillant pourra découvrir l'œuf étranger avec une probabilité  $P_a$ .

Par souci de simplicité, cette dernière hypothèse peut être approchée par une probabilité  $P_a$  de  $n$  nids qui vont être remplacés par des nouveaux nids (Des solutions aléatoires dans des nouveaux emplacements aléatoires).

Yang et Deb [33] supposent que chacun des œufs dans les nids représente une solution, les nouvelles solutions sont les œufs des coucous. Les meilleures de ces dernières seront utilisées pour remplacer celles qui sont mauvaises dans les nids [33].

Sur la base des trois règles, citées précédemment, les étapes basiques de l'ARC peuvent être résumées dans le pseudo-code suivant (figure 1).

```

Début
Fonction objective  $f(x)$ ,  $x=(x_1, \dots, x_d)^T$ ;
Générer une population initiale de  $n$  nids hôtes
 $x_i (i=1, 2, \dots, n)$ ;
Tant que ( $t < \text{Génération Max}$ ) ou (critère d'arrêt);
  Choisir un coucou ( $i$ ) aléatoirement;
  Générer une nouvelle solution;
  Evaluer la qualité de la solution fitness;  $F_i$ 
  Choisir un nid  $j$  parmi les  $n$  aléatoirement;
  Si ( $F_i > F_j$ ),
  Remplacer  $j$  par la nouvelle solution;
Fin
Abandonner une fraction ( $P_a$ ) des pires nids;
Générer de nouveaux nids dans de nouveaux emplacements;
Garder les meilleures solutions;
Classer les solutions et choisir la meilleure;
Fin Tant que
Fin
  
```

Fig. 1. Le pseudo code de l'ACS

Lors de la génération de nouvelles solutions  $x(t+1)$  pour un coucou (solution)  $i$ , un vol de Lévy (*Lévy flight*) (1) ou marche aléatoire (*random walk*) (2) est effectuée :

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \otimes \text{Lévy}(\lambda) \quad (1)$$

$$x_i^{(t+1)} = x_i^{(t)} + s\varepsilon \quad (2)$$

## III. APPROCHE UTILISEE

L'approche proposée est décrite en détail dans les sections suivantes.

### A. Le codage

Une des parties les plus importantes d'un algorithme d'optimisation est la représentation des solutions, mais il n'y a pas de règles ou de méthodes générales qui expliquent la façon de codage d'une solution [36]. Dans notre approche, les nids (solutions) représentent les valeurs RGB des centres des régions.

Un processus d'initialisation aléatoire est adopté pour la génération des nids.

### B. Evaluation des solutions

Il existe plusieurs choix de fonctions objectives dans la littérature, dans ce papier trois indices sont utilisés comme fonction à optimiser à savoir, DB, XB et PBMF. Dans les sections suivantes, les notations ci dessous seront adoptées:



$N$ : nombre d'objets à classifier dans l'ensemble des données.

$K$ : nombre de classes.

$c_j$ : le centre de la classe  $j$ .

$x_i$ : l'élément  $i$  de l'ensemble des données.

### C. Indice Davies-Bouldin (DB)

Il a été défini par Davies et Bouldin (DB) [29] et représente la proportion de la dispersion intra-cluster sur la dispersion inter-cluster, il nécessite seulement l'utilisation de deux paramètres donné par l'utilisateur à savoir-  $p$  (la mesure de distance) et  $q$  (la mesure de dispersion). L'indice DB est calculé en suivant la formule ci-dessous :

$$DB = \frac{1}{K} \sum_{i=1}^K R_i \quad (3)$$

Avec

$$R_i = \max_{i,i \neq j} \left\{ \frac{S_i + S_j}{M_{ij}} \right\} \quad (4)$$

Où

$$S_i = \left\{ \frac{1}{T_i} \sum_{j=1}^{T_i} \|x_j - c_i\|^q \right\}^{\frac{1}{q}} \quad (5)$$

Et

$$M_{ij} = \left\{ \sum_{k=1}^K \|c_{ki} - c_{kj}\|^p \right\}^{\frac{1}{p}} \quad (6)$$

Avec

$c_{ki}$  :  $k^{\text{th}}$  composant du  $n$ -dimensionnel vecteur  $c_i$ ,

$c_i$  : le centre du cluster  $i$ ,

$M_{ij}$  : la distance de Minkowski,

$T_i$  : le nombre de vecteurs (pixels) dans un cluster  $i$ .

### D. Indice Xie-Beni (XB)

L'indice XB (nommé aussi fonction S) est défini comme une fonction du rapport de la compacité globale et de la séparation des clusters [26] et utilisé généralement dans les algorithmes de classification floue:

$$XB = \frac{\sum_{i=1}^K \sum_{j=1}^N (u_{ij})^m \|x_j - c_i\|^2}{N \min \|c_l - c_i\|^2} \quad (7)$$

Avec :

$m$ : nombre réel supérieur ou égal à 1 ( $1 \leq m < \infty$ )

$u_{ij}$ : degré d'appartenance du pixel  $x_i$  au cluster  $j$ .

Notons que l'indice XB souffre d'une décroissance monotone lorsque le nombre de classes (clusters) s'approche du nombre d'éléments dans l'ensemble des données [37]. Une des solutions adoptées pour résoudre ce problème est de fixer un point  $c_{max}$ , ou le comportement décroissant de l'indice

commence à être observé. Le nombre de centres appartiendra à l'intervalle  $0, c_{max}$ .

### E. Indice Pakhira- Bandyopadhyay-Maulik Fuzzy (PBMF)

Cet indice doit être maximisé. Il est employé avec le regroupement flou et a été développé par [23] et formulé comme suit:

$$PBMF = \frac{1}{K} \times \frac{E_1}{\sum_{i=1}^K \sum_{j=1}^N (u_{ij})^m \|x_j - c_i\|^2} \times D_K \quad (8)$$

Où

$$D_K = \max \|c_i - c_j\| \quad (9)$$

Avec  $E_1$  est une constante pour un ensemble de données.

## IV. RESULTATS EXPERIMENTAUX

Dans cette section, nous décrivons les expérimentations menées sur des données synthétiques et des données réelles. L'algorithme de recherche coucou est utilisé sous sa forme mono-objective conjointement avec les indices de validation DB, XB et PBMF. Ces indices, largement utilisés dans le regroupement de données, représentent dans notre étude des fonctions objectives.

Le premier jeu de données synthétiques de deux dimensions est généré en utilisant une distribution gaussienne et est formé de 600 points répartis en six groupes distincts (fig.2). Le second jeu de données synthétiques contient 5000 points répartis en 15 groupes présentant un chevauchement partiel. Ces jeux de données ont été téléchargés à partir de la page web <http://cs.uef.fi/sipu/datasets> de l'unité SIPU (Speech & image processing unit) de l'université de Eastern Finland (Joensuu) (fig.3).

Les ensembles de données réelles se composent de quatre multi-spectrale sous-scènes d'images satellitaires.

### A. Données synthétiques

Pour le premier jeu de données, les trois fonctions objectives ont été testées en variant le nombre de centres  $K$  de 2 à 9 par pas de 1. Les paramètres de l'algorithme coucou ont été fixés, après plusieurs essais, comme suit :  $\alpha = 10$ ,  $P_a = 0.25$  et  $Iter = 500$ . Les résultats sont reportés dans la Table I. Notons que les tests pour  $\alpha = 0.01$  sont concluants pour DB et XB, cependant dans le cas de la fonction objective PBMF, le résultat se rapproche de la solution (Table II).

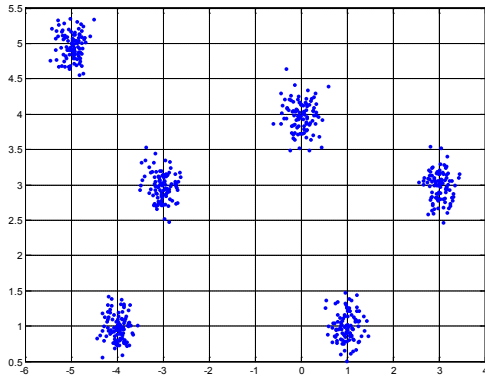


Fig. 2. Premier jeu de données synthétiques

TABLE I. VALEURS DE DB, XB ET PBMF EN FONCTION DE K POUR ALPHA=10

<i>K</i>	$\alpha=10$		
	<i>DB</i>	<i>XB</i>	<i>PBMF</i>
2	0.4496	0.0660	23.8304
3	0.3921	0.0730	41.5129
4	0.3318	0.0604	65.5808
5	0.2640	0.0346	79.8415
6	<b>0.1752</b>	<b>0.0168</b>	<b>86.4457</b>
7	0.2948	0.0180	73.0564
8	0.4092	0.0236	66.8955
9	0.4557	0.0295	62.4733

TABLE II. COMPARAISON ENTRE DB, XB ET PBMF

<i>K</i>	$\alpha=0.01$		
	<i>DB</i>	<i>XB</i>	<i>PBMF</i>
2	0.4501	0.0660	20.8196
3	0.3920	0.0730	36.3472
4	0.3323	0.0604	58.3984
5	0.2613	0.0346	<b>86.8071</b>
6	<b>0.1538</b>	<b>0.0168</b>	79.5622
7	0.2540	0.0180	77.6447
8	0.3181	0.0236	72.8335
9	0.4039	0.0295	58.7983

Pour Alpha=10, nous constatons que l'algorithme coucou délivre le nombre adéquat de centres, et ce avec les trois fonctions objectives. Cependant, l'algorithme est stable et converge rapidement (Nombre d'itération=100) dans le cas de la fonction objective DB.

Pour alpha=0.01, l'algorithme coucou donne les mêmes résultats pour les deux fonctions objectives DB et XB que dans le cas pour alpha=10. Néanmoins, les résultats ne sont pas stables lors de l'utilisation de la fonction objective PBMF malgré accroissement du nombre d'itérations. En effet, la

valeur maximale du PBMF est atteinte avec un nombre de centres =5.

Concernant le second jeu de données, le nombre de centres *k* a été varié de 10 à 20 par pas de 1. Les paramètres de l'algorithme coucou ont été fixés, après plusieurs essais, comme suit :  $\alpha = 10^4$  et  $\alpha = 9 \times 10^4$ , Pa = 0.25 et Iter = 500. Les tables III et IV présentent les résultats obtenus.

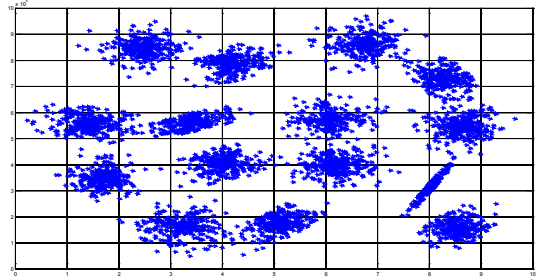


Fig. 3. Deuxième jeu de données synthétiques

TABLE III. COMPARAISON ENTRE DB, XB ET PBMF

<i>K</i>	$\alpha=100000$		
	<i>DB</i>	<i>XB</i>	<i>PBMF</i>
10	0.54364	0.0971	5.3224e+009
11	<b>0.527183</b>	0.0938	<b>5.8880e+009</b>
12	0.613775	<b>0.0904</b>	5.6509e+009
13	0.611052	0.1016	5.5154e+009
14	0.629659	0.1140	5.3791e+009
15	0.604309	0.1136	5.4031e+009
16	0.679735	0.1224	5.0501e+009
17	0.648885	0.1207	5.3373e+009
18	0.617109	0.1395	5.1324e+009
19	0.730713	0.1615	5.1250e+009
20	0.678703	0.1392	4.7958e+009

La lecture du tableau III montre que l'algorithme coucou n'a pas donné le nombre escompté de centres, pour les trois fonctions objectives utilisées avec  $\alpha = 10^4$ . Cela est dû probablement au chevauchement des différents groupes. Afin de pallier la complexité des données, le paramètre Alpha a été fixé à  $9 \times 10^4$  en vue d'améliorer les résultats. Le tableau IV résume les tests effectués.

Suite à la modification de la valeur de  $\alpha$ , nous avons constaté que l'algorithme coucou délivre le bon nombre de centres, à savoir 15 pour les deux fonctions objectives DB et PBMF. Par contre, les résultats sont insatisfaisants lors de l'utilisation de la fonction objective XB (*k*=10 pour une valeur minimale de XB).

TABLE IV. COMPARAISON ENTRE DB, XB ET PBMF

K	$\alpha=900000$		
	DB	XB	PBMF
10	0.568299	<b>0.0836</b>	1.0817+010
11	0.56559	0.0908	1.0857+010
12	0.577802	0.0910	1.0618+010
13	0.573899	0.1011	1.0752+010
14	0.624536	0.1163	1.0854+010
15	<b>0.648406</b>	0.1101	<b>1.0973+010</b>
16	0.556099	0.1411	1.0404+010
17	0.614442	0.1440	1.0157+010
18	0.701075	0.1177	0.9578+010
19	0.709853	0.1565	0.9471+010
20	0.74781	0.1505	0.9359+010



Fig. 5. Région d'Arzew.

### B. Données réelles

Les images utilisées pour les différentes expérimentations sont issues du satellite Landsat8, 3 bandes ont été combinées (Vert, Rouge, Proche infrarouge). Des fenêtres de 400 x 400 pixels et de 600 x 800 pixels ont été sélectionnés représentant la région d'Oran (Fig.4) et la région d'Arzew (Fig.5) respectivement (Ouest d'Algérie) et contenant les thèmes suivants : Eau, urbain, sol nu, jachère et végétation.

Les différents résultats obtenus en utilisant les trois fonctions objectives DB, XB et PBMF sont représentés et discutés dans les paragraphes suivants.



Fig. 4. Région d'Oran.

Plusieurs tests ont été établis sur les données réelles, le but étant d'obtenir une bonne classification assurée par des valeurs minimales pour les fonctions objectives DB et XB, et valeurs maximales par la fonction objective PBMF. Le tableau suivant résume les meilleurs et les pires résultats obtenus pour chaque fonction objective. Les différents paramètres de l'algorithme coucou changent selon la solution (Nombre de centres, nombre d'itération et  $\alpha$ ), le paramètre Pa a été fixé à 0.25 pour tous les tests.

TABLE V. RESULTATS POUR L'IMAGE REPRESENTANT LA REGION D'ORAN

DB			
Paramètres	Min	Paramètres	Max
$\alpha=5$ K=5 Iter=400	0.45969	$\alpha=20$ K=8 Iter=100	0.69472
XB			
Paramètres	Min	Paramètres	Max
$\alpha=10$ K=8 Iter=400	0.0466	$\alpha=5$ K=8 Iter=100	0.0777
PBMF			
Paramètres	Min	Paramètres	Max
$\alpha=10$ K=8 Iter=200	3569.54	$\alpha=10$ K=5 Iter=400	15856.9

TABLE VI. RESULTATS POUR L'IMAGE REPRESENTANT LA REGION D'ARZEW

DB			
Paramètres	Min	Paramètres	Max
$\alpha=5$ K=5 Iter=500	0.4142	$\alpha=20$ K=8 Iter=200	0.6174
XB			
Paramètres	Min	Paramètres	Max
$\alpha=5$ K=6 Iter=500	0.0362	$\alpha=10$ K=8 Iter=100	0.0691
PBMF			
Paramètres	Min	Paramètres	Max
$\alpha=5$ K=8 Iter=300	3320.51	$\alpha=10$ K=5 Iter=300	32355.2

Un ensemble de tests a été effectué sur les deux images. Une population initiale de nids aléatoire est générée pour chaque test, et cinq valeurs de alpha ont été testé à savoir : 0.01, 5, 10, 15 et 20. La condition d'arrêt est le nombre d'itération qui a été pris à 100, 200, 300, 400 et 500 pour chaque valeur de alpha. Les résultats obtenus variaient pour chaque indice de validité. Les tables (V) et (VI) résument les

meilleures et les pires résultats. On remarque que l'ARC a pu obtenir le nombre optimal de centres (K=5) pour la fonction objective DB pour les deux images. Les deux autres indices ont juste permis d'approcher le nombre de centre optimal (K=6 et K=8) pour XB et PBMF. On remarque aussi qu'en augmentant le nombre des centres pour ces images des petites classes peuvent apparaître (contenant juste quelques pixels et qui ne sont pas visibles à l'œil nu), on prend comme exemple le cas de XB min où le nombre de centres est à 6 contrairement au nombre de centres apparent dans l'image qui est égale à 3 (Fig 6). Les images représentant les résultats des tables (V) et (VI) sont données en fin de l'article.

## V. CONCLUSION

Dans ce papier l'ARC est utilisé pour le regroupement des données. Trois indices de validité ont été utilisés autant que fonction objective pour l'ARC à savoir : DB, XB et PBMF. Différents tests ont été établis sur des données synthétiques ainsi que sur des données satellitaires. Les résultats obtenus diffèrent selon la fonction objective utilisée et semblent plus intéressants avec la fonction objective DB. Néanmoins, des petites classes peuvent apparaître, ceci nuira au bon fonctionnement de l'algorithme et devra être réglé dans les travaux futurs. Aussi, l'utilisation d'autres nouveaux indices plus performants ou la combinaison des indices en utilisant l'optimisation multi-objective semblent intéressantes pour l'ARC.

## REFERENCES

- [1] Abdul Khalid NE, Ariff N Md, Yahya S, and Noor NM: A Review of Bio-inspired Algorithms as Image Processing Techniques, In proceeding of: Software Engineering and Computer Systems - Second International Conference, ICSECS 2011, Kuantan, Pahang, Malaysia, June 27-29. (2011)
- [2] Yang X-S: Nature-Inspired Metaheuristic Algorithms, Second Edition, Luniver Press (2010).
- [3] Dariusz K and Heitor SL: Nature-inspired collective intelligence in theory and practice, Editorial, Information Sciences 182 (2012).
- [4] H. Tong, M. Zhao, X. Li, Applications of computational intelligence in remote sensing image analysis, in: ISICA '09: Proceedings of the Fourth International Symposium on Intelligence Computation and Applications, Springer, pp. 171–179 (2009).
- [5] Goel,L, Gupta,D, Panchal .V.K. and Abraham.A, "Taxonomy of Nature Inspired Computational Intelligence A Remote Sensing Perspective", NaBIC, page 200-206. IEEE, (2012).
- [6] Atkinson, P.M., Tatnall, A.R.L.: Neural networks in remote sensing—Introduction. International, Journal of Remote Sensing 18(4), 699–709 (1997)
- [7] Wei, W., Guanglai, G.: An application of neuro-fuzzy system in remote sensing image classification. In: International Conference on Computer Science and Software Engineering, Wuhan, China, December 2008, pp. 1069–1072 (2008).
- [8] MAHLH, Kaouadji.M: Shape-texture features for the VHSR satellite images classification using the MLP neural net. EARSeL eProceedings, ISSN: 1729-3782, 13(2), pp: 67-76, (2014).
- [9] Awad.M.M, Chehdi.K: Satellite Image Segmentation Using Hybrid Variable Genetic algorithm, International Journal of Imaging Systems and Technology, Vol 2, p 199–207 (2009)
- [10] Bandyopadhyay, S., Pal, S.K.: Pixel classification using variable string genetic algorithms with chromosome differentiation. IEEE Transactions on Geoscience and Remote Sensing 39(2), 303–308 (2001).
- [11] Maulik, U., Bandyopadhyay, S.: Fuzzy partitioning using a real coded variable-length genetic algorithm for pixel classification. IEEE Transactions on Geoscience and Remote Sensing 41(5), 1075–1081 (2003).
- [12] Bandyopadhyay.S, Maulik.U and Mukhopadhyay.A: Multiobjective Genetic Clustering for PixelClassification in Remote Sensing Imagery. IEEE Transactions on Geoscience and Remote Sensing, Vol. 45, No. 5, p.1506-1511. ( 2007)
- [13] Bandyopadhyay, S., Maulik, U., Mukhopadhyay, A.: Multiobjective genetic clustering for pixel classification in remote sensing imagery. IEEE Trans. Geosci. Remote Sensing 45, 1506–1511 (2007).
- [14] Song, Q., Guo, P., Jia, Y.: Ant colony optimization algorithm for remote sensing image classification using combined features. In: International Conference on Machine Learning and Cybernetics, July 2008, pp. 3478–3483 (2008).
- [15] Zhang, L., Zhong, Y., Li, P.: Applications of artificial immune systems in remote sensing image classification. In: Proc. 20th Congr. Int. Soc. Photogrammetry Remote Sens., Istanbul, Turkey, July 2004, vol. 35, pp. 397–401 (2004).
- [16] Yang X-S and Deb S: Cuckoo search via Lévy flights, In: Proceedings of World Congress on Nature & Biologically Inspired Computing. IEEE Publications, USA, pp 210–214 (2009).
- [17] Iztok Fister Jr., Xin-She Yang, Dušan Fister and Iztok Fister, "Cuckoo Search: A Brief Literature Review", in Cuckoo Search and Firefly Algorithm, X.-S. Yang, Ed., vol. 516 of Studies in Computational Intelligence, pp. 49–62 (2014).
- [18] Azizah Binti Mohamad, Azlan Mohd Zain and Nor Erne Nazira Bazin, "Cuckoo Search Algorithm for Optimization Problems— A Literature Review and its Applications", Applied Artificial Intelligence, 28:419–448 (2014).
- [19] Zhao,J, Lei,X, Wu,Z, and Tan,Y, " Clustering Using Improved Cuckoo Search Algorithm", The Fifth International Conference on Swarm Intelligence (ICSI 2014) , Hefei, China, October 17-20. Springer, LNCS 8794, pp. 479-488 (2014).
- [20] Goel, S. Sharma, A. & Bedi, P. (2011). Cuckoo Search Clustering Algorithm: A novel strategy of biomimicry. World Congress on Information and Communication Technologies, 916 – 921.
- [21] O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J. M. Prez, and I. Perona, "An extensive comparative study of cluster validity indices," Pattern Recognition, vol. 46, no. 1, pp. 243–256, January 2013.
- [22] G. Gan, C. Ma, and J. Wu, "Data Clustering: Theory, Algorithms, and Applications," ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA, 2007.
- [23] M. K. Pakhira, S. Bandyopadhyay and U. Maulik, "A Study of Some Fuzzy Cluster Validity Indices, Genetic Clustering and Application to Pixel Classification," Fuzzy Sets and Systems, vol. 155, pp. 191-214, 2005.

- [24] M. Halkidi, Y. Batistakis and M. Vazirgiannis, "Clustering validity checking methods: Part II," SIGMOD Record , vol. 31 no. 3, pp. 19-27, ACM, 2002.
- [25] Q. Zhao, "Cluster validity in clustering methods," Ph.D. dissertation, University of Eastern Finland, 2012.
- [26] X.L. Xie and A. Beni, "Validity measure for fuzzy clustering," IEEE Trans. Pattern Anal. Mach. Intell. Vol.3, pp. 841–846, 1991.
- [27] J.C. Bezdek, "Cluster validity with fuzzy sets," J. Cybernet. Vol. 3, pp.58–73, 1974.
- [28] J.C. Bezdek, "Mathematical models for systematics and taxonomy, " in: Eighth International Conference on Numerical Taxonomy, San Francisco, CA, pp. 143–165, 1975.
- [29] D. Davies and D. Bouldin, "A cluster separation measure," IEEE PAMI, Vol. 1, no. 2, pp. 224–227, 1979.
- [30] J.C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well separated clusters," J Cybernet. Vol 3, pp. 32–57, 1973.
- [31] R.B. Calinski and J. Harabasz, "Adendrite method for cluster analysis," Commun. Statist. pp 1–27, 1974.
- [32] Labeled. K, FIZAZI.H, Mahi. H: Comparative Study between Three Cluster Validity Indices, ICDSO, Oran, Algeria, (2014).
- [33] Yang X-S: Nature-Inspired Metaheuristic Algorithms, Second Edition, Luniver Press (2010).
- [34] Yang X-S and Deb S: Cuckoo search via Lévy flights, In: Proceedings of World Congress on Nature & Biologically
- [35] Inspired Computing. IEEE Publications, USA, pp 210–214 (2009).
- [36] Payne RB, Sorenson MD and Klitz K: The Cuckoos, Oxford University Press, New York, (2005).
- [37] Speed ER : Evolving a Mario Agent Using Cuckoo Search and Softmax Heuristics, 2nd International IEEE Consumer Electronics Society's Games Innovations Conference, (2010).
- [38] O.Z. Maimon and Rokach.L, "Data Mining and Knowledge Discovery Handbook", Springer Science & Business Media, 1383 pages, (2005).

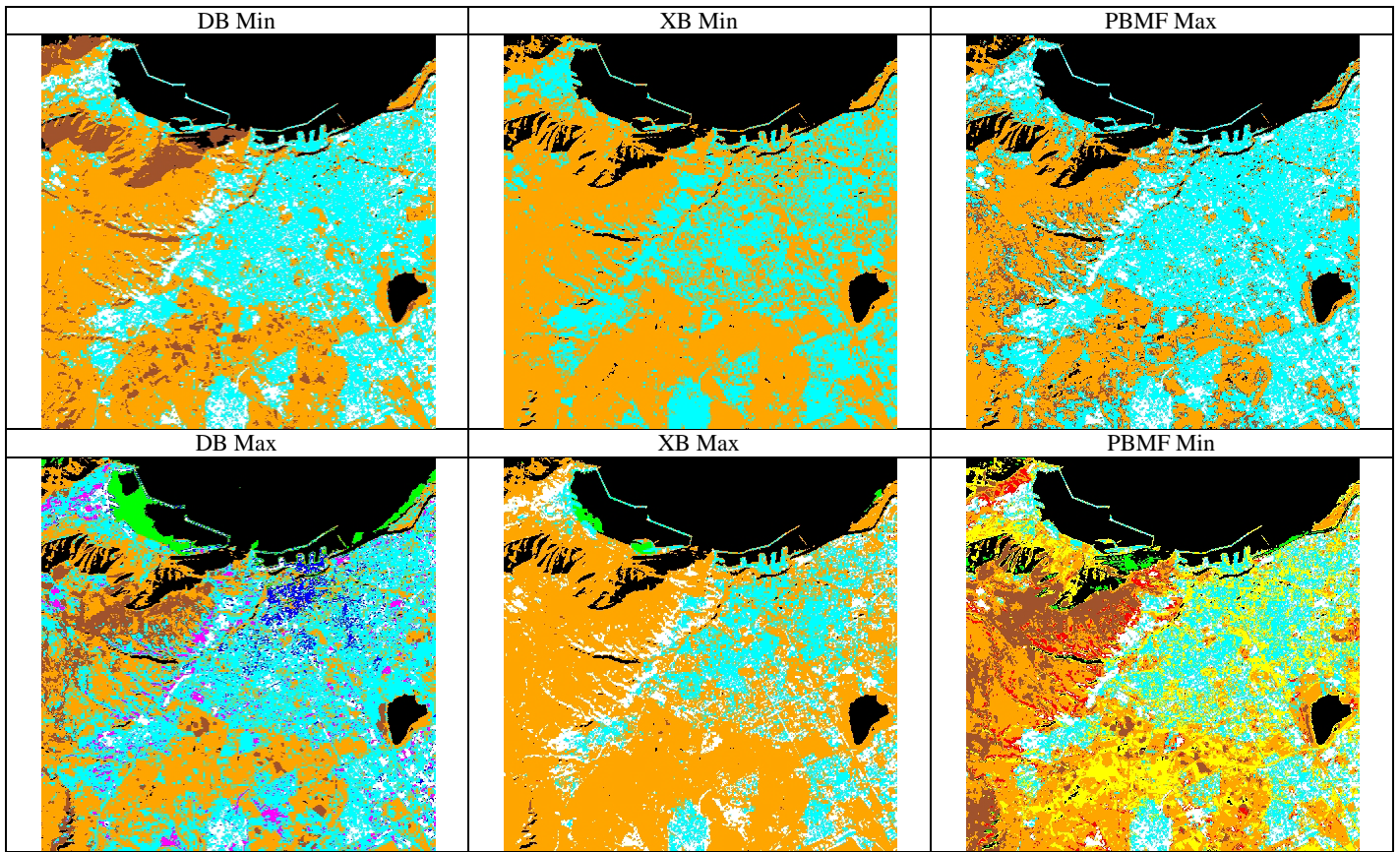


Fig. 6. Images représentantes de la région d'Oran.

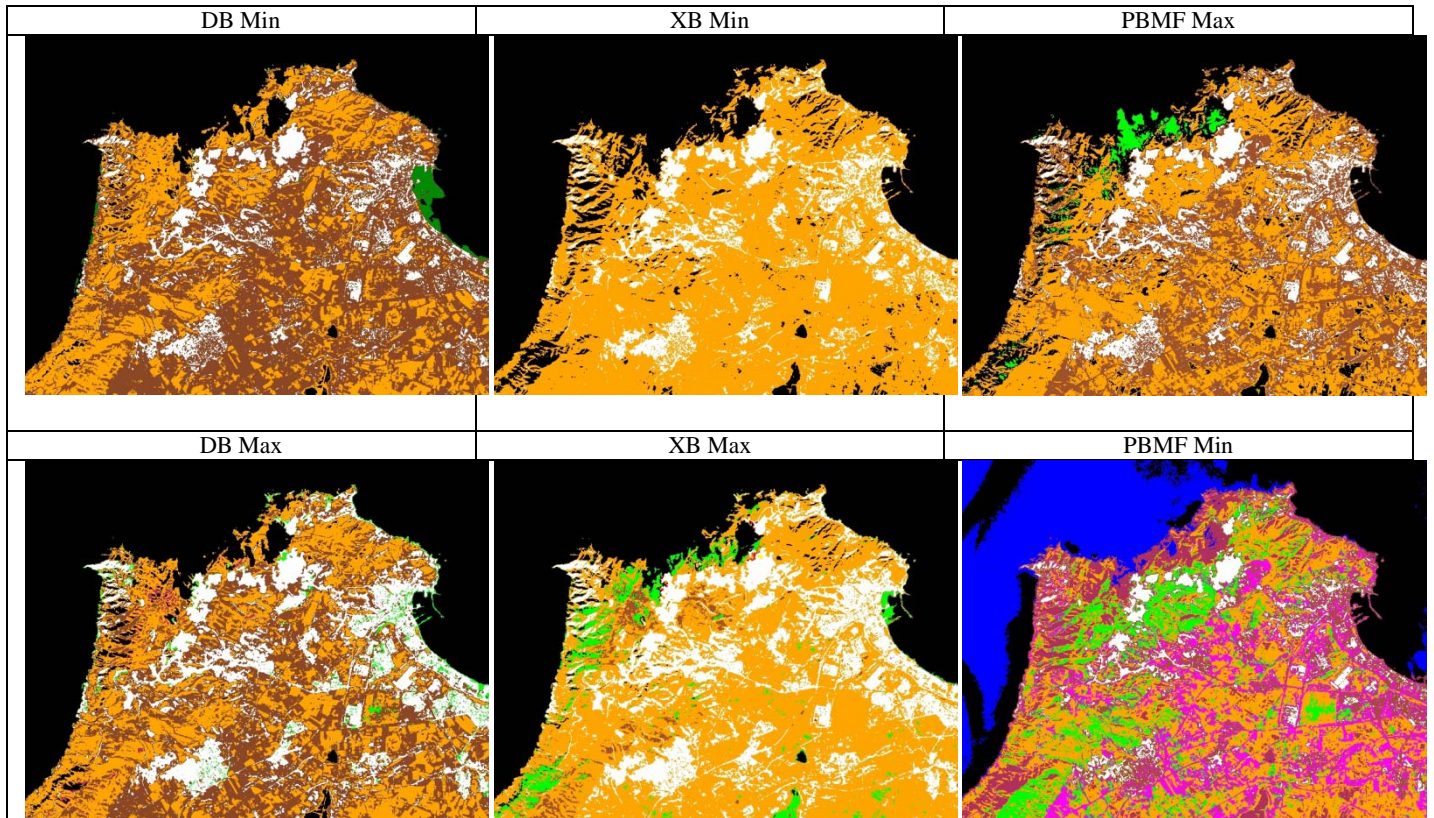


Fig. 7. Images représentantes de la région d'Arzew.

# Etude comparative entre l'algorithme de Firefly et les comportements collectifs des fourmis dans la Classification des Images satellitaires.

Rabhi Seddik<sup>1</sup>, Fizazi Hadria<sup>2</sup>

Université USTO Laboratoire SIMPA, Signal Image Parole Equip Imageries  
BP1505 El M'NAOUAR 31000 Oran – Algérie,

[rabhi\\_seddik@yahoo.fr](mailto:rabhi_seddik@yahoo.fr)

[hadriafizazi@yahoo.fr](mailto:hadriafizazi@yahoo.fr)

## Résumé

Bien que le problème de classification soit la pierre de base de tous système de traitement de données, on n'arrive jamais a trouvé des méthodes de résolution pour une classification idéale. Pour cela, les chercheurs en informatique ont demandé l'aide à la nature et plus particulièrement à la «biologie», pour avoir des solutions optimales ou proche de l'optimale. Parmi les techniques utilisent la biologie comme source d'inspiration l'algorithme à base des fourmis et la méta-heuristique des Firefly. Les algorithmes inspirés par les fourmis qui sont arrivés à maturité, et connu un succès inattendu et porte le nom : « les algorithmes de colonies de fourmis artificielles ». Ces algorithmes peuvent être appliqués à toutes problèmes, ces algorithmes ont été utilisés dans le domaine de la classification automatique et prouvent leur succès. L'algorithme de Firefly est une méta-heuristique récente bio-inspirée conçu pour résoudre les problèmes d'optimisation difficiles, il s'inspire le comportement des lucioles naturelles. Dans notre contexte nous utilisons les deux méta--heuristique à la classification des images satellitaires, Par suite, on présente une étude comparative entre ces deux visions; ces algorithmes découvrent automatiquement les classes dans les données numériques sans connaître le nombre de classes à priori, sans partition initiale et sans paramétrage délicat, ils utilise les principes exploratoires stochastiques d'une colonie de fourmis pour l'une et l'éclate de la lumière chez les lucioles naturelles dans l'autre.

*mots clés*—algorithme de Firefly, fourmis, méta-heuristique bio-inspirée, classification des images.

## I. INTRODUCTION

Des nombreux algorithmes de partitionnement déterministes existent dans la littérature telle que le K-means et le Kppv et leurs variantes. Ces algorithmes sont très simples à implémenter et convergent rapidement avec une solution localement optimale. Cependant leur majeur inconvénient est qu'ils nécessitent de fournir en entrée une partition initiale de bonne qualité ainsi que le nombre possible des classes. Ces contraintes rendent l'utilisation de ces algorithmes peu

intéressante quand on veut classifier automatiquement une image [2].

Dans ces dernières années, une évolution croissante du nombre d'études menées sur les animaux vivants en groupe ou en société et plus particulièrement les insectes sociaux. Ces études dans la théorie de l'auto-organisation ont inspiré un grand nombre des chercheurs pour développer une nouvelle approche appelé les méta-heuristiques.

Les méta-heuristiques sont des algorithmes généraux d'optimisation applicables à une grande variété des problèmes. Elles sont apparues dans le but de résoudre au mieux des problèmes d'optimisation, elles sont généralement inspirées de la nature : de la biologie (Algorithme génétiques, système immunitaire, etc.), l'éthologie (colonie de fourmis, essaim particulaire etc.). L'application des méta-heuristiques en classification d'image a connu une évolution particulière ces dernières années grâce aux avancées technologiques en matière de calcul de machine. [1,3]

Dans ce document, nous avons donc cherché d'appliquer deux méta-heuristiques à la classification de l'image satellitaire d'Oran: La première technique appartient aux algorithmes de colonies de fourmi appelée Antclust proposée pour les problèmes de classification, et la deuxième c'est une méta-heuristique récemment conçu pour résoudre les problèmes d'optimisation difficiles, il s'inspire le comportement des lucioles naturelles nommée l'algorithme de Firefly. Afin de voir leurs résultants et de sortir en dernier lieu d'une étude comparative qui nous permet de voir leurs points forts et les points faibles. Pour cela, nous décrivons dans la deuxième section les deux méta-heuristiques, son origine et ses paramètres, Les algorithmes proposés sont décrits dans la section 3. La section 4 présente nos résultats expérimentaux préliminaires. Le travail conclusion et l'avenir s'ensuit.

## II. ALGORITHME DE FIREFLY

### II.1 LUCIOLES NATURELLE

Les lucioles(en anglais firefly) sont des petits coléoptères ailés capables de produire une lumière clignotante froide pour une attraction mutuelle. Les femelles peuvent imiter les signaux lumineux d'autres espèces afin d'attirer des mâles qu'elles capturent et dévorent. Les lucioles ont un mécanisme de type condensateur, qui se décharge lentement jusqu'à ce que certain seuil est atteint, ils libèrent l'énergie sous forme de lumière. Le phénomène se répète de façon cyclique.

L'Algorithme de lucioles développé par [7] est inspiré par l'atténuation de la lumière sur la distance et l'attraction mutuelle mais il considère toutes les lucioles comme unisexes.

## II.2 L'ALGORITHME FIREFLY

L'algorithme Firefly est une méta-heuristique, bio-inspirée, introduite par Dr Xin-She Yang à l'université Cambridge en 2007. L'algorithme est basé sur le principe d'attraction entre les lucioles et simule le comportement d'un essaim de lucioles dans la nature, ce qui lui donne beaucoup de similarités avec d'autres méta-heuristiques basées sur l'intelligence collective du groupe, tel que l'algorithme PSO (Particle Swarm Optimisation), l'algorithme d'optimisation par colonies d'abeilles (ABC), et l'algorithme des bactéries de fourrages (BFA) [13, 15]. Selon des bibliographies récentes, les performances de l'algorithme Firefly dans la résolution des problèmes d'optimisation dépassent celles des autres algorithmes, tel que les algorithmes génétiques. Ceci a été justifié par des recherches récentes, où les performances de cet algorithme ont été comparées avec celles de quelques algorithmes connus [15].

Inspiration biologique des comportements artificielle :

L'algorithme prend en considération les trois points suivant [14] [15] :

1. Toutes les lucioles sont unisexes, ce qui fait l'attraction entre celles-ci n'est pas en fonction de leur sexe.
2. L'attraction est proportionnelle à leurs luminosités, donc pour deux lucioles, la moins lumineuse se déplacera vers la plus lumineuse. Si aucune luciole n'est lumineuse qu'une luciole particulière, cette dernière se déplacera aléatoirement.
3. La luminosité des lucioles est déterminée en fonction d'une fonction objective (à optimiser)

Pseudo code de L'algorithme : En se basant sur les trois règles précédentes, l'algorithme Firefly se présente comme suit :

```

Définir une population de lucioles  $x_i$ , ( $i = 1, \dots, n$ )
Définir l'intensité de lumière  $I$  à un point  $x_i$ , par la fonction objective  $f(x_i)$  Déterminer le coefficient d'absorption  $\gamma$ 
Tant que ( $t < \text{Max Génération}$ )
Pour  $i = 1$  jusqu'à  $n$ 
Pour  $j = 1$  jusqu'à  $n$ 
  Si ( $I_i < I_j$ )
    Déplacer la luciole  $j$  vers la luciole  $i$ 
  Fin Si
  Varier l'attraction en fonction de la distance  $r$  via  $\exp[-\gamma r]$ 
  Evaluation des nouvelles solutions et mettre à jour l'intensité de lumière
Fin Pour j
Fin Pour i
  Classer les lucioles et trouver la meilleure solution
Fin Tant que
Visualiser les résultats

```

L'algorithme Firefly est formulé avec deux choses importantes : La variation de l'intensité de la lumière et la formulation de l'attraction. Pour simplifier, l'attraction des

lucioles est déterminée en fonction de la luminosité, où la luminosité est déterminée avec la fonction objective.

Dans le cas d'un problème de minimisation, la luminosité  $I$  d'une luciole à une position  $x$  peut être définie comme  $I(x) \propto f(x)^{-1}$ . Cependant, l'attraction  $\beta$  est relative à la position des autres lucioles. Par conséquent, elle varie en fonction de la distance  $r_{ij}$  entre la luciole  $i$  et la luciole  $j$ . D'un autre côté, l'intensité de la lumière diminue avec la croissance de la distance par rapport à la source. Ce qui fait que l'attraction peut varier selon le degré d'absorption. Pour simplifier, l'intensité de la lumière  $I(r)$  varie en fonction de la loi  $I(r) = I_s / r^2$  où  $I_s$  est l'intensité à la source. Pour une valeur constante de  $Y$ , l'intensité varie en fonction de la distance  $r$ , ce qui donne  $I = I_0 e^{-\gamma r}$ , où  $I_0$  est l'intensité de la lumière de la source. La combine des deux effets de la loi carrée inverse et l'absorption peut être approximer avec la formule Gaussienne suivante [16]:

$$I(r) = I_0 e^{-\gamma r^2} \quad (1)$$

Parfois il est nécessaire d'utiliser une fonction monotone décroissante. Dans ce cas, nous pouvons utiliser l'approximation :

$$I(r) = \frac{I_0}{1 + \gamma r^2} \quad (2)$$

Sachant que l'attraction d'une luciole est proportionnelle à l'intensité des lucioles adjacentes, La formule de cette attractivité  $\beta$  d'une luciole peut être définie comme :

$$\beta(r) = \beta_0 e^{-\gamma r^2} \quad (3)$$

Où  $\beta_0$  est l'attraction à  $r = 0$ . Pour généraliser, le calcul de  $\beta(r)$  est défini par :

$$\beta(r) = \beta_0 e^{-\gamma r^m}, (m \geq 1) \quad (4)$$

D'autre part, la distance entre deux lucioles  $i$  et  $j$  à des positions  $x_i$  et  $x_j$  est définie par la distance Cartésienne suivante :

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (5)$$

Où  $x_{i,k}$  représente le  $k^{\text{ème}}$  composant spatiale de la coordonné  $x_i$  de la luciole  $i$ . Dans un repaire 2-D, la formule  $r$ , devient :

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (6)$$

Le mouvement d'une luciole  $i$  attirée par une autre luciole  $j$  (plus lumineuse que  $i$ ) est déterminé par :

$$x_i = x_i + \beta_0 e^{\gamma r_{ij}^2} (x_j - x_i) + \alpha (\text{rand} - 1/2) \quad (7)$$

Le second terme dans l'équation est dû à l'attraction. Tant dis que le troisième terme rajoute de l'aléatoire à l'équation, où  $a$  est aléatoire,  $rand$  une fonction de génération de nombre aléatoire uniforme dans l'intervalle  $[0, 1]$ . Le paramètre  $\gamma$  caractérise la variation de l'attractivité, sa valeur est cruciale dans la détermination de la vitesse de convergence et le comportement de l'algorithme.

## III LA CLASSIFICATION PAR COLONIES DE FOURMI

Les fourmis réelles ont inspirés les chercheurs en informatique dans des nombreux domaines. Cela se justifie particulièrement quand on connaît la richesse



comportementale de ces animaux [24].

Trois domaines s'occupent de près à la modélisation des fourmis en terme mathématique et informatique et à l'utilisation concrète de ces modèles.

1. La première approche est l'exploitation du modèle d'optimisation pour résoudre le problème de partitionnement.
2. La deuxième approche développée et appliquée à de nombreux problèmes d'optimisation combinatoire et numérique.
3. La troisième approche, qui nous intéresse dans notre travail, s'intéresse à la modélisation des fourmis dans le domaine du tri d'objets artificielles et par la classification.

Pour cette dernière approche, les algorithmes de classification automatique sont inspirés du comportement de tri collectif observé chez les fourmis. En effet, certains travaux ont montré que certaines espèces de fourmis parviennent à organiser divers éléments du couvain tels que les œufs, les larves,...etc. [17][18].

Le principe de base de ce comportement est le suivant :

- Lorsqu'une fourmi rencontre un élément du couvain, plus cet élément est isolé, plus elle a de chance de le ramasser;
- Lorsqu'une fourmi transporte un élément du couvain, la probabilité qu'elle le dépose est d'autant plus grande que la densité d'éléments de même type dans le voisinage est grande.

Deneubourg et son équipe [17] furent les premiers à modéliser ce genre de comportement. Lors des expériences de simulation, les objets à rassembler sont placés aussi aléatoirement sur une grille. Les fourmis sont modélisées par des simples agents qui sont placés eux aussi aléatoirement sur la grille représentant l'environnement dans lequel elles évoluent. Chaque agent fourmi n'a qu'une perception locale de son environnement et a pour tâche de déplacer les objets en fonction de la concentration des objets de même type dans leur environnement proche appelé « voisinage ». Le principe est de regrouper les objets similaires en des groupes sur une grille. Chaque fourmi peut prendre un objet avec une probabilité fonction de sa similarité avec les objets présents dans son voisinage et le déposer selon la même probabilité. Après un certain nombre d'itérations, des groupes d'objets similaires se forment sur la grille. La principale caractéristique de ces algorithmes est leur coté non supervisé qui permet de découvrir automatiquement le nombre de groupe adéquat sans intervention extérieure comme dans les algorithmes classiques de classification. Les opérations de dépôt et de ramassage des objets sont biaisées par les probabilités  $P_p$  et  $P_d$  représentées par :

$$P_p = \left( \frac{k_1}{k_1 + f} \right)^2 \quad P_d = \left( \frac{f}{k_2 + f} \right)^2 \quad (8)$$

f est une estimation du nombre d'objets placés dans le voisinage de la fourmi.  $k_1$  et  $k_2$  sont des constantes positives.

Quand f est plus petit que  $k_1$  cela signifie qu'il y a peu d'objets dans le voisinage de l'objet et donc la probabilité de

le prendre est élevée ( $P_p$  est proche de 1). Inversement quand f est plus grand que  $k_1$  la probabilité de prendre l'objet est faible s'il est entouré de plusieurs objets [26].

L'algorithme proposé par Deneubourg a été repris et étendu par Lumer et Faieta [23] pour la classification des données numériques. Les travaux de Lumer et Faieta ont inspiré d'autres auteurs pour la résolution de problème de classification par les fourmis. Aussi les travaux de Kuntz et al [19] se sont inspirés pour le partitionnement de graphes. Dans [20] un algorithme de classification basé fourmis est proposé pour la minimisation de communication entre les processeurs dans un système de simulation où les traitements sont répartis sur plusieurs processeurs. Dans [22] Monmarché introduit AntClass un algorithme de classification utilisant des populations de fourmis. AntClass se base sur l'algorithme de Lumer et Faieta avec des modifications de base. AntClass utilise une grille toroïdale et chaque fourmi à la possibilité de transporter plusieurs objets à la fois et de déposer un tas d'objets sur une même case de la grille.

AntClust [24] est une autre reprise d'AntClass avec des améliorations concernant le support des objets à classer et le déplacement des fourmis était abordé dans la segmentation des images. Le pseudo code d'AntClust sera donné comme suite :

/\* Initialisation\*/

**Pour** chaque pixel  $p_i$  **faire** Placer  $p_i$  dans une cellule du tableau

**Fin pour**

**Pour** chaque fourmi  $a_k$  **faire** Placer dans une cellule choisie aléatoirement et lui affecter son pixel;

Etat [ $a_k$ ] := porteuse ;

**Fin pour**

Déplacer toutes les fourmis vers le nid

/\* Boucle principale\*/

**Pour**  $t=1$  à  $t_{max}$  **faire**

**Pour** chaque fourmi  $a_k$  **faire**

**Si** état[ $a_k$ ] = porteuse **alors**

Déplacer  $a_k$  vers une cellule  $c_k$

Dépôt = faux ;

Dépôt = déposer le pixel  $p_i$  qu'elle transporte dans  $c_k$  avec une probabilité  $P_{Dépôt}(p_i, c_k)$

**Si** Dépôt = vrai **alors** Etat [ $a_k$ ] := libre ;

**Finsi**

**Sinon**

Choisir aléatoirement un pixel  $p_i$  ;

Déplacer  $a_k$  vers la cellule  $c_k$  contenant  $p_i$  ;

Porter = faux;

Porter = Porter  $p_i$  de sa cellule avec une probabilité

$P_{portere}(p_i, c_k)$  ;

**Si** Porter = vrai **alors** Etat [ $a_i$ ] := porteuse ;

**Finsi**

Déplacer les fourmis vers le nid

**Fin pour Retourner la partition obtenue**

#### IV. LES APPROCHES UTILISEE :

##### IV.1 « FClust » PROCESSUS DE CLASSIFICATION PAR L'ALGORITHME DE FIREFLY :

Puisque la méta-heuristique des Firefly est un algorithme d'optimisation comme on a vu précédemment il faut d'abord reformuler le problème de classification d'image, en un problème d'optimisation et essayer de les résoudre.

Donc les méthodes de classification des images définissent les partitions présents dans l'image d'une façon que chaque classe rassemble des pixels ayant des vecteurs de caractéristiques (souvent réduit au niveau de gris) aussi similaires que possible et les classes obtenues doivent être aussi distincts que possible. Ce problème de classification constitue un problème NP-difficile [1], donc, il ne peut pas être résolu exactement dans un temps de calcul raisonnable. Nous ajoutons à ce problème des fonctions qui indiquent une cohésion des classes que nous optimiserons avec l'algorithme Firefly.

Pour classifier les images, Nous initialisons la fonction objective par

$$f_i = \frac{1}{D_{\text{terain}}} \sum_{j=1}^{D_{\text{terain}}} d(x_j, P_i^{\text{CLknown}}(x_j)) \quad (9)$$

Et nous suivrons la démarche expliquée dans la partie II de chapitre précédent, concernant les algorithmes Firefly pour l'optimisation (minimiser l'inertie intra classe) comme suite:

**POPULATION INITIALE:** Pour mettre en œuvre l'algorithme Firefly, une population initiale de  $N$  individus (lucioles) est générée. Chaque individu est représenté par une cellule constitué de  $K$  case qui correspondent à  $K$  centres de classes de  $D$  dimensions (figure IV.2).

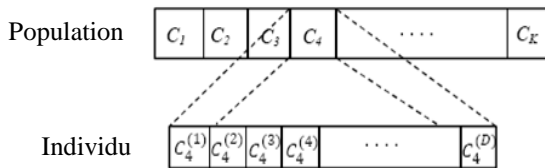


Fig.1: représentation d'une population des individus (Firefly)

La génération des centres des classes se fait aléatoirement. Les centres des classes dans l'espace de recherche, après avoir détecté les bornes inférieures et supérieures de chaque dimension, vu que les attributs ne sont pas de même ordre de grandeur.

**EVALUATION DES LUCIOLE** Après la génération de la population initiale, chaque pixel est affecté à la classe dont le centre est plus proche (La mesure de distance est utilisée pour évaluer les similitudes entre les objets). Tous les Centre de chaque individus sont, alors, évalués selon la fonction illustrée par la formule:

$$J(K) = \sum_{k=1}^k \sum_{i \in c_k} (x_i - c_k) \quad (10)$$

Et des performances proportionnelles a leur intensité lumineuse trouvé par la relation de la fonction objective (équation 9).

**PROCESSUS EFFECTUEE:** Comparez les valeurs d'intensité d'une luciole, si ( $I_2 < I_1$ ) puis déplacez luciole 2 vers 1, et par suite comparer tous les agents et mettre à jour intensité de la lumière en utilisant la fonction objective

(Equation 9). Cette procédure est suivie jusqu'à la convergence de l'algorithme ou un critère d'arrêt est atteint.

**CRITERE D'ARRET:** Les processus d'optimisation par l'algorithme Firefly est itératif: La solution optimale est approchée de génération en génération. Après un nombre de génération fixé au préalable le processus est arrêté. Une seule solution sera retenue: Celle qui a le plus grande valeur de la fonction d'adaptation  $F$  dans la dernière génération.

**RESULTATS:** L'image classifiée est produite par la construction d'une image index de l'appartenance de pixels à chaque classe dont les centres sont représentés par l'individu retenu.

#### IV.2 « ANTClust » PROCESSUS DE CLASSIFICATION PAR FOURMIS:

L'algorithme Antclust utilise le principe exploratoire stochastique d'une colonie de fourmis. Ces dernières se déplacent sur une grille à deux dimensions et peuvent transporter des objets. La saisie ou le dépôt d'un objet sur un tas dépend de la similarité entre cet objet et les objets du tas. Pratiquement dans notre réalisation la grille est représenté par un tableau bidirectionnelle de taille proportionnel au taille de l'image de type cellule, chaque cellule du tableau contient un support dynamique pour placer une ou plusieurs pixel de l'image et un nid pour assurer le déplacement des fourmis. Le ramassage et le dépôt des pixels et le déplacement des fourmis ne sont pas effectués d'une manière physique mais virtuellement suivant des mesures de probabilités calculées dans le voisinage spatial (contexte spatial) du pixel sur lequel la fourmi est déposée.

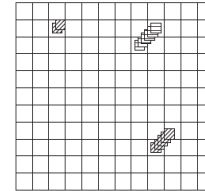


Fig.2 L'environnement de travail.

Le Principe général de l'algorithme Antclust sera donné par l'organigramme suivant:

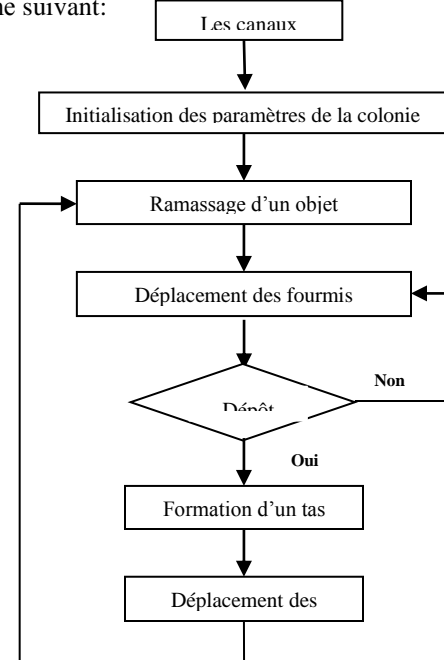


Fig.3: Schéma de Principe de l'algorithme AntClust

## V. RÉSULTATS EXPÉRIMENTAUX

Notre travail expérimental a pour but de comparer les deux algorithmes AntClust et FClust.

### V.1 DONNEE UTILISEE :

Pour tester les approches étudiées, nous avons utilisé une image satellitaire de la région d'Oran (Ouest de l'Algérie) donnée par le satellite LANDSAT TM du 15 mars 1993, acquise du Centre National des Techniques Spatiales CNTS d'Arzew. Cette image a été choisie pour la diversité des thèmes qui la composent. Les canaux utilisés sont : rouge, infrarouge et bleu.



**Fig.4:** La zone d'étude

### V.2 MESEURE D'EVALUATION DES RÉSULTATS

Pour étudier la performance des approches étudiée ici et comparés les résultats de classification des images de test satellitaires, nous avons décidé d'utiliser les mesures d'évaluation suivantes:

- ✓ Nombre des tas obtenus.
- ✓ Calculer l'erreur de classification ( $E_c$ ) par l'expression suivante :

$$E_c = \frac{\text{nombre des pixel mal classer}}{\text{nombre total des pixels}}$$

- ✓ Taux de classification

$$T_c = (1 - E_c) * 100 \% \text{ ou}$$

$$T_c = \left( \frac{\text{nombre des pixel bien classer}}{\text{nombre total des pixels}} * 100 \right) \%$$

### V.3 PARAMÈTRES DES ALGORITHMES:

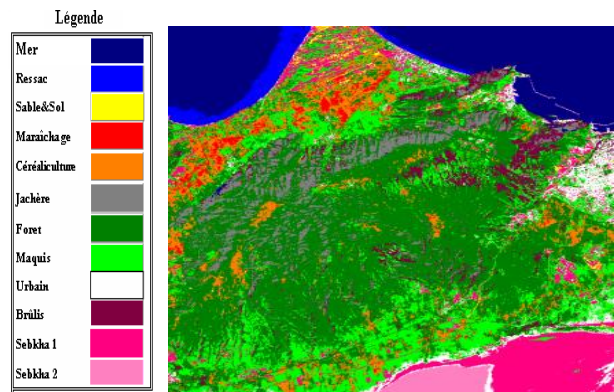
La performance de ces deux méta-heuristiques dépend d'un certain nombre des paramètres dont les valeurs peuvent dépendre ou non de l'image à classifier, Toute ces paramètres ont été choisis d'une manière empirique c'est-à-dire par tâtonnement jusqu'à obtenir les bons résultats.

Le tableau suivant résume les valeurs des paramètres et leur description:

Algo	V. Paramètre	Description
ANTClust	A=20	Nombre des fourmis.
	$K_d=0.3$	Contrôle de la probabilité de déplacer un pixel d'une case.
	$K_p=0.1$	Contrôle la probabilité de déposer un pixel dans une case.
	Q=0.7	Contrôle la probabilité de choisir un pixel dans une case contenant deux pixels.
	B=50	Paramètre qui contrôle la dilatation de la fonction de similarité.
	$T_{max}=3*10^6$	Nombre des itérations de l'algorithme AntClust.
FCLUST	N=20	Nombre de luciole.
	$\beta=1$	Attractivité
	$\gamma=1$	Coefficient d'absorption lumineuse
	$T_{max}=1000$	Nombre des itérations de l'algorithme Firefly.

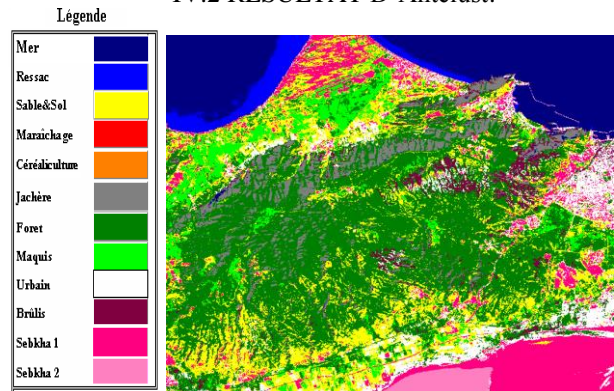
**Tableau 1:** paramètres des algorithmes

### IV.1. RÉSULTAT DE FCLUST:



**Fig.3** image classifiée par l'algorithme de firefly.

### IV.2 RÉSULTAT D'Antclust:



**Fig.4** image classifiée par l'algorithme d'Antclust.

### IV.3 INFLUENCE DE VARIATION DE NOMBRE DES GÉNÉRATIONS

Le tableau suivant illustre le taux et l'erreur de classification par rapport au changement de nombre de génération pour les deux algorithmes:

Algo	Nombre d'itération	Tc	Ec	Nombre des classes obtenus
Antclust	10 <sup>2</sup>	01.02 %	0.989	1
	10 <sup>3</sup>	09.50 %	0.905	2
	10 <sup>5</sup>	35.61 %	0.644	6
	3*10 <sup>6</sup>	96.21 %	0.038	12
Fclust	05	70.86 %	0.291	12
	10	90.31 %	0.194	12
	20	94.83 %	0.052	12
	50	94.33 %	0.057	12
	100	94.00 %	0.060	12

**Tableau 2:** Résultat, changement de nombre de génération

### IV.4 INFLUENCE DE VARIATION DU NOMBRE DE LA POPULATION:

Algo	Nombre des agents	Tc	Ec	Nombre des classes obtenus
Antclust	5	94.00 %	0.060	12
	10	93.06 %	0.069	11
	25	96.21 %	0.038	12
	50	93.50 %	0.065	12
Fclust	5	92.00 %	0.080	12
	10	93.60 %	0.064	12
	20	92.50 %	0.075	12
	25	94.83 %	0.052	12
	50	94.00 %	0.060	12

**Tableau 3:** Résultat, changement de nombre des agents

Pour l'ensemble des tests nous avons fixé les paramètres des algorithmes et nous faisons varier le nombre des agents. Les résultats sont consignés dans le tableau suivant :

### IV.4 DISCUSSIONS ET COMPARAISONS :

A Travers une lecture des tableaux précédents on remarque les points suivants:

- Si le nombre d'itérations augmente le temps d'exécution augmente, et le taux de classification s'améliore et sa valeur de l'erreur de classification sera réduite pour les deux méta-heuristiques.
- Nous remarquons que le nombre des agents utilisées influe un petit peu sur la précision totale de la classification, de plus nous observons que l'approche

des fourmis donne des meilleurs résultats avec un nombre entre 15 et 20 fourmis et nous constatons qu'avec 25 lucioles on obtient le meilleur taux de classification pour l'approche des firefly d'après les différents tests effectués.

On constate que les algorithmes arrive à découvrir le nombre de classes exact avec un taux de reconnaissance élevé sans connaître le nombre de classes à priori mais le fait que le choix des paramètres des méta-heuristiques de façon empirique, c'est-à-dire par tâtonnement. Pour quelqu'une d'utilisateur débutant (ignorants de la philosophie des algorithmes) il nous a fallu de faire des heuristiques pour la détermination des valeurs optimales des paramètres des algorithmes, et aussi le caractère aléatoire n'assure pas leurs convergences dans les images de grandes tailles. Par exemple, le déplacement aléatoire des agents, risque de présence de pixels non classifiés. Comme on a vu les algorithmes bio-inspiré sont connus pour être très gourmands en temps de calcul sur les machines série (PC), car ils sont des algorithmes destinés principalement pour tourner sur des machines parallèles (stations multiprocesseurs). A la fin, on a vécu des difficultés de classification des images satellitaires, on estime un traitement allégé pour d'autres types de données (photo aérien, images médicales, données sonores,...).

En comparant les différents résultats on remarque la supériorité de l'approche des fourmis par rapport au algorithme de Firefly, C'est ce qui justifiera par la diversité des travaux des fourmis existant ce qui facilite leur adaptation et utilisation.

### V CONCLUSION

*AntClust* et *FCLUST* sont des nouvelles approches pour la classification non supervisée de différents types de données. Elles sont inspirées du comportement de tri de couvain observé chez les fourmis réelles et l'attraction de la lumière chez les firefly. L'intérêt principal de ces algorithmes est leur capacités d'extraire automatiquement les classes présent dans l'ensemble de données sans connaître le nombre de classe à priori et sans partition de départ.

Nous avons appliqué ces méthodes sur une image réelle de télédétection, les résultats obtenus montrent que ces deux algorithmes donnent des meilleurs résultats avec une petite supériorité de l'algorithme *AntClust*.

### REFERENCES

- [1] A. Benyamina and H. Fizazi "Efficacité de l'algorithme hybride ACOClust pour le partitionnement des images satellitaires » International Conference on Signal, Image, Vision and their Applications SIVA'11, November 21-24, 2011, Guelma Algeria
- [2] A.Benyamina, H.FIZAZI. «Conception d'une approche hybride pour une classification multisource de donn e de télédétection». M emoire de Magister, U.S.T.Oran Mars 2005.
- [3] A.M.HANNANE, H.FIZAZI « Segmentation des images satellitaires par une approches Neuro-ACO » Les 1 eres Journ ees Scientifiques sur l'Informatique et ses Applications, Guelma, Alg erie, 2009
- [4] J. Kennedy, R. C. Eberhart, "Particle Swarm Optimization". In: IEEE International Conference on Neural network ,pp. 1942—194, 1995.
- [5] Hajo Broersma "Application of the Firefly Algorithm for Solving theEconomic Emissions Load Dispatch Problem,"Hindawi Publishing Corporation, International Journal of Combinatorics, Volume 2011.

- [6] L.DEKHICI, K.BELKADI « Métaheuristique des Lucioles Discrète pour les Flow Shop Hybrides » **C A R I** - O c t o b r e 2 0 1 2
- [7] L.DJEROU, M.BATOUCHE« resolution collective du probleme de segmentation » **Courrier du Savoir** – N°04, Juin 2003, pp. 89-94
- [8] N. Monmarché, F. Guinand et P. Siarry, «Fourmis artificielles : nouvelles directions pour une intelligence collective». Lavoisier 2009
- [10] T.Hassanzadeh, M.R.Meybodi « A New Hybrid Approach for Data Clustering using Firefly Algorithm and K-means
- [11] S.Y. Shin et al. "Improving Image Segmentation Algorithms with Differential Evolution" (eds.), *Reliable and Autonomous Computational Science, Autonomic Systems*, DOI 10.1007/978-3-0348-0031-020, © Springer Basel AG 2011
- [12] Saibal K. Pal, C.S Rai Amrit Pal Singh "Comparative Study of Firefly Algorithm and Particle Swarm Optimization for Noisy Non-Linear Optimization Problems" *I.J. Intelligent Systems and Applications*, 2012, 10, 50-57 Published Online September 2012 in MECS (<http://www.mecs-press.org/>) DOI: 10.5815/ijisa.2012.10.06
- [13] X.S.Yang "Nature-Inspired Metaheuristic Algorithms". Luniver Press, UK.(2008)
- [14] X.S.Yang: "Firefly Algorithms for Multimodal Optimization, Stochastic Algorithms: Foundations and Applications", SAGA 2009, *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 5792:169-178.(2009).
- [15] X.S.Yang: "Firefly algorithm, stochastic test functions and design optimization", *International Journal of Bio-Inspired Computation*, 2(2):78-84.(2010)
- [16] X.S.Yang: "Firefly algorithm, Lévy flights and global optimization", in: *Research and Development in Intelligent Systems XXVI* (Eds M. Bramer, R. Ellis, M. Petridis), Springer
- [17] Deneubourg J.-L., Aron, S., Goss, S., et Pasteels, J.-M. «*The self-organizing exploratory pattern of the argentine ant*». Dans *Journal on insect Behavior*, 3: 159-168, 1990.
- [18] Deneubourg, J. L., Goss, S., Franks, N., Sendova-Franks, A., Detrain, C. and Chretien, L. [1991]. «*The dynamic of collective sorting robot-like ants and ant-like robots*», in J. A. Meyer and S. W. Wilson (eds), *SAB 90 – 1st Conference On Simulation of Adaptive Behavior: From Animals to Animats*, MIT Press.
- [19] P. Kuntz, P. Layzell et Snyers. D. «A colony of ant- like agents for partitioning in vlsi technology». In P. Husbands et I. Harvey, éditeurs, *Proceedings of the Fourth European Conference on Artificial Life*, pages 417-424. MIT Press, 1997.
- [20] A.E. Langham, P.W. Grant, «Using competing ant colonies to solve k-way partitioning problems with foraging and raiding strategies», in: D. Floreano et al. (Eds.), *Advances in Artificial Life, Lecture Notes in Computer Science*, 1674, Springer, 1999, pp. 621–625.
- [21] David Banks, Leanna House, Frederick R. McMorris, Phipps Arabie, Wolfgang Gaul; *Classification, Clustering, and Data Mining Applications* ; Springer-Verlag Berlin Heidelberg ; 2004 (disponible sous <http://scholar.google.com/>)
- [22] N. Monmarché, G. Venturini, M. Slimane. «Classification non supervisee par une population de fourmis artificielles». *Actes Coll. Insectes Sociaux*, 13 : 43-52 (2000).
- [23] Lumer, E. and Faieta, B., 1994, «*Diversity and Adaptation in Populations of Clustering Ants*». *Proceedings of the third International Conference on Simulation of Adaptive Behaviour: From Animals to Animates*. MITPress, 501508.
- [24] P.S. Shelokar, V.K. Jayaraman and B.D. Kulkarni, «An ant colony approach for clustering». *Analytica Chimica Acta*, Vol. 509, Issue 2, 3 May 2004, pp. 187-195.
- [25] A.Benyamina, H.Fizazi. «ANTClust adapté pour le partitionnement des images satellitaires par colonies de fourmis». *COSI'2010*, 18-20 Avril 2010, Ouargla, Algérie
- [26] Ouadfel S., M. Batouche. «*Des fourmis pour la segmentation des images*». *Journées d'étude en Informatique graphique JIG05*, Biskra novembre 2005.

# Comparison of Metaheuristic Algorithms for Unconstrained Global Optimization

Ghania KHENSOUS<sup>1,2</sup>, Belhadri MESSABIH<sup>2</sup>, Abdallah CHOUARFIA<sup>2</sup>, Bernard MAIGRET<sup>3</sup>

**Abstract**— Metaheuristic Algorithms are widely recognized as one of most practical approaches for Global Optimization Problems. This paper presents a comparison between two metaheuristics to solve three standard benchmark functions. Among the most representative single solution metaheuristics was selected Tabu Search Algorithm (TSA), to compare with a novel population-based metaheuristic: Cuckoo Search Algorithm (CSA). Empirical results reveal that the problem solving success of the TSA algorithm is better than the CSA. However, the runtime complexity for acquiring global minimizer by the CSA is generally smaller than the TSA.

**Index Terms**— Metaheuristic Algorithm; CSA; TSA; Global Optimization, Nature inspired algorithms.

## I. INTRODUCTION

Global Optimization has been an active area of research for several decades since optimization problems are inherent in nearly every research area, ranging from engineering to the natural sciences such as biology or chemistry. It is also an active research topic in many other areas such as mathematics, business, and the social sciences [1]. As many real-world optimization problems become more complex, better optimization algorithms were needed.

In all optimization problems the goal is to find the minimum or maximum of the objective function. Therefore, optimization aims to obtain the relevant parameter values allowing an objective function the generation of the minimum or maximum value. Thus, unconstrained optimization problems can be formulated as minimization or maximization of D-dimensional function [2]:

$$\text{Min (or max)} f(x), x=(x_1, x_2, x_3, \dots, x_D)$$

The challenge of developing new methods, baptised Metaheuristics, which are better able to solve difficult problems, still attracts the interests of current researchers. Metaheuristic optimization is therefore a field of growing interest since a single metaheuristic optimization algorithm

which can solve all optimization problems of different types and structures does not exist.

The metaheuristic optimization algorithms use two basic strategies while searching for the global optimum: exploration and exploitation [3]. The exploration process succeeds in enabling the algorithm to achieve the best local solutions within the search space, whereas the exploitation process expresses the ability to reach the global optimum solution around the obtained local solutions.

A metaheuristic algorithm must have some characteristics such as [4]: it must be able to rapidly reach the global optimum solution. The total calculation amount and the runtime required to reach the optimum must be acceptable for practical applications. The algorithmic structure of a metaheuristic has to be simple enough to allow for its easy adaptation to different problems. Besides, it is desired that the metaheuristic has very few algorithmic control parameters excluding the general ones like: size of population, total number of iterations.

There are a wide variety of metaheuristics and a number of properties allowing their classification. One classification dimension is single solution vs population-based [5]: Single solution approaches focus on modifying and improving a single candidate solution such as Simulated Annealing (SA) and Tabu Search Algorithm (TSA). Whereas population-based approaches maintain and improve multiple candidate solutions such as genetic algorithms (GA) and Cuckoo Search Algorithm (CSA).

Several comparisons of the efficiency of metaheuristic algorithms have been published [6-11]. It has been shown that TSA represents one of the most efficient heuristic techniques to find good quality solutions in a short running time compared to population algorithms such as GA or Ant Colony Optimization (ACO) [12, 13]. It has been shown also that CSA gave superior results compared to GA, Particle Swarm Optimization (PSO) and Artificial bee colony (ABC) [4, 14, 15].

In this paper, we use therefore CSA and TSA for solving three standard test functions with diverse properties in terms of modality, separability, and valley landscape to compare their effectiveness in terms of both solution quality and runtime.

The paper is organized as follows: In Section 2, are described the principles of two mentioned metaheuristics: CSA and TSA. In Section 3, Experiments have been

<sup>1</sup>Université Abdelhamid Benbadis de Mostaganem, Faculté des Sciences Exactes et Informatique, P.O. Box 227 & 118 Mostaganem 27000, Algérie.

<sup>2</sup>Université des Sciences et de la Technologie d'Oran USTO-MB, Faculté des Mathématiques et Informatique, B.P 1505 El M'Naouer 31000 Oran Algérie.

<sup>3</sup>CNRS, LORIA, Campus Scientifique - B.P.239 54506- VANDOEUVRE-LES-NANCY Cedex France.

Corresponding author: ghania.khensous@univ-usto.dz

expressed. Finally, conclusions have been given in Section 4.

## II. MATERIAL AND METHODS

### A. Cuckoo Search Algorithm

CSA is a population based stochastic search metaheuristic proposed by Yang and Deb in 2009 [16-18]. It is inspired by a natural mechanism; the parasitic breeding behavior of some cuckoo species that lay their eggs in the nests of host birds. Therefore, a pattern corresponds to a nest and similarly each individual attribute of the pattern corresponds to a Cuckoo-egg and a cuckoo egg represents a new solution. In each computation steps, the new and potentially improved solutions replace worse solutions (eggs in the nests). CSA can be briefly described using the following three rules [17-19]:

- i. Each cuckoo lays one egg at a time and dumps it in a randomly chosen nest.
- ii. Best nests with high quality of eggs will be passed to the next generations.
- iii. The number of available host nests is fixed, and a host bird can discover a foreign egg with a probability  $p_a$ . In this case, the host bird can throw the egg away or abandon the nest and build a new nest in a new location.

In exploring the search space, Yang and Deb discovered that the performance of the CSA can be improved by using Lévy Flights instead of simple random walk [18].

We have chosen this novel population based stochastic global search meta-heuristic algorithm because it has been shown that CSA is superior with respect to GAs and PSO and ABC [12-14]. Besides, several studies indicate that Cuckoo Search is a powerful algorithm and successful results have been achieved in various applications such as manufacturing optimization [20], physically - based runoff - erosion model [21] Query Optimization [22], Training Artificial Neural Networks [23] and PCB (*Printed Circuit Board*) Drill Path Optimization [19], Performing Phase Equilibrium Thermodynamic Calculations [15].

Although this metaheuristics is novel, many improvements are proposed in the literature such as: ICS (*Improved Cuckoo Search*) which is proposed to enhance the accuracy and convergence rate of this algorithm. In this version, a proper strategy for tuning the cuckoo search parameters is used instead of keeping these parameters constant [24]. Another modified cuckoo search algorithm is also presented in [25], the authors implemented a CSA version where the step size is determined from the sorted, rather than only permuted fitness matrix.

In its original version, CSA is proposed for continuous problems; however, it can be extended for combinatorial discrete optimization problems [19, 26]. It can also be combined with others metaheuristics such as: TSA [22], Scatter Search [23] and Greedy Randomized Adaptive Search Procedure (GRASP)[27].

The pseudo-code can be summarized as follows [16-19]:

*begin*

*Objective function*  $f(x)$ ,  $x = (x_1, \dots, x_d)T$

*Generate initial population of n host nests*  $x_i$  ( $i = 1, 2, \dots, n$ )

*while* ( $t < \text{MaxGeneration}$ ) *or* (*stop criterion*)

*Get a cuckoo randomly by Lévy flights*

*evaluate its quality/fitness*  $F_i$

*Choose a nest among n* (say,  $j$ ) *randomly*

*if* ( $F_i > F_j$ ),

*replace j by the new solution;*

*end*

*A fraction* ( $p_a$ ) *of worse nests are substituted by new built ones;*

*Keep the best solutions* (or nests with quality solutions);

*Rank the solutions and find the current best*

*end while*

*Post process results and visualization*

*end*

### B. Tabu Search Algorithm

TSA was first proposed by Fred Glover in 1986 [28]. It is inspired by human memory. It is so-called because it avoids returning to recently visited solutions. At each iteration, the best neighbor is selected as a current solution. To avoid cycles, i.e.; the infinite repetition of a sequence of movements, the  $L$  latest movements are forbidden.  $L$  is the length of the tabu list, which is a short-term memory. It contains the best conformations already visited. Then, the selected movements must be the best ones and not in the tabu list.

Although it might seem simple to reject a solution to a discrete combinatorial problem if it appears in the tabu list, this is not the case for continuous problems. As for other metaheuristics, a random candidate solution within a neighborhood can be defined. If this solution has an objective value higher than the current solution (minimization), the decision whether to accept it or not is based on the content of the tabu list. However, rather than checking if the solution is already tabu it should be checked if the solution is within a certain distance of a solution in the tabu list [29]. A TSA with this property called Enhanced Continuous Tabu Search (ECTS) [30]. It results from an adaptation of combinatorial TSA which aims to follow Glover's basic approach as closely as possible. In order to cover a wide domain of possible solutions, this algorithm first performs the diversification: it locates the most promising areas, by fitting the size of the neighborhood structure to the objective function and its definition domain. For each located promising area, the algorithm continues the search by intensification within one promising area of the solution space.

We have chosen this variant for its advantages [30]: first, its principle is rather basic, directly inspired from combinatorial Tabu Search. Secondly it shows a good performance. Indeed, the authors tested the efficiency of ECTS using a set of benchmark multimodal functions, of which global and local minima are known. ECTS was also compared to other published versions of continuous Tabu Search and to some alternative algorithms like Simulated Annealing.

Among the neighborhood search methods, TSA is considered as one of the most prominent, being widely used

and providing a powerful approach for solving a large range of optimization problems [31]. TSA, which also has the advantage that only function values are used, (differentiability and continuity being not required), is characterized by the use of “memories” during the search [32]. Additionally, TSA needs fewer parameters to be adjusted than the SA algorithm. Unlike other metaheuristics, TSA is not trapped in local minimum [29]. TSA is subject to several developments such as: Directed Tabu Search (DTS) which is a continuous TSA [33], Enhanced Continuous Tabu Search (ECTS) proposed for the global optimization of multi-minima functions [30], Memory Models also introduced in order to improve Tabu Search with Real Continuous Variables [34].

The pseudo-code of the TSA can be summarized as follows:

**Data:** An initial solution "Sol"

**Result:** Best solution

Initialization of memories

repeat

*Sol = best neighbor of Sol;*

*Updating the tabu list, the aspiration criterion and memories;*

*if intensification criterion is checked then intensification*

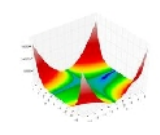
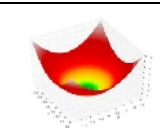
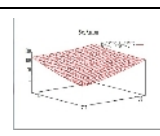
*if diversification criterion is checked then diversification*

until stop criterion

return the best solution found;

### C. Test functions

Test functions are important to validate new optimization algorithms and to compare the performance of various algorithms. There are many test functions in the literature [35-39], but there is no standard list or set of benchmark functions to be followed. In order to make sure whether or not the tested algorithm can solve certain type of optimization efficiently, test functions should have diverse properties. So, we select a list of 3 test problems usually used for checking the properties of optimizers for unconstrained optimization with diverse properties in terms of modality, separability and valley landscape. Tables 1 and 2 below show the continuous test function used in our work [35-39]:

<b>Beale Function</b>	
$f_1(x,y) = (1.5 - x + xy)^2 + (2.25 - x + xy^2)^2 + (2.625 - x + xy^3)^2$	
<b>De Jong Function</b>	
$f_2(x,y) = x^2 + y^2$	
<b>Step Function</b>	
$f_3(x,y) = (x + 0.5)^2 + (y + 0.5)^2$	

Tab. 1. Test Functions.

<b>Test Function</b>	<b>Properties</b>	<b>Search Space</b>	<b>The global Minimum</b>
$f_1$	Differentiable, Non-Separable, Non-Scalable, Unimodal	[-4.5;4.5]	f(3 ;0.5)=0
$f_2$	Differentiable, Separable, Scalable, Multimodal	[-5.12;5.12]	f(0;0)=0
$f_3$	Non-Differentiable, Separable, Scalable, Unimodal	[-100;100]	f(0.5;0.5)=0

Tab. 2. Test Functions properties.

### III. Experiments

To verify the reliability of the CSA and TSA algorithms, several well known test functions as shown in tables above, are considered. In the experiments, the parameters of CSA and TSA are shown in the table 3 below:

<b>Algorithm</b>	<b>Parameters</b>
CSA	Number of Nests =25 Discovery Rate = 0.25
TSA	Tabu List Length = 10 Neighborhood Size = 10

Tab. 3. CSA and TSA parameters.

The table below shows the run times of the algorithms CSA and TSA for 1000, 10000 and 100000 iterations. From this table, we can note that the CSA is better than TSA in terms of running time.

<b>Function</b>	<b>Algorithm</b>	<b>Iteration</b>	<b>Runtime (seconds)</b>
$f_1$	CSA	1000	0.181127
		10000	1.491567
		100000	1.468508
	TSA	1000	999.355000
		10000	10005.400000
		100000	100109.773000
$f_2$	CSA	1000	0.174693
		10000	1.474619
		100000	15.487769
	TSA	1000	999.476000
		10000	10001.728000
		100000	104953.158000
$f_3$	CSA	1000	0.203946
		10000	2.186290
		100000	25.910908
	TSA	1000	999.305000
		10000	10002.178000
		100000	100142.609000

Tab. 4. CSA and TSA Runtimes.

We have then executed each algorithm ten times for each benchmark function. The table below shows the experimental results of the comparative performances in terms of best, worst and average solutions between the algorithms CSA and TSA.



Test Function	CSA		
	Best	Worst	Average
$f_1$	0.00099507	0.1729	0,0420371609
$f_2$	0,0003587200	0,0032968000	0,0012535720
$f_3$	0,12119	2,5204	0,87182667
TSA			
$f_1$	0.003130	9.819423	1,3011135
$f_2$	0,0000010000	0,0001410000	0,0000360909
$f_3$	0,0019300000	0,0195940000	0,0111310909

Tab. 5. CSA and TSA results.

From the table above, we can note that TSA is more efficient since TSA solutions are better than the CSA ones for two functions.

#### IV. CONCLUSION

In this study, we have selected 3 test functions usually used for checking the properties of optimizers for unconstrained optimization with diverse properties in terms of modality, separability and valley landscape. We have also chosen two promising nature-inspired metaheuristic algorithms for the solution of these three test problems. These test problems were solved by the different metaheuristics and the results were tracked and compared. The results clearly show that TSA is more reliable whereas CSA is better in terms of runtime.

Others test functions and/or metaheuristics could be tested in a future study in an attempt to find the most reliable and efficient algorithm.

#### REFERENCES

- [1] C. Grebner, "New Tabu Search Algorithms for the Exploration of Energy Landscapes of Molecular Systems", WÜRZBURG 2012.
- [2] M. TUBA, M. SUBOTIC, N. STANAREVIC, "Modified cuckoo search algorithm for unconstrained optimization problems", Proceedings of the European Computing Conference, 978-960-474-297-4.
- [3] E. Rashedi, H. Nezamabadi-pour, S. Saryzadi, GSA: A Gravitational Search Algorithm, Information Sciences 179 (2009) 2232–2248.
- [4] P. Civicioglu, E. Besdok, A conceptual comparison of the Cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms, Artif Intell Rev DOI 10.1007/s10462-011-9276-0 (2011).
- [5] A. LAYEB, "Introduction aux métaheuristiques", Université Mentouri de Constantine (2009).
- [6] Bajeh, A. O., Abolarinwa, K. O., "Optimization: A Comparative Study of Genetic and Tabu Search Algorithms", International Journal of Computer Applications (0975- 8887) Volume 31– No.5, October 2011.
- [7] S. Kannan, S. M. R. Slochanal and N. P. Padhy, "Application and Comparison of Metaheuristic Techniques to Generation Expansion Planning Problem", IEEE Transactions on Power Systems, Vol. 20, No. 1, February 2005.
- [8] M. Madi, D. Markovi and M. Radovanovi, "Comparison of MetaHeuristic Algorithms for Solving Machining Optimization Problems", Facta Universitatis Series: Mechanical Engineering Vol. 11, No 1, 2013, pp. 29 - 44.
- [9] J. Silberholz and B. Golden, "Comparison of Metaheuristics", University of Maryland.
- [10] Z. N. AZIMI, "Comparison of Metaheuristic Algorithms for Examination Timetabling Problem", J. Appl. Math.&Computing Vol.16 (2004), No 1-2, pp. 337-354.
- [11] T. Arora, Y. Gigras, "A Survey of Comparison between Various Metaheuristic Techniques for Path Planning Problem", International Journal of Computer Engineering & Science, Nov. 2013.
- [12] Pirim H, Bayraktarn E, Eksioğlu B, "Tabu Search: A Comparative Study".
- [13] C. Pappas, R. K. Srinivasan, J. W. Barnes, and S. Sarkar, "A Comparison of Heuristic, Meta-Heuristic and Optimal Approaches to

the Selection of Conservation Area Networks", University of Texas at Austin.

- [14] Sangita Roy, Sheli Sinha Chaudhuri, Cuckoo Search Algorithm using Lévy Flight: A Review, I.J. Modern Education and Computer Science, 2013, 12, 10-15.
- [15] Seif-Eddeen K. Fateen, Adrian Bonilla-Petriciolet, On the Effectiveness of Nature- Inspired Metaheuristic Algorithms for Performing Phase Equilibrium Thermodynamic Calculations, The Scientific World Journal (2014).
- [16] Yang, X.-S., and Deb, S., "Engineering Optimisation by Cuckoo Search", Int. J. Mathematical Modelling and Numerical Optimisation, Vol. 1, No. 4, 330–343 (2010).
- [17] X.-S. Yang, S. Deb, "Cuckoo search via Lévy flights", in: Proc. of World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), December 2009, India. IEEE Publications, USA, pp. 210-214 (2009).
- [18] Xin-She Yang (2010), "Nature-Inspired Metaheuristic Algorithms Second Edition", Luniver Press.
- [19] Wei Chen Esmonde Lim, G. Kanagaraj and S. G. Ponnambalam, PCB Drill Path Optimization by Combinatorial Cuckoo Search Algorithm, Scientific World Journal Volume 2014, Article ID 264518. <http://dx.doi.org/10.1155/2014/264518>
- [20] A Syberfeldt and S. Lidberg, "Real-World Simulation-Based Manufacturing Optimization using Cuckoo Search", Proceedings of the 2012 Winter Simulation Conference C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher, eds.
- [21] Celso A. G. Santos, Paula K. M. M. Freire and Sudhanshu K. Mishra, "Cuckoo Search via Lévy Flights for Optimization of a Physically -Based Runoff- Erosion Model", Journal of Urban and Environmental Engineering (JUEE), v.6, n.2, p.123-131, 2012.
- [22] Mukul Joshi, Praveen Ranjan Srivastava, "Query Optimization: An Intelligent Hybrid Approach using Cuckoo and Tabu Search, International Journal of Intelligent Information Technologies, 9(1), 40-55, January-March 2013.
- [23] Ahmed T. Sadiq Al-Obaidi, "Improved Scatter Search Using Cuckoo Search", (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol.2, No. 2, 2013.
- [24] Ehsan Valian, Shahram Mohanna and Saeed Tavakoli "Improved Cuckoo Search Algorithm for Global Optimization", International Journal of Communications and Information Technology, IJCIT
- [25] M. TUBA, M. SUBOTIC and N. STANAREVIC, "Performance of a Modified Cuckoo Search Algorithm for Unconstrained Optimization Problems", WSEAS TRANSACTIONS on SYSTEMS, E-ISSN: 2224-2678, Issue 2, Volume 11, February 2012.
- [26] Amira Gherboudj, "Méthodes de résolution de problèmes difficiles académiques", Université de Constantine2, 2013.
- [27] Hongqiang Zheng, Yongquan Zhou and Qifang Luo, "A hybrid Cuckoo Search Algorithm-GRASP for Vehicle Routing Problem", Journal of Convergence Information Technology(JCIT) Volume8, Number3, Feb 2013.
- [28] F. GLOVER, "Tabu Search: A Tutorial", University of Colorado (1990).
- [29] P. G. Schulz, "Creative Design in Optimization Metaheuristics Applied to Multi- modal Continuous Functions", Kongens Lyngby 2006.
- [30] R. Chelouah, P. Siarry, "Tabu Search applied to global optimization", European Journal of Operational Research 123 (2000) 256-270.
- [31] Y. PY, "Local Search Techniques: Focus on Tabu Search", W Jaziri Ed., ISBN 978-3-902613-34-9, pp. 278, 2008.
- [32] L. B. Morales, R. Garduño-Juárez, J. M. Aguilar-Alvarado, F. J. Riveros-Castro, "A Parallel Tabu Search for Conformational Energy Optimization of Oligopeptides", J Comput Chem 21: 147–156, 2000.
- [33] A. Hedar, M. Fukushima, "Continuous Optimization: Tabu Search directed by direct search methods for nonlinear global optimization", European Journal of Operational Research 170 (2006) 329–349.
- [34] D. Andrew M. Connor, "Memory Models for Improving Tabu Search with Real Continuous Variables", Proceedings of the Sixth International Conference on Hybrid Intelligent Systems (HIS'06).
- [35] M. Molga, C. Smutnicki, "Test functions for optimization needs", 2005.
- [36] N. Andrei, "An Unconstrained Optimization Test Functions Collection", Advanced Modeling and Optimization, Volume 10, Number 1, 2008.
- [37] X. S. Yang, "Test Problems in Optimization", Department of Engineering, University of Cambridge, UK, 2010.
- [38] M. Jamil, X. S. Yang, "A Literature Survey of Benchmark Functions For Global Optimization Problems", Int. Journal of Mathematical Modelling and Numerical Optimization, Vol. 4, No. 2, pp. 150–194 (2013).

- [39] Ernesto P . Adorio, “MVF - Multivariate Test Functions Library in C for Unconstrained Global Optimization”, Department of Mathematics U.P . Diliman, 2005

# An Hybrid Method to Improve Segmentation of Brain MR Images

Lamiche Chaabane

Department of Computer Science, M'sila University  
Algeria  
Lamiche07@gmail.com

Moussaoui Abdelouahab

Department of Computer Science, Setif 1 University  
Algeria  
Moussaoui.abdel@gmail.com

**Abstract**—In this work, we present a fully unsupervised segmentation process of magnetic resonance image (MRI) of the brain using a data fusion technique and some of ideas of the possibility theory context. The fusion methodology is decomposed into three fundamental phases. We modeling information coming from T2 and PD weighted images in a common framework, in this step an hybridization between FCM and PCM algorithms is retained. In the second phase an operator of fusion is used to combine then these information. Finally, an image of fusion is generated when a decision rule is applied. Some results are presented and discussed using a set of simulated MR image.

**Keywords**—fusion; possibility theory; segmentation; FCM; PCM; MR images

## I. INTRODUCTION

Fully automatic brain tissue classification from magnetic resonance images (MRI) is of great importance for research and clinical study of much neurological pathology. The accurate segmentation of MR images into different tissue classes, especially gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), is an important task.

In medical imaging field, segmenting MR images has been found a quite hard problem due to the existence of image noise, partial volume effects, the presence of smoothly varying intensity inhomogeneity, and large amounts of data to be processed. To handle these difficulties, a large number of approaches have been studied, including fuzzy logic methods [1], neural networks [2], Markov random field methods with the maximum expectation [3], statistical methods [3], and data fusion methods [4], to name a few.

In recent years, the need for data fusion in medical image processing increases in relation to the increase of acquisition techniques such as magnetic resonance imaging (MRI), tomography(CT), the newer positron emission tomography (PET) and a functional modality SPECT. These techniques are more and more jointly used to give access to a better knowledge[5].

As one typical data fusion problem, the segmentation of multispectral brain MR images aims at achieving improved segmentation performance by taking advantage of redundancy and complementariness in information provided by multiple sources. There have existed many data fusion methodologies, which are capable of reasoning under various types of uncertainty. Typical ones include probability theory based

approaches, possibility theory based approaches, and Dempster-Shafer evidence theory based approaches [5].

Traditionally probabilities theory was the primary model used to deal with uncertainty problems, but they suffer from drawbacks. Whereas the Dempster-Shafer theory also allows to representing these two natures of information using functions of mass but the set of operators used by this theory is very restricted.

Alternative to this approach is the possibility theory where uncertainty and imprecision are easily modeled and it allows to combining information coming from various sources by the use a wide range of available combination operators [5].

In this work we aim to improve the segmentation of the human brain tissues using a multispectral fusion approach. This approach consists of the computation of fuzzy tissue maps in each of tow modalities of MR images namely T2 and PD as an information source, the creation of fuzzy maps by a combination operator and a segmented image is computed in decision step. This paper is organized as follows : In section II, we summarize the main ideas of FCM algorithm. In section III, we introduce the principals of possibility theory reasoning. Section IV outlined the fusion process methodology. Steps of our proposed method are described in section V. Section VI presents some experimental results. Finally, conclusion and perspectives of our work are suggested in section VII.

## II. THE FCM ALGORITHM CLUSTERING

Clustering is a process of finding groups in unlabeled dataset based on a similarity measure between the data patterns (elements). A cluster contains similar patterns placed together. The fuzzy clustering technique generates fuzzy partitions of the data instead of hard partitions. Therefore, data patterns may belong to several clusters, having different membership values with different clusters. The membership value of data pattern to a cluster denotes similarity between the given data pattern to the cluster. Given a set of  $N$  data patterns  $X=\{x_1, x_2, x_3, \dots, x_n\}$  the Fuzzy C-Means (FCM) clustering algorithm minimizes the objective function [26][27]:

$$J(B, U, X) = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m d^2(x_j, b_i) \quad (1)$$

Where  $x_j$  is the  $j$ -th P-dimensional data vector,  $b_i$  is the center of cluster  $i$ ,  $u_{ij}$  is the degree of membership of  $x_j$  in the  $j$ -th cluster,  $m$  is the weighting exponent  $d^2(x_j, b_i)$  is the Euclidean distance between data  $x_j$  and cluster center  $b_i$ .

The minimization of objective function  $J(B, U, X)$  can be brought by an iterative process in which updating of membership  $u_{ij}$  and the cluster centers are done for each iteration.

$$u_{ij} = \left[ \sum_{k=1}^c \left( \frac{d^2(x_j, b_i)}{d^2(x_j, b_k)} \right)^{2/(m-1)} \right]^{-1}. \quad (2)$$

$$b_i = \frac{\sum_{k=1}^N u_{ij}^m x_k}{\sum_{k=1}^N u_{ik}^m}. \quad (3)$$

Where :

$$\forall i \in \{1..C\}, \quad \forall j \in \{1..N\} \quad \begin{cases} u_{ij} \in [0,1] \\ 0 < \sum_{i=1}^N u_{ij} < N \end{cases}. \quad (4)$$

$$\forall j \in \{1..N\} \quad \sum_{i=1}^c u_{ij} = 1. \quad (5)$$

The algorithm of the FCM consists then of the reiterated application of (2) and (3) until stability of the solutions.

### III. THE POSSIBILITY THEORY

Possibilistic logic was introduced by Zadeh (1978) following its former works in fuzzy logic (Zadeh, 1965) in order to simultaneously represent imprecise and uncertain knowledge. In fuzzy set theory, a fuzzy measure is a representation of the uncertainty, giving for each subset  $Y$  of the universe of discourse  $X$  a coefficient in  $[0,1]$  assessing the degree of certitude for the realization of the event  $Y$ . In possibilistic logic, this fuzzy measure is modeled as a measure of possibility  $\Pi$  satisfying:

$$\Pi(X) = 1 \quad \text{et} \quad \Pi(\emptyset) = 0$$

$$(\forall (Y_i)) \Pi(\bigcup_i Y_i) = \sup_i \Pi(Y_i)$$

An event  $Y$  is completely possible if  $\Pi(Y) = 1$  and is impossible if  $\Pi(Y) = 0$ . Zadeh showed that  $\Pi$  could completely be defined from the assessment of the certitude on each singleton of  $X$ . Such a definition relies on the definition of a distribution of possibility  $\pi$  satisfying :

$$\pi : X \rightarrow [0,1]$$

$$x \rightarrow \pi(x) / \sup_x \{\pi(x) = 1\}$$

Fuzzy sets  $F$  can then be represented by distributions of possibility, from the definition of their characteristic function  $\mu_F : (\forall x \in X) \mu_F(x) = \pi(x)$

Distributions of possibility can mathematically be related to probabilities, and they moreover offer the capability to declare the ignorance about an event. Considering such an event  $A$  (e.g., voxel  $v$  belongs to tissue  $T$ , (where  $v$  is at the interface between two tissues), the probabilities would assign  $P(A) = P(\bar{A}) = 0.5$ , whereas the possibility theory allows fully possible  $\Pi(A) = \Pi(\bar{A}) = 1$ . We chose to model all the information using distributions of possibility, and equivalently we represented this information using fuzzy sets.

The literature classically distinguishes three modes for combination of uncertainty and imprecise information in a possibility theory framework:

*The conjunction:* gather the operators of t-norms (fuzzy intersection), this mode of combination must be used if measurements are coherent, i.e. without conflict.

*The compromise:* gather the median operator and some average operators, it must be used when measurements are in partial conflict.

*The Disjunction:* gather the operators of t-conorms (fuzzy union), it must be used when measurements are in disaccord, i.e. in severe conflict.

### IV. THE FUSION PROCESS STEPS

A general information fusion problem can be stated in the following terms : given  $l$  sources  $S_1, S_2, \dots, S_l$  representing heterogeneous data on the observed phenomenon, take a decision  $d_i$  on an element  $x$ , where  $x$  is higher level object extracted from information, and  $d_i$  belongs to a decision space  $D = \{d_1, d_2, d_3, \dots, d_n\}$ . In numerical fusion methods, the information relating  $x$  to each possible decision  $d_i$  according to each source  $S_j$  is represented as a number  $M_{ij}$  having different properties and different meanings depending on the mathematical fusion framework. In the centralized scheme, the measures related to each possible decision  $i$  and provided by all sources are combined in a global evaluation of this decision, taking the form, for each  $i$  :  $M_i = F(M_{i1}, M_{i2}, M_{i3}, \dots, M_{in})$ , where  $F$  is a fusion operator. Then a decision is taken from the set of  $M_i$ ,  $1 \leq i \leq n$ . In this scheme, no intermediate decision is taken and the final decision is issued at the end of the processing chain. In decentralized scheme decisions at intermediate steps are taken with partial information only, which usually require a difficult control or arbitration step to diminish contradictions and conflicts [5]. The three-steps fusion can be therefore described as :

Modeling of information in a common theoretical frame to manage vague, ambiguous knowledge and information imperfection. In addition, in this step the  $M_{ij}$  values are estimated according to the chosen mathematical framework.

Combination : the information is then aggregated with a fusion operator F. This operator must affirm redundancy and manage the complementarities and conflicts.

Decision : it is the ultimate step of the fusion, which makes it possible to pass from information provided by the sources to the choice of a decision  $d_i$ .

## V. PROPOSED METHOD

According to the data fusion process, our method consists on three steps below :

### A. Modeling step

In the framework of possibility theory and fuzzy sets [7], the  $M_{ij}$ 's represent membership degrees to a fuzzy set or possibility distribution  $\pi$ , taking the form for each decision  $d_i$  and source  $S_i$ :  $M_{ij} = \pi_j(d_i)$ . Particularly, in our study this step consists in the creation of WM, GM, CSF and background (BG) fuzzy maps for both T2 and PD images using the FPCM algorithm then  $u_{ij} = \pi_j(d_i)$

### B. Fusion step

For the aggregation step in the fusion process, the advantages of possibility theory rely in the variety of combination operators, which must affirm redundancy and manage the complementarities. And may deal with heterogeneous information. It is particular interest to note that, unlike other data fusion theories like Bayesian or Dempster-Shafer combination, possibility theory provides a great flexibility in the choice of the operator, that can be adapted to any situation at hand [4]. If  $\pi_T^{T2}(v)$  and  $\pi_T^{PD}(v)$  are the memberships of a voxel  $v$  to tissue T resulting from step 1 then a fusion operator F combine these values to generate a new membership value and can managing the existing ambiguity and redundancy. The possibility theory propose a wide range of operators for the combination of memberships[8].

For our MR images fusion, we chose a context-based conjunctive operator because in the medical context, both images were supposed to be almost everywhere concordant, except near boundaries between tissues and in pathologic areas. In addition, the context-based behavior allowed to take into account these ambiguous but diagnosis-relevant areas. Then we retained an operator of this class, this one is introduced in [8]:

If  $\pi_T^{T2}(v)$  and  $\pi_T^{PD}(v)$  are the gray-levels possibility distributions of tissue T extracted from T2 and PD fuzzy maps respectively and F design the fusion operator, then the fused possibility distribution is defined for any gray level  $v$  as :

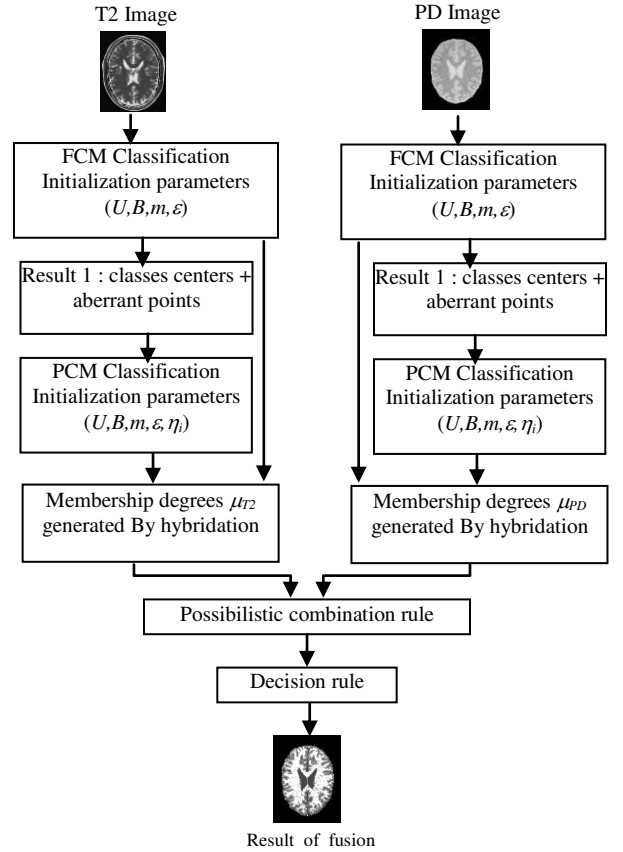
$$\pi_T(v) = \max\left(\frac{\min(\pi_T^{I_b}(v), \pi_T^{I_j}(v))}{h}, \min(\max(\pi_T^{I_b}(v), \pi_T^{I_j}(v)), 1-h)\right)$$

where  $I_b, I_j \in \{T2, PD\}$ , and  $h$  is a measure of agreement

between  $\pi_T^{I_b}$  and  $\pi_T^{I_j}$  :  $h = 1 - \sum_{v \in \text{Image}} |\pi_T^{I_b}(v) - \pi_T^{I_j}(v)| / |\text{Image}|$

### C. Decision step

A segmented image was finally obtained using the four maps computed in step 2 by assigning to the tissue T any voxel for which it had the greatest degree of membership (i.e maximum of possibility rule)[5].



It should be noted that the stability of this approach depend to the stability of the algorithm used in the modeling step[8].

## VI. EXPERIMENTAL RESULTS

Since the ground truth of segmentation for real MR images is not usually available, it is impossible to evaluate the segmentation performance quantitatively, but only visually. However, Brainweb<sup>1</sup> provides a simulated brain database including a set of realistic MRI data volumes produced by an MRI simulator. These data enable us to evaluate the performance of various image analysis methods in a setting where the truth is known .

To have tests under realistic conditions, one volume was generated with a thickness of 1 mm and a level of noise of 3%. We fixed at 20% the parameter of heterogeneity.

<sup>1</sup> [www.bic.mni.mcgill.ca/brainweb](http://www.bic.mni.mcgill.ca/brainweb).

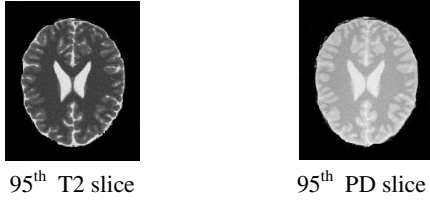


Figure 1. Simulated T2 and PD images illustrate the fusion.

The different segmentations are showing in figure 2 below :

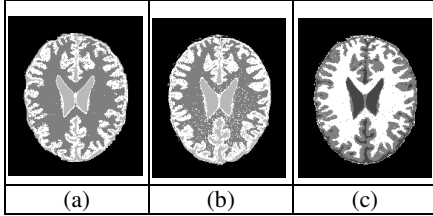


Figure 2. (a) T2 segmented with FCM algorithm. (c) PD segmented with FCM algorithm. (d) Image of fusion.

To compare the performance of these three models of fusion produced by F operator, we compute different coefficients reflecting how well two segmented volumes match. We use a different performance measures :

$$Overlap(Ovrl) = \frac{TP}{TP + FN + FP}.$$

$$Similarity(SI) = \frac{2TP}{2TP + FN + FP}.$$

Where TP and FP stand for true positive and false positive, which were defined as the number of voxels correctly and incorrectly classified as brain tissue by the automated algorithm. TN and FN stand for true negative and false negative, which were defined as the number of voxels correctly and incorrectly classified as non-brain tissue by the automated algorithm. The comparative results are presented in table 1 below :

TABLE I. COMPARATIVE RESULTS.

	T2/PD Fusion			PD alone			T2 alone		
	CSF	WM	GM	CSF	WM	GM	CSF	WM	GM
Ovrl.	<b>0.90</b>	<b>0.93</b>	<b>0.92</b>	0.58	0.76	0.70	0.67	0.90	0.83
Si.	<b>0.94</b>	<b>0.96</b>	<b>0.95</b>	0.78	0.83	0.80	0.83	0.92	0.86

The results in Table 1 show a considerable improvement for all tissues using T2/PD fusion than T2 only and PD only.

Finally, we have also compared the performance of our proposed algorithm to that of well-known methods and other published reports that have recently been applied on brain tissue segmentation on Brainweb datasets for the segmentation of MR images in CSF, WM and GM tissues. They are summarized in table 2, these include the proposed work in [9] and the published approach of fusion in [10]. The results are reported in table 2 below using Accuracy coefficient [9], and Dice coefficient [11].

TABLE II. RESULTS ON BRAINWEB PHANTOM IMAGES FOR NINE METHODS AND THE APPROACH WE PROPOSE.

Measurement	Approach	Inu. Noise	20%		
			0%	3%	5%
Accuracy coefficient (%)	Published work in [9] (FDS1)	Min Acc.	-	95.95	-
		Acc.	-	96.95	-
		Max Acc.	-	97.51	-
	Published work in [9] (FDS2)	Min Acc.	-	96.11	-
		Acc.	-	97.04	-
		Max Acc.	-	97.58	-
Dice coefficient	Published work in [10]	CSF	0.87	0.85	0.83
		WM	0.96	0.95	0.88
		GM	0.90	0.88	0.78
Dice coefficient	Our proposed approach	CSF	<b>0.96</b>	<b>0.94</b>	<b>0.91</b>
		WM	<b>0.97</b>	<b>0.96</b>	<b>0.93</b>
		GM	<b>0.96</b>	<b>0.95</b>	<b>0.90</b>
Accuracy coefficient (%)	Our proposed approach	Min Acc.	-	<b>97.09</b>	-
		Acc.	-	<b>97.16</b>	-
		Max Acc.	-	<b>97.86</b>	-

'-' that means no result is given in this case on the reference

The methods compared in table 2 have been run on images which have 0%, 3% and 5% of noise, 20% of intensity inhomogeneity (Inu.) and voxel size of 1mm3.

Regarding the performance of the fusion based methods, the proposed evidential fusion approach described in [9] is the worst (in terms of average accuracy  $\overline{Acc}$ , minimum accuracy Min acc. And maximum accuracy Max acc.) and the published work in [10] is the next worst, because the first one use focal elements and masses to represent data and the Dempster-Shafer rule to combine evidence. And in the second one, the data is modeled by the FCM algorithm, that is considered poorly to classify the pixels when they are situated very far to cluster centers. However, our approach is close to those proposed in [9] and [10]. Results of comparison show clearly the potential interest of our approach for magnetic resonance imaging (MRI) brain segmentation

## VII. CONCLUSION

In this paper, a new multispectral fusion approach for the segmentation of MR images is discussed. We outlined in here some features of possibility theory context, which can be very useful for medical images fusion. And which constitute advantages over classical theories. The results reported in this paper show the superior capabilities of fusion approach compared to the taking into account of only one weighting in MR image segmentation. As a perspective of this work, the cooperation of the algorithms of classification to modeling a data is desired. In fusion step, further studies to construct other adaptive operators is necessary. In addition, we can integrate other numerical, symbolic information, experts' knowledge or images coming from other imaging devices include computer tomography(CT), the newer positron emission tomography (PET) or a major functional modality SPECT in order to improve the segmentation of the MR images or to detect anomalies in the pathological images.

## REFERENCES

- [1] Y. Hata, S. Kobashi, and S. Hirano, "Automated segmentation of human brain mr images aided by fuzzy information granulation and fuzzy inference," *IEEE Trans. SMC*, vol. 30, 1998, pp. 381–395.
- [2] D. Goldberg-Zimring, A. Achiron, and S. Miron, "Automated detection and characterization of multiple sclerosis lesions in brain mr images," *Magnetic Resonance Imaging*, vol. 16, 1998, pp.311–318.
- [3] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of mr images of the brain," *IEEE Trans. Medical Imaging*, vol. 18, 1999, pp.897–908.
- [4] Y. Wang, T. Adali, J. Xuan, and Z. Szabo, "Magnetic resonance image analysis by information theoretic criteria and stochastic models," *IEEE Trans. Infor. Tech. in Biom.*, vol. 5, 2001, pp.150–158.
- [5] I. Bloch, and H. Maitre, "data fusion in 2D and 3D image processing: an overview," In *Proceedings of the X Brazilian symposium on computer graphics and image processing*, Brazil, 1997, pp.127–134.
- [6] C. Bezdek, J. Keller, R. Krishnapuram, and N. R. Pal *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Kluwer Academic, TA 1650, F89. 1999.
- [7] D. Dubois, and H. Prade, *Fuzzy Sets and Systems: Theory and Application*, New-York: Academic Press, 1980.
- [8] I. Bloch, "Information combination operators for data fusion : a comparative review with classification", *IEEE Transactions en systems, Man. and Cybernetics*, vol. 1, 1996, pp.52–67.
- [9] J. Z. Hongwei, and O. Basir, "Adaptive Fuzzy Evidential Reasoning for Automated Brain Tissue Segmentation," In *Proceedings of the 7<sup>th</sup> International Conference of Information Fusion*. Stockholm, Sweden, 2004.
- [10] C. Lamiche, and A. Moussaoui, "Improvement of brain tissue segmentation using information fusion approach," *Int. Jou. of Advan. Comp. Sci. and Applications*, vol. 2, 2011, pp.84–90.
- [11] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, 1945, pp.297–302.

# *EEG–fMRI fusion approaches and its applications*

*S.BELHADJ<sup>1</sup>, A.BENALI<sup>1</sup>, B.A.ADNANE<sup>2</sup>, Z.AHMED-FOITIH<sup>1</sup>, A.TALEB-AHMED<sup>3</sup>*

1. University of science and technology MOHAMED BOUDIAF, LEPESA Oran, Algeria
2. University Hospital of Oran CHUO, Department of neurosurgery, Oran, Algeria
3. University of Valencian and Hainaut Cambrésis LAMIH UMR CNRS 8201, France  
sabrina\_belhadj@live.fr

**Abstract**—The Many studies are currently collecting multiple types of brain imaging data and information from the same subject. Each imaging method reports on a limited domain and is provide some common information and some unique information. There is two way for utilize this data the first one is to analyze independently each type of information and then demonstrate the relationships between the different types (for e.g. EEG signals and functional images), the second way called information fusion which utilize multiple images-types together in order to take advantages of the cross data. In this paper, we will represent a state of the art of EEG-fMRI data integration. We will see why it is important to fuse these types of modalities, what are the limitations, solutions and finally in any field this technique is applied.

**Keywords**—*data integration; brain activity; medical imaging; EEG; fMRI.*

## I. INTRODUCTION

Brain Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) have been established as the most widely used non-invasive tools to map sensory, motor, cognitive, and emotional processes in the human brain [1]. They provide complementary spatio-temporal information about the brain function, it is obvious that each modality has its own advantages for example EEG reflect the brain changes on timescale of millisecond [2], which mean that it can capture the brain dynamic changes very well. But, it has poor spatial resolution due to being recorded from a limited number of electrodes. On the other hand, fMRI is powerful in investigation the brain function but it is slow to follow the brain activities because it relies on the brain blood flow response rather than electrical activities [3]. That's way the integration of information from these two modalities promises to provide a better understanding of brain function [4]. Various methods have been developed to combine high spatial resolution data provided by fMRI with high temporal resolution generated by EEG [5][6].

## II. LIMITATION AND SOLUTION

The principal limitations on multimodal EEG-fMRI integration are imposed by physiology: One reason why EEG

and fMRI sources might be dislocated is the distance between the neuronal population whose electrical activity is generating the EEG signal and the vascular tree, which provides the blood supply to these neurons, since BOLD signal changes are essentially hemodynamic [7]. Similarly, in addition to pre- and post synaptic electrochemical dynamics, a number of physiological processes also require energetic support. This differential sensitivity to neuronal activity and energetic can also arise whenever hemodynamic activity is caused by non-synchronized electrophysiological activity or if the latter has a closed source configuration that is invisible to EEG. Conversely, if the electrophysiological activity is transient, it might not induce any significant (i.e. detectable) metabolic activity changes [8]. Another important potential source of bias in EEG-fMRI fusion is experimental variability. In some situations, it might be necessary to acquire the EEG and fMRI data in separate sessions. In this case, habituation effects variations in the stimulation paradigm, or any other difference between sessions might lead to differential activity of neural networks [9]. Simultaneous EEG/fMRI acquisition techniques have been developed specifically to address these issues. Nevertheless, and despite advances in simultaneous EEG/fMRI hardware and software, the signal-to-noise ratio (SNR) of these signals is still significantly lower than the corresponding unimodal paradigms. This is mainly due to reciprocal electromagnetic perturbations [10]. Many data analysis have argued that dedicated modeling and signal processing tools should be used to combine the advantages of EEG and fMRI [11].

## III. EEG FMRI INFORMATION FUSION APPROACHES

During the past decade, an extensive literature has focused on analysis techniques devoted to EEG-fMRI integration, they classify them as in two categories: Asymmetric approaches and symmetric approaches (see figure 1).



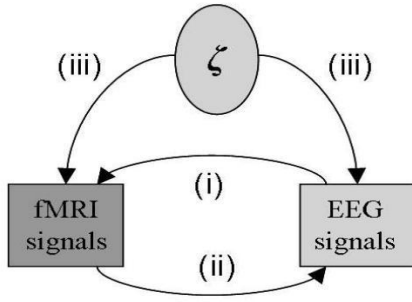


Figure 1: EEG-fMRI information fusion approaches (i) asymmetric EEG to fMRI approaches, (ii) asymmetric fMRI to EEG approaches, and (iii) symmetric fusion approaches (from EEG-fMRI fusion: biophysics and data analysis)

### 1. A. Asymmetric approaches

Asymmetric approaches for EEG-fMRI integration are characterized by a biased weighting of modalities in as much as information from one modality is used to guide the analysis of the other. Most influential of these approaches have been fMRI to EEG approaches and EEG and fMRI.

#### A.1 Asymmetric EEG to fMRI approaches

The objectives of these techniques is to localize, using fMRI, brain regions whose response is temporally correlated with given EEG-defined event or feature. In other words temporal information from the EEG signal is used as constraint or predictor variable in the fMRI time-series model. This type of EEG-fMRI integration is necessarily implemented within a simultaneous EEG-fMRI acquisition paradigm [12]. Other EEG to fMRI asymmetrical approaches have investigated the neuronal correlates of spontaneous cerebral activity occurring when the subject is not exposed to any extrinsic stimulation [13], these techniques try to find a linear decomposition of fMRI data which covariates with a time-frequency decomposition of the EEG [14].

#### A.2 Asymmetric fMRI to EEG approaches

Aims to alleviating the spatial EE inverse problem by guiding electromagnetic source imaging using results obtained from fMRI. To do this the head geometry and most relevant biophysical characteristics of the brain (e.g. Tissue conductivity) are first estimated to establish of forward model from which given a simulated neural event, the paths of currents to the scalp can be calculated. This approach can be divided into two classes, associated with the EEG source model employed: the equivalent current dipole (ECD) model and the distributed source model.

##### A.2.1 The equivalent current dipole (ECD) model

Dipolar fMRI to EEG approaches [15] simply associate each fMRI focus with the ECD whose position lies a prior at the centre of the activation [8]. This type of a prior constraint is hard in the sense that the results of fMRI analysis are not questioned, it has been shown that many ECDs are required to model spatially extended regions correctly [16].

##### A.2.2 The distributed source model [17]

In this technique, the model uses the fMRI activation as a prior on the spatial profile of cortically distributed sources. As a consequence, the estimation penalizes sources whose fMRI derived activation probability is low. This approach has been shown to estimate the position and extent of underlying sources robustly. These techniques require the tuning of a “hyperparameter” which regulate the weight of the penalty term, in relation to accuracy or model fit. This hyperparameter is cortical because it determines the balance between the effects of the fMRI and the EEG data on the estimation. Therefore, some authors [18] have proposed principal Bayesian techniques to optimize standard regularization procedures by estimating this hyperparameter from the EEG data.

### 2. B. Symmetric approaches

Symmetrical multimodal EEG-fMRI information fusion has been applied to the analysis of event-related bioelectric and hemodynamic responses. The objectives of any symmetric fusion multimodal EEG-fMRI information fusion should be twofold. First the approach should be able to identify the parts of EEG and fMRI signals that convey complementary information to decrease uncertainty when inferring on this common subspace. As a consequence a symmetrical fusion approach requires the explicit definition of the common neuronal states that engenders both EEG and fMRI measurements [11].

## IV. EEG FMRI INFORMATION FUSION APPLICATIONS

Since the first simultaneous acquisition of fMRI-EEG in humans [19], this non-invasive method is gaining attention in several research areas.

### 3. A. fMRI-EEG in epilepsy

The initial motivation for EEG-fMRI was in the field of epilepsy research [20], and in particular the study of interictal epileptic discharges (IED or interictal spikes), and their generators, and seizures. FDI are random events and subclinical patients with epilepsy that cannot be observed using the EEG (or MEG). Therefore EEG recording during fMRI acquisition allows us to study hemodynamic changes correlated in the brain during seizures. [21] The method can reveal hemodynamic changes associated with IED attacks, and proved to be a powerful scientific tool.

### B. fMRI- EEG for sleep study

Few studies have used fMRI- EEG during sleep. Activation of the occipital lobe and deactivation of frontal lobe could be detected during paradoxical sleep. [22] FMRI / EEG has also been used to characterize different types of areas during sleep and compare their location [23]. Other studies

regarding the auditory stimulation [24, 25] and visual stimulation [26] during sleep. This type of study is faced to the noise generated by the acquisition that can disturb sleep despite the use of quieter sequences [21]. The second difficulty is the length of acquisitions that can reach 5-7 hours to make recordings during paradoxical sleep.

### C. fMRI-EEG and evoked potentials

Because of their low amplitude which requires excellent quality of the EEG signal, the study of evoked potentials (EP) in fMRI-EEG is quite delicate. EP obtained in MRI appear similar to those obtained outside of the MRI [27]. Most dedicated Evoked Potentials fMRI / EEG studies have investigated a spatial location thereof. The first study of PE performed simultaneously with fMRI concerned visual evoked potentials [28]. It showed that the electrical sources of evoked potentials corresponded to the fMRI hemodynamic activations. However, the various components of evoked potentials may be associated with either activation or deactivation of a hemodynamic point of view. [29]

### D. fMRI-EEG and spontaneous rhythmic activity

Few studies have examined the localization of alpha rhythm using fMRI-EEG [30-31]. This rate appears on the EEG, especially in the posterior regions, when the subject closes his eyes. [32] Those studies have demonstrated the activation of the thalamus and sometimes the occipital lobe accompanied by deactivation of the parietal and frontal regions corresponding the active regions during the conscious resting state, called "default mode". [33] This state corresponds to the periods during which the brain works without specific instruction.

## V. CONCLUSION

We have present a state of the art of data integration, its remarkable that a variety of different methods have already been proposed the concurrent analysis of EEG and fMRI data. The methods chosen will strongly depend on the research question addressed interested (applications) in the neural generators of scalp EEG phenomena; fMRI informed EEG is the method of choice. Analysis based on EEG informed fMRI have already shown the ability not only to link physiological measures of different modalities with each other but also to expose associations between physiology and variations in cognition, perception and behavior. We have also seen that the EEG-fMRI data fusion in human systems, cognitive and clinical neuroscience is rapidly evolving and has received substantial attention.

## REFERENCES

[1] C Mulert, L.L. Lemieux, *Simultaneous EEG and fMRI*, Oxford, 2010.

- [2] M Danilo, M. Gozal, and al, *Signal and processing for knowledge extraction and information fusion*, Springer, 2008.
- [3] A.N Steibers, C.I. Bowman, and E.F. White, "Revision to the JDL data fusion model" *Sensor Fusion. Algorithms and applications*, vol 3 pp. 430-441, 1999.
- [4] V.D Calhoun and T. Adali "Feature Based Fusion of Medical Imaging Data" *biomedicine*.
- [5] G Bonmassar et al "Spatiotemporal brain imaging of visual evoked activity using interleaved EEG and fMRI recording" *Neuroimage*, vol 13 pp. 1035-1043, 2001.
- [6] V.D. Calhoun, T. Adali, G.D Pearlson, and E.F. Roberts, "Neuronal chronometry of target detection: Fusion of hemodynamic and event related potential data" *Neuroimage*, vol.30 pp. 544-553, 2006.
- [7] R. Beisteiner, M. Erdlerl and al, "Magnetoencephalography may help improve functional MRI brain mapping," *Eur .J. Neurosci*", vol 9 pp. 1072-1077, 1997.
- [8] A.B. Smith, C.D. Jones, and E.F. Roberts, "On the relationship of synaptic activity to macroscopic measurements: do coregistration of EEG with fMRI make sense," *Brain Topogr*, vol.13, pp. 79-96, 2000.
- [9] M. Wagner, M. Fuchs and al, "Integration of functional MRI, structural MRI, EEG and MEG," *Int. J. Bioelectromag*, vol 3, 2001.
- [10] J. Daunizeau, H. Laufs and K Friston, "EEG-fMRI fusion: biophysics and data analysis," *Welcome trust Centre of Neuroimaging*.
- [11] Z. Liu, L. Ding and al, "Integration of EEG/MEG with MRI and fMRI in functional neuroimaging," *IEEE Eng.Med.Biol. Mag*, vol 25, pp.46-53, 2006.
- [12] J. Daunizeau, C. Grova and al, "Symmetrical event related EEG-fMRI information fusion in variational Bayesian framework," *Neuroimage*, vol 3, pp.69-87, 2007.
- [13] H. Laufs, K. Krakow and al, "Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuation at rest," *PNAS*, vol 100, pp.11053-11068, 2003.
- [14] J.M. Kilner, J. Mattout, R. HENSON, K.J Friston, "Hemodynamic correlates of EEG: a heuristic," *Neuroimage*, vol 28, pp.280-286 2005.
- [15] S.J. Kiebel, M. Garrido and al, "Dynamic causal modeling of evoked response: the role of intrinsic connections," *Neuroimage*, vol 36, pp.332-345, 2007.
- [16] H. Shiraishi, S. Ahlfors and al, "Application of magnetoencephalography in epilepsy patients with widespread spike or slow-wave activity," *Epilepsia*, vol 46, pp.1264-1272, 2005.
- [17] J.K Friston, L.M. Harrison and al, "Multiple sparse priors for the M/EEG inverse problem," *Neuroimage*, vol 39, pp.1104-1120, 2008.
- [18] J. Mattout, C. Phillips and al, "MEG source localization under multiple constraints: an extended Bayesian framework," *Neuroimage*, vol 30, pp.753-767, 2006.
- [19] J.R. Ives, S. Warach, F. Schmitt and al. "Monitoring the patient's EEG during echo planar MRI". *Electroencephalogr Clin Neurophysiol*, vol.87: pp.417-420, 1993.
- [20] G. Bonmassar et al "Spatiotemporal brain imaging of visual evoked activity using interleaved EEG and fMRI recording" *Neuroimage*, vol .13 pp. 1035-1043, 2001.
- [21] J. Gotman, et F. Pittau, "Combining EEG and fMRI in the study of epileptic discharge," *Epilepsia*, vol 52, pp.38-42, 2011.
- [22] K.O. Lovblad, R. Thomas, P.M. Jakob and al. "Silent functional magnetic resonance imaging demonstrates focal activation

- in rapid eye movement sleep". *Neurology*, vol. 53: pp.2193-2195, 1999.
- [23] M. Schabus, T.T. Dang-Vu, G. Albouy and al. "Hemodynamic cerebral correlates of sleep spindles during human non-rapid eye movement sleep". *Proc Natl Acad Sci U S A*, vol. 104: pp.13164-13169, 2007.
- [24] M. Czisch, R. Wehrle, C. Kaufmann and al. "Functional MRI during sleep: BOLD signal decreases and their electrophysiological correlates". *Eur J Neurosci*, vol. 20: pp.566-574, 2004.
- [25] C.M. Portas, K. Krakow, P. Allen and al. "Auditory processing across the sleep/wake cycle: simultaneous EEG and fMRI monitoring in humans". *Neuron*, vol. 28: pp.991-999, 2000.
- [26] A.P. Born, I. Law, T.E. Lund and al. "Cortical deactivation induced by visual stimulation in human slow-wave sleep". *Neuroimage*, vol. 17: pp.1325-1335, 2002.
- [27] F. Kruggel, C.J. Wiggins, C.S. Herrmann, and D.Y. Von Cramon. "Recording of the event-related potentials during functional MRI at 3.0 Tesla field strength". *Magn Reson Med*, vol. 44: pp.277-282, 2000.
- [28] G. Bonmassar, K.J. Anami, Ives, and J.W. Belliveau. "Visual evoked potential (VEP) measured by simultaneous 64-channel EEG and 3T fMRI". *Neuroreport*, vol.10: pp.1893-1897, 1999.
- [29] K. Whittingstall, D. Wilson, M. Schmidt, and G. Stroink. "Correspondence of visual evoked potentials with FMRI signals in human visual cortex". *Brain Topogr*, vol. 21: pp.86-92, 2008.
- [30] R.I. Goldman, J.M. Stern, J.Jr. Engel, and M.S. Cohen. "Simultaneous EEG and fMRI of the alpha rhythm". *Neuroreport*, vol.13: pp.2487-2492, 2002.
- [31] M. Moosmann, P. Ritter, I. Krastel and al. "Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy". *Neuroimage*, vol.20: pp.145-158, 2003.
- [32] E. Niedermeyer. "Alpha rhythms as physiological and abnormal phenomena". *Int J Psychophysiol*, vol.26: pp.31-49, 1997.
- [33] M.E. Raichle, A.M. MacLeod, A.Z. Snyder and al. "A default mode of brain function". *Proc Natl Acad Sci U S A*, vol. 98: pp.676-682, 2001.

# L'utilisation de la réalité virtuelle pour la telerobotique

MASMOUDI. MOSTEFA <sup>(1)</sup>, L. KADDOUR EL BOUDADI <sup>(1)</sup>, Khelf. MOHAMED <sup>(2)</sup>, Dahane.AMINE <sup>(1)</sup>, A. LOUKIL <sup>(1)</sup>

<sup>(1)</sup>Intelligent Systems Research Laboratory (LARESI)

P.O. Box. 1505 - 31 000 Oran – Algeria

E-mail: safa-auto@hotmail.fr

<sup>(2)</sup> Laboratoire d'Électronique de Puissance, d'Énergie Solaire et d'Automatique (LEPESA)

P.O. Box. 1505 - 31 000 Oran – Algeria

E-mail: khelf\_mohamed@yahoo.fr

**Abstract :** *Cet article présente la commande à distance d'un robot mobile via Internet. Pour résoudre le problème de temps de retard qui est imprévisible, une architecture de téléopération est proposée. Cette architecture nous permet de minimiser l'erreur de trajectoire du mouvement de robot qui est commandé via internet. Ce travail présente aussi l'utilisation de la réalité virtuelle dans le contexte de la commande à distance. La réalité virtuelle peut être utilisée d'une manière classique afin de simuler le comportement d'un système, mais aussi en parallèle avec le système réel afin d'améliorer la qualité du contrôle. Pour valider notre travail, nous avons mené des expériences de téléopération dans divers lieux, et les résultats expérimentaux montrent l'efficacité de l'architecture proposée.*

**Keywords-** *teleoperation, Virtual Reality, mobile robot, qualité de service*

## I. INTRODUCTION

La télérobotique est un nouveau domaine de la robotique qui a attiré beaucoup d'attention des chercheurs au cours des dernières années [1, 2,3]. Le terme de télérobotique Internet se réfère à un Système de robot commandé à distance et surveillé à travers l'Internet. L'intérêt croissant dans ce domaine est stimulé par la promotion de l'Internet, qui donne accès à diverses ressources informatiques pratiquement de partout dans le monde. L'augmentation de la demande de l'Internet comme moyen de communication pour le système de télérobotique vient aussi du fait qu'il utilise un protocole de communication standard, et que le support physique est disponible pour l'application de télérobotique, éliminant le besoin pour développer un système de communication dédié, exclusif, et cher. La télérobotique via Internet diffère de celle-ci classique en ce qu'ils sont disponibles, lorsque cela est souhaité, pour le cas de multi utilisateurs. L'internet ne fournit pas une qualité de service garantie (QoS); il donne un certain nombre de limitations et difficultés, telles que contraintes de bande passante, délais de transmission, paquet perdu, connexion rupture, etc la situation ci-dessus influence la performance des systèmes de télérobotique basés sur Internet. [4, 22, 23], Ce qui est un nouveau champ dans la dernière décennie [24].

Il y a deux générations des robots commandés par Internet. Le projet Mercury [18] a été l'un des premières implémentations de télérobotique sur Internet, avec l'Australie

Télé robot [19] mis en ligne presque en même temps. Depuis, environ 40 de ces systèmes ont été mis en ligne par les équipes de recherche à travers le monde [21]. La première génération de robots commandés par internet a été principalement basée sur des bras robotisés ou des robots mobiles simples contrôlés directement par des opérateurs humains [18, 19, 20]. Ces robots commandés par Internet, opérés dans un environnement bien structuré avec peu d'incertitude, et n'ont pas l'intelligence locale. En revanche, la recherche sur la deuxième génération des robots commandés par internet a récemment commencé à se concentrer sur les robots mobiles autonomes qui naviguent dans un environnement dynamique et incertain [25, 26, 27]. Les principales caractéristiques de cette génération sont que les robots autonomes et réactifs, ce qui leur permet de naviguer et de faire face à l'incertitude dans le monde réel. L'objectif principal dans la construction de cette génération de robots commandés par internet est le contrôle à distance et la surveillance [28].

La téléopération de systèmes complexes en utilisant Internet comme canal de communication génère plus de difficultés telles que l'indisponibilité du système, l'imprécision et l'instabilité. Cela est dû à la qualité de service sur internet non garantie et surtout le retard instable de bout en bout [5, 6, 7]. Pour résoudre ce problème (la latence), Riyanto T. et Bambang développaient une architecture client –serveur basé sur la REALITY VIRTUAL pour commander à distance un manipulateur de robot à cinq degrés de liberté, appelé Mentor [8]. La majorité de ces recherches utilisent un système de contrôle de surveillance qui permet à l'opérateur d'émettre des commandes de haut niveau. Comme le délai internet est imprévisible, la conception d'un système de contrôle direct qui permet à l'utilisateur de contrôler le mouvement des robots continu peut ne pas être facile. De tels systèmes de contrôle ont été proposés, mais ne sont pas suffisants pour atténuer l'influence de retard sur l'internet. Pour pallier le problème de retard de temps, trois approches [23] sont souvent utilisées : the predictive aiding approach, the simulating, and planning display approach [5,9,15,16,17], and the event-based approach [29,30]. On trouve aussi d'autres approches; la stratégie "Move et Wait" [4], ce qui est typique pour les robots de l'espace a été appliqué en premier avec la communication à longue distance. Il y a d'autres chercheurs proposaient la stratégie « Speed Limit Module and Delay Approximator » [4] et le GEMMA-Q [10,11], ces solutions permettent de contrôler

la vitesse d'un robot mobile suivant la qualité de service afin que le robot atteigne la position désirée.

Notre objective est mis sur le domaine de la téléopération de robot à roues. Dans ce domaine, certains systèmes de télérobotique sur Internet ont utilisé le contrôle direct [20, 31,32]. De toute évidence, le contrôle direct n'est pas adapté pour les robots mobiles contrôlés par internet à cause de la limitation de la bande passante et des retards imprévisibles lors de la transmission à longue distance. Afin de surmonter ces problèmes, nous devons améliorer considérablement notre architecture (client-serveur) de système et de l'interface homme-machine pour commander à distance un robot mobile (Pioneer P3AT) dans un réseau sont qualité de service garantir. L'application client contient une java applet, retour visuel et une application 3D pour simuler le comportement de notre system (Pioneer P3at), et l'application serveur contient une architecture qui nous permet de retarder les commandes envoiaient ver le robot avec un retard connaît et constant afin d'éviter les variations aléatoires de RTT.

Cet article est organisé comme suit: en section II, nous mettrons L'architecture du système de téléopération. La solution proposée est décrite en section III et les résultats des expériences sont discutés en section IV. Finalement, les conclusions et perspectives sont présentées en section V.

## II. ARCHITECTURE DU SYSTEME DE TELEOPERATION

Notre système de téléopération est composé d'un PC hôte distant et d'un robot mobile (pioneer P3AT) communiquant entre eux par internet de type client serveur en utilisant le protocole de communication Internet TCP/IP (figure 1).

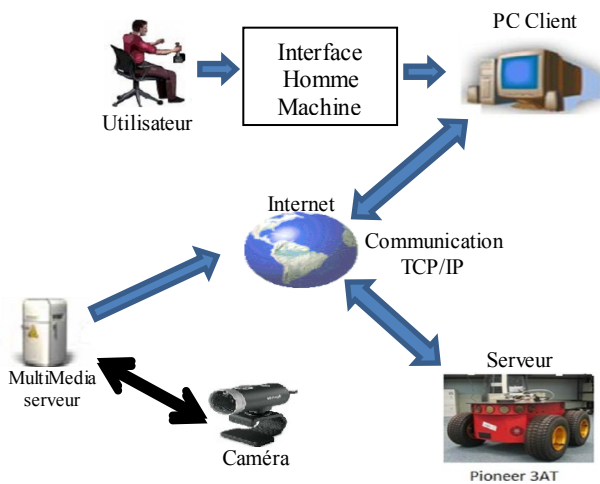


Figure 1 Structure détaillée du système de téléopération.

L'opérateur via une interface graphique de communication homme machine contrôle à distance les mouvements du robot (avant ,arrière ,gauche et Droite) par l'intermédiaire d'un PC client, tout en ayant la possibilité de voir la scène grâce à la caméra et transmises via le canal de transmission (internet). Par la suite, nous donnons une explication de chaque module.

### II.1. L'interface de communication homme machine :

C'est l'espace par lequel l'opérateur dialogue avec le robot. Ce module assure pour l'opérateur trois tâches essentielles, il prend en charge les commandes de l'utilisateur, affiche les images vidéo et gère le Simulateur virtuel du robot pour permettre à l'utilisateur de voir les tâches que le robot doit exécuter. La figure ci-dessous (figure 6) illustre bien la fenêtre principale de l'interface développée. Cette dernière se composait de deux parties.

#### A. Le client

L'application client a été développée par le langage de programmation java (NetBeans), est installée sur le PC hôte, elle assure à son tour les fonctions suivantes :

- La gestion d'une interface de communication graphique homme machine
- La gestion de la communication depuis et vers le robot, par l'envoi des consignes de l'utilisateur.
- La réception, l'affichage, des images vidéo issues de la caméra.
- La prise en charge et l'interprétation des commandes données par l'utilisateur.

#### B. La simulation du robot

Nous avons développé un simulateur graphique (figure 2) de notre système robotique, que l'on a appelé (simulateur du robot mobile Pioneer PA3T). Ce simulateur a été développé, avec Unity 3D logiciel de moteur de jeu. On a conçu ce simulateur dans le but de pouvoir tester les différentes stratégies possibles dans les conditions réelles d'utilisation. Le robot virtuel reproduit fidèlement le comportement réel du robot.

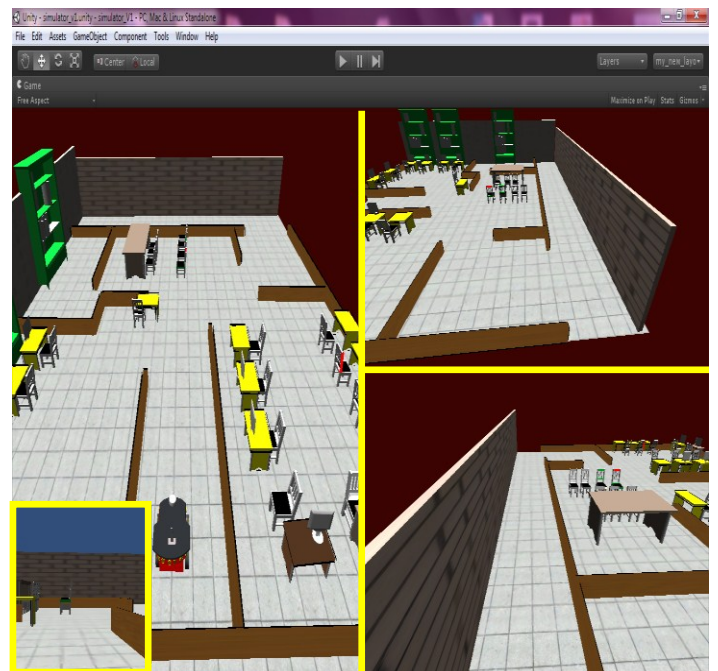


Figure 2 le simulateur développé par Unity 3D

## II.2 Le serveur

Nous avons développé une application serveur par le langage de programmation C++, Cette application est installée sur le PC embarqué dans le robot (figure 3), elle assure les fonctions suivantes :

- La gestion de la communication depuis et vers le pc distant, par l'envoi des informations sur l'état du robot et la réception des consignes de l'utilisateur.
- Le contrôle du mouvement du robot.
- La gestion des capteurs du robot.

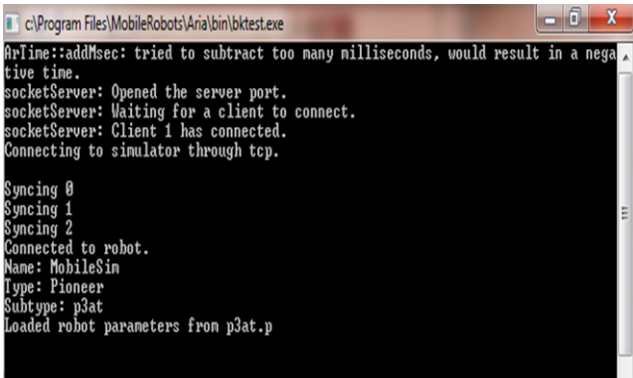


Figure 3 L'interface serveur du Pioneer-3AT

## II.3 communication internet

Le problème souci principal des applications de télérobotique est la limitation de la bande passante et des retards imprévisibles lors de la transmission à longue distance. Afin de surmonter ces problèmes, nous devons améliorer considérablement notre architecture de système et de l'interface homme-machine comme nous allons le montrer dans la suite. Nous avons utilisé une architecture client-serveur avec un protocole TCP / IP. Tous les liens entre l'opérateur et le lieu de travail sont mis en œuvre en utilisant les sockets TCP / IP pour permettre une téléopération de longue distance grâce à Internet.

## III. LA SOLUTION PROPOSÉE

L'architecture générique de la télérobotique est la même dans la plupart des applications. Le problème principal est d'atténuer l'influence de la variation du retard sur internet (RTT), l'impact de la QoS apparait clairement sur un robot mobile, car à travers le réseau internet la commande atteint le robot avec un retard imprévisible  $\Delta t$  ; le robot qui se déplace à la vitesse  $V$  peut se déplacer d'une distance supplémentaire  $\Delta d$  avant de recevoir la commande pour tourner ou bien s'arrêter comme illustré sur la (figure 4). On remarque que l'erreur sur la trajectoire s'accumule, le robot n'atteint pas la position désirée. Pour résoudre ce problème les chercheurs dans la REF [4] ont proposé une architecture Speed Limit Module (SLM) qui permet de varier la vitesse de robot suivant la variation du RTT, mais il n'est pas facile de trouver un SLM uniforme pour tous les utilisateurs. À ce problème, nous avons proposé une architecture qui nous permet de minimiser l'erreur sur la trajectoire du robot commandé via internet ;

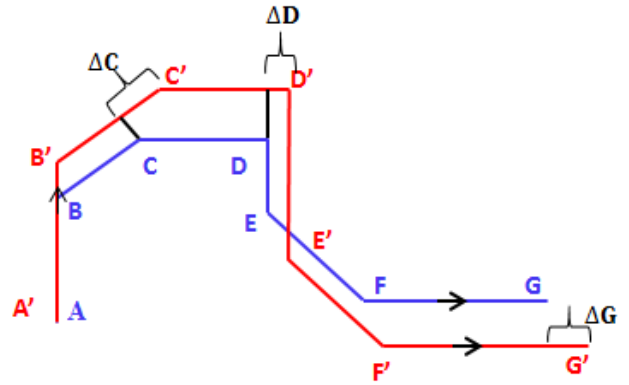


Figure 4 Impact du retard sur internet sur le déplacement d'un robot mobile

$$\Delta d = V \cdot \Delta t \quad 1$$

le principe est pour atteindre à la position désirée « G », il doit être avoir une valeur constante pour la vitesse du robot «  $V$  » et le temps d'aller de l'information de client vers le robot «  $T$  », (équation 1). Pour la vitesse du robot est facile de donner une valeur fixe, mais le problème dans la valeur «  $T$  » dus elle est variée en fonction de temps et de façon aléatoire, pour la fixer nous avons proposé un algorithme qui nous permet de retarder les commandes envoyaient vers le robot avec un temps de retard constant (TMC=500ms), comme montre sur (figure 5).

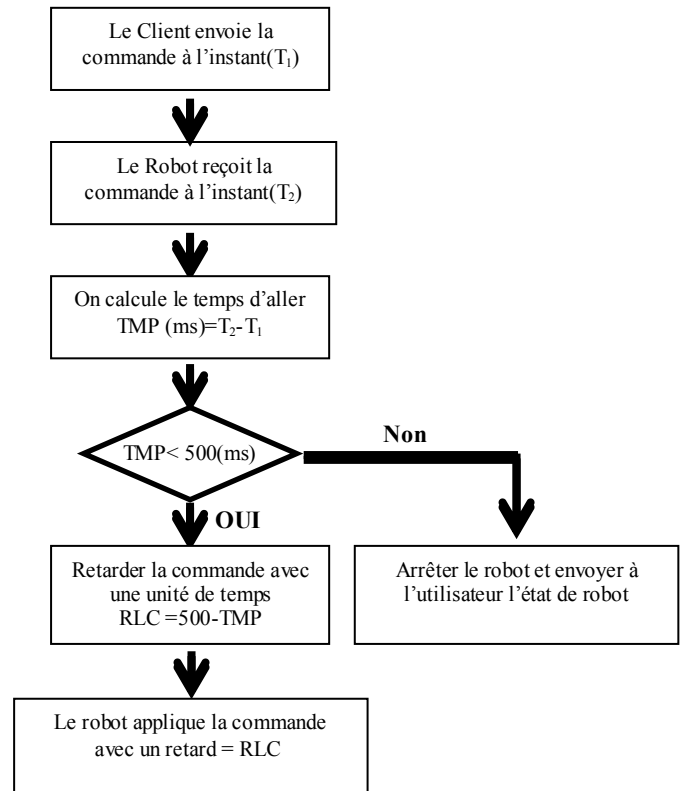


Figure 5 Procédé et système de retardement des commandes

Te d'abord le client à distant envoie les commandes (avant, arrière, gauche et Droite) à l'instant T1 puis le robot les reçoit à l'instant T2 ensuite on calcule le temps d'aller de la commande  $TMP=T1- T2$  au niveau de serveur(Robot) puis on teste si  $TMP<500$  (ms) on retardera la commande avec un temps ( $RLC= 500-TMP$ ), sinon on arrêtera le robot. Cette solution capable pour résoudre la problématique de la latence, comme montre dans la section suivante.

#### IV. EXPERIMENTS

Dans ce travail, deux types d'expérimentation ont été réalisés. La première est la synchronisation entre les mouvements de l'objet virtuelle (simulateur) et le mouvement réel. Les résultats montrent que la simulation 3D correspond bien avec le mouvement de robot réel. Cependant, nous avons trouvé un problème, pour déplacer le robot surtout quand il est dans une position difficile et effectue un mouvement complexe.



Figure 6- Commande basée sur un modèle virtuel

La seconde expérimentation porte sur l'implémentation de l'architecture proposée sur le robot, et pour ce faire nous avons mené des expériences de téléopération sur différentes distance entre éloignent l'opérateur et le matériel respectivement (100m, 1km, 5km et 130km) (figure 7).

La principale nouveauté de ce travail est l'architecture qui nous permet de retarder les commandes qui ont envoyé, de client vers le serveur (Robot) avec une unité de temps connu, cette solution nous permet de connaître avec une très petite erreur le temps de retard qui est imprévisible afin de l'opérateur d'ajuster et localiser la position de robot, mais la

précision a une relation avec la distance entre l'opérateur et le robot, et ce dernier peut atteindre la destination avec succès (figure 8). Dans l'ensemble du processus, aucune collision arriver. Les expériences montrent que cette architecture a une bonne praticabilité et de sécurité.

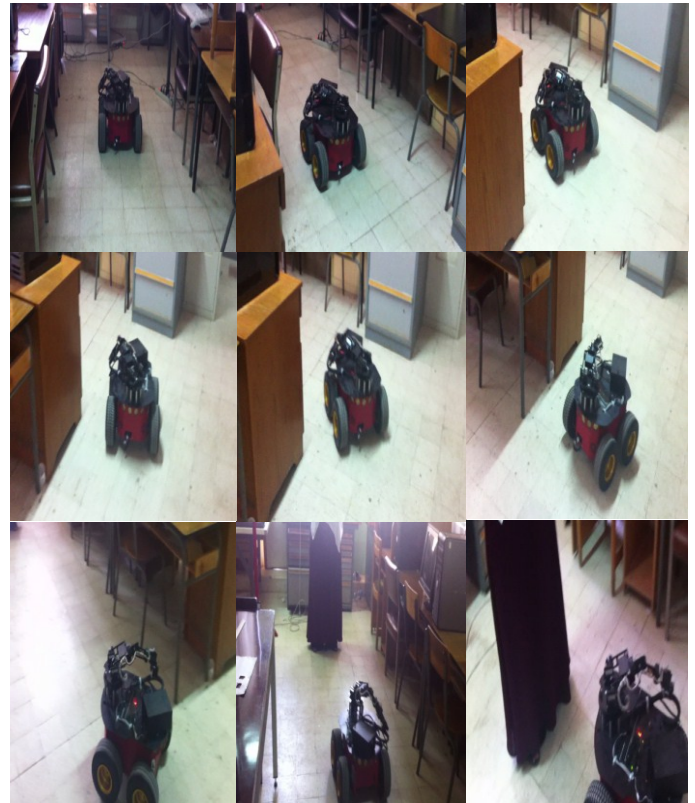


Figure 7 Résultat de la Télé-opération

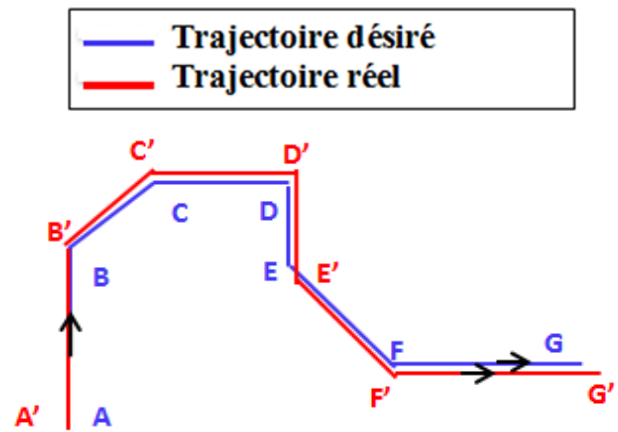


Figure 8 résultats de chemin en utilisant l'architecture proposée

#### V. CONCLUSION:

Le temps de retard sur internet est l'un des plus grands obstacles de nombreuses applications en temps réel, en particulier pour les systèmes de contrôle basés sur Internet. Pour résoudre ce problème, nous avons proposé une

architecture pour améliorer la qualité du contrôle d'un système à distant. Dans un premier temps, nous avons établi une communication réseau entre le pc client et le robot Pioneer P3AT, et pour compenser la latence induite par le réseau. Nous avons développé une architecture qui nous permet de retarder les commandes qui ont envoyé, de client vers le robot avec une unité de temps connu, cette solution autorise le robot d'atteindre la destination désirée avec succès.

Nous avons aussi développé un simulateur, qui est un banc d'essai virtuel, il suscite pour nous beaucoup d'intérêt il, nous a permis de voir la configuration du robot avant et après et lors l'exécution d'une tâche. L'opérateur donc aura une imagination complète de l'action à faire par le robot.

Les expériences qui ont fait nous permettent de conclure que l'architecture proposée donne nous des résultats prometteurs. Dans les futurs travaux, il serait intéressant d'améliorer les performances de l'algorithme proposé afin de commander le robot, même quand le temps de retard dépassera la seule de 500milli seconde.

## REFERENCES

- [1] Conway L, Volz R A, Walker M W. Teleautonomous systems: Projecting and coordinating intelligent action at a distance. IEEE Transactions on Robotics and Automation, 1990, 6, 146–158.
- [2] Bambang R. Development of Internet telerobotics systems. Proceedings of Information Systems Management Research Concentration, France, 2002.
- [3] FU Xiu-hui, HUA Jian-ning, ZHENG Wei etc. Study on task produced the method of Internet-based teleoperation robot. Micro-Computer Information [J], 2006, 3-2: 185-187.
- [4] Shengtong Zhong, Philippe Le Parc, Jean Vareille, "INTERNET-BASED TELEOPERATION: A CASE STUDY TOWARD DELAY APPROXIMATION AND SPEED LIMIT MODULE", ICINCO 2007, Proceedings of the Fourth International Conference on Informatics in Control, Automation and Robotics, Robotics and Automation 1, Angers : France (2007)
- [5] L. Kaddour-El-Boudadi, J. Vareille, P.L. Parc, N. Berrached, "Remote control on Internet, long distance experiment of remote practical works, measurements and results", International Review on Computers and Software (IRECOS), vol. 2, n°3, May 2007.
- [6] E. Witrant, C. Anudas-de-wit, D. Georges, and M. Alamir," Remote stabilization via time-varying communication network delays: Application to TCP networks", IEEE Conference on Control Applications, Taipei, Taiwan, and September 2-4, 2004.
- [7] I.Dilaneh, L. Laval,"Stabilization of networked control systems with uncertain time-varying delays", 17thMediterranean Conference on Control & Automation, MED'09, Greece, June 24-26, 2009.
- [8] Riyanto T. Bambang , Development of Architectures for Internet Telerobotics Systems, Journal of Bionic Engineering (2007) Vol.4 No.4
- [9] Philippe Le Parc, , E. Pardo, A. Touil, J. Vareille, " Virtual Reality to Improve Remote Control in Presence of Delays", published in "2010 IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems (VECIMS 2010), Tarente : Italy (2010)"
- [10] Ogor, P., 2001. Une architecture générique pour la supervision sûre à distance de machines de production avec Internet. Ph.D. thesis, Université de Bretagne Occidentale.
- [11] ADEPA, 1981. GEMMA: Guide d'études des modes de marche et d'arrêt.
- [12] X. Yang, D. C. Petriu, T. E. Whalen, E. M. Petriu, "A Web-Based 3D Virtual Robot Remote Control System", CCECE 2004- CCGEI 2004, Niagara Falls, May/mai 2004-7803-8253-6/04/517.00 02004 IEEE.
- [13] Ding Cheng-jun , Duan Ping , Zhang Ming-lu Zhang ,Yan-fang," Design of Mobile Robot Teleoperation System Based on Virtual Reality", Proceedings of the IEEE International Conference on Automation and Logistics Shenyang, China August 2009, 978-1-4244-4795-4/09/\$25.00 © 2009 IEEE.
- [14] Feng Cui, Ming-Lu Zhang, Bing-Qing Liu ,," DESIGN OF THE MENDING ROBOT BASED ON VIRTUAL REALITY AND INTELLIGENT DECISION SYSTEM", Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 13-16 August 2006, 1-4244-0060-0/06/\$20.00 ©2006 IEEE.
- [15] Tiam Hee Tee, Kok Seng Eu, Kian Meng Yap, Alan Marshall , Tsung-Han Lee;" 3D Smart User Interactive System with Real-Time Responding Tele-Robotic Proprioceptive Information", 2013 4th International Conference on Intelligent Systems, Modelling and Simulation, 2166-0662/13 \$26.00 © 2013 IEEE DOI 10.1109/ISMS.2013.90
- [16] Feng Cui, Minglu Zhang, and Shenghui Guo," Design of the Intelligent Control System of Telerobot Based on Virtual Reality", Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06) 0-7695-2616-0/06 \$20.00 © 2006
- [17] Lahouari KADDOUR–EL BOUDADI, Nassreddine BERRACHED,"MULTI USERS INTERNET BASED TELEOPERATION STATION WITH AUGMENTED REALITY ",Acta Electrotechnica et Informatica No. 2, Vol. 7, 2007
- [18] K. Goldberg, S. Gentner, et al., The Mercury project: a feasibility study for Internet robots, IEEE Rob. Autom. Mag. 7 (March (1)) (2000) 35–40.
- [19] K. Taylor, J. Trevelyan, Australia's telerobot on the web, in:International Symposium on Industrial Robots, 1995, pp. 39–44.
- [20] P. Saucy, F. Mondada, KhepOnTheWeb: open access to a mobile robot on the Internet, IEEE Rob. Autom. Mag. 7 (March (1)) (2000) 41–47.
- [21] K. Goldberg, R. Siegwart, Introduction, in: K. Goldberg, R. Siegwart (Eds.), An Introduction to Online Robots, The MIT Press, Cambridge, England, 2002, pp. XV–XXI.
- [22] K. Brady, T.J. Tarn, Handling latency in Internet-based teleoperation, in: K. Goldberg, R. Siegwart (Eds.), An Introduction to Online Robots, The MIT Press, London, England, 2002, pp.171–192.
- [23] R.C. Luo, K.L. Su, Networked intelligent robots through theInternet: issues and opportunities, IEEE Proc. 91 (3) (2003)371–382.
- [24] R. Siegwart, K. Goldberg, Robots on the web, IEEE Rob. Autom. Mag. 7 (March (1)) (2000) 4.
- [25] R. Simmons, Fernandez, et al., Lessons learned from Xavier,IEEE Rob. Autom. Mag. 7 (2) (2000) 33–39.
- [26] S. Thrun, M. Bennewitz, et al., MINERVA: a second-generation museum tour-guide robot, IEEE Int. Conf. Rob. Autom. 3(1999) 1999–2005.
- [27] R.C. Luo, T.M. Chen, Development of a multi-behavior based mobile robot for remote supervisory control through the Internet, IEEE/ASME Trans. Mechatronics 5 (December (4)) (2000)376–385.
- [28] T.B. Sheridan, Telerobotics, automation and human supervisory control, The MIT Press, London, England, 1992.
- [29] I. Elhajj, N. Xi, et al., Supermedia-enhanced Internet-based telerobotics, Proc. IEEE 91 (3) (2003) 396–421.
- [30] P. Li, W. Lu, Implementation of an event-based Internetrobot teleoperation system, in: Fourth World Congress on Intelligent Control and Automation, vol. 2, 2002, pp. 1296–1300.
- [31] H. Huosheng, Y. Lixiang, et al., Internet-based robotic systems for teleoperation, Int. J. Assembly Autom. 21 (2) (2001) 1–10.
- [32] K.H. Han, S. Kim, Y.J. Kim, J.H. Kim, Internet control architecture for Internet-based personal robot, Autonom. Rob. 10(2001) 135–147.



# SVM based gender classification

OULAD KADDOUR Mohamed<sup>1</sup>, BENATCHBA Karima<sup>2</sup>

<sup>1</sup>STIC doctoral school,

High school of computer science,

Alger, Alegria

<sup>1</sup>m\_ouled\_kaddour@esi.dz

<sup>2</sup>LMCSI

High school of computer science,

Alger, Algeria

**Abstract**— Gender classification is an important and challenging task for computer vision because it is recommended for a lot of fields. In this paper, we propose an approach for gender classification from faces images that is based on support vector machine (SVM). In this approach a majority voting combining technique is applied for the output of three SVMs that have been trained for three thumbnails features. In the experimentation step we have used the FERET with a WWW datasets and an accuracy of 92% was obtained.

**Keywords**— Gender classification, Face detection, Support vector machine, Face identification, Biometrics systems, Majority voting

## I. INTRODUCTION

Gender classification is a biometric task that consists to predict the gender of a person from a biometric modality. Formally, gender classification is a binary classification problem in the role is to affect an example data to one of two complementary classes: man or woman [01].

Recently, gender classification has attracted several searchers and various systems have been proposed. Indeed, gender classification is greatly recommended for a lot of domains in which it allows a considerable advantage. Among these domains we can mention the computer-human interaction, the biometrics systems, the security, the target commerce and many others [04].

In the practice, divers biometrics modalities have been used for performing gender classification systems as iris images [09], hand shape [10], body analyse [11], etc. However, the face modality is the most exploited because it's the most powerful and vigorous.

Classically, a gender classification system is organized in three modules: pre-processing, features extraction and classification [02]. In the pre-processing step, preliminary filters are applied to the input image for improving details and face is extracted. In the literature, the size normalization, histogram equalization and Viola-Jones algorithm [5] for face detection are useful techniques for pre-processing. In the features extraction step, a determinist vector is calculated

from the pre-processed image. Features extraction based, a gender classification approach is categorized as global if a same operator is applied for the whole image to compute the feature and as local if the image is decomposed into subcomponents [03]. The Principal Components Analyse, Local Binary Pattern and Gabor filters are useful techniques for feature extraction. In the classification phase, a binary classifier is used to predict the gender of the input image. For this step, the Neural networks, Support Vector Machines and the Adaboost algorithms are powerful and robust techniques [04].

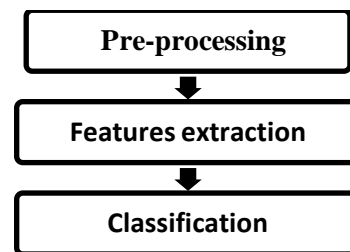


Fig. 1: Steps for gender classification

In this paper, we propose a gender classification approach from face images based on the support vectors machines classifiers with the vote majority combination technique. For each classifier a thumbnail feature is extracted from different masked face images which differ in the contended face regions.

The rest of this paper is organized as follows: (2) The proposed system overview, (3) Pre-processing and face detection algorithm presentation, (4) Features extraction (5) Classifiers and majority voting presentation (6) Experimentations and results discussion.

## II. SYSTEM OVERVIEW

It was observed that the classifiers performances are affected by the included information's exploited for features extraction and the use of some parts of face can decrease the robustness. In the same, a single classifier can be judged as

insufficient and the combination of classifiers ensemble can be utile to increase the global performance [04].

So, knowing that a well accuracy can be obtained for gender classification from face image with an intensity deduced feature the overview of our approach is illustrated by the following figure.

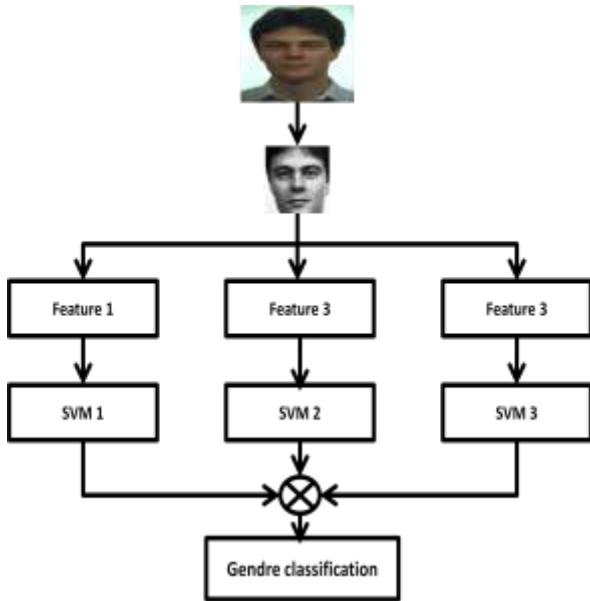


Fig. 2: Overview of the proposed approach

After preliminaries filters and face detection, three features are calculated. The first feature is corresponding to the intensity pixels values of the integral detected face, the two others features are corresponding to the intensity pixels values of the masked faces using two different masks. The first mask is exploited to eliminate the hair part and the second mask is exploited to obtain the internal face by eliminating the border and hail regions.

Thus, the idea of our proposition is to performing a profitable features gender classifications approach with taken into consideration different scenarios of prediction according to the regions contained in the face and the final decision is relative to the integrity of these divers' scenarios.

### III. PRE-PROCESSING AND FACE DETECTION

In the proposed approach the histogram equalization is used as a pre-processing technique to make the intensity distribution more equitable in the input image. After the histogram equalization, the face is detected using the viola Jones technique that is efficacy and rapid algorithm for face extraction. In this algorithm, the face detection is performed using a supervised boosted classifier that exploit the Haar features for doing a cascades scans in a computed matrix for the input image, named integral image. The integral image allows a rapid detection of preliminary interesting rectangles

for which the Haar features are used as mask to take intermediary decisions about face localisation.

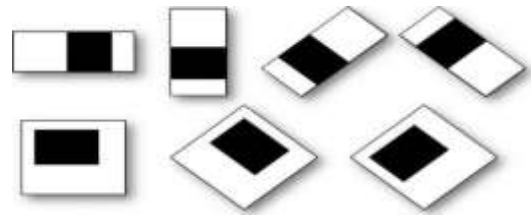


Fig. 3: Examples of Haar features

The following figure Fig. 4 shows an example for face localization using the Viola-Jones algorithm.

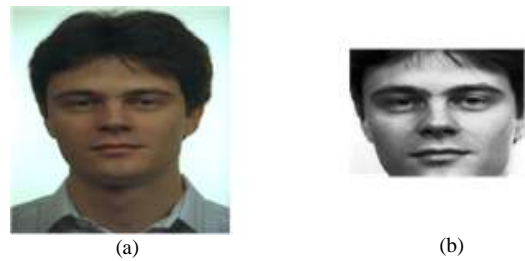


Fig. 4: Face extraction using the Viola-Jones algorithm  
(a) input image (b) detected face

The face extraction is an important pre-processing step because it allows to reducing information quantity and to eliminate the non facials input data for each an error message can be indicated.

### IV. FEATURE EXTRACTION

For feature extraction, we have used an intensity based features. After pre-processing, the gray faces images are normalized into  $24*24$  and simply converted into integer's vectors, the principle of these features is showed by the following figure.

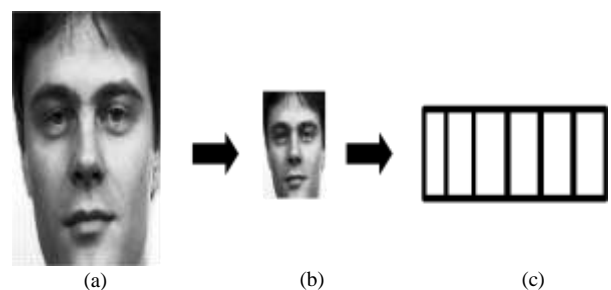
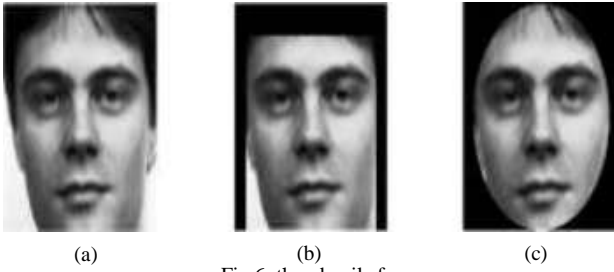


Fig. 5: Thumbnail feature  
(a) - Detected face (b) - Thumbnail face (c) - Intensity feature vector

In the same, the figure Fig 6 illustrates the exploited thumbnails images for computing the three features which correspond respectively to the detected face and two others masked faces.



(a) – Integral detected face (b) – First mask: face without hair Face (c) – Second mask: internal face

After vectors construction and as mentioned above, the computed features for the different thumbnails will be passed to different classifiers and each classifier return an intermediary gender prediction class.

## V. CLASSIFICATION

For classification, we have used the Support Vector Machines classifiers that are supervised machines learning. The SVM was firstly proposed by Vapnik [06] and it was introduced for the gender classification problem by Moghadem [07].

We have opted for the SVM classifiers because it has been proven that they are robust and powerful for gender classification with face images, in comparison with others classifiers [04].

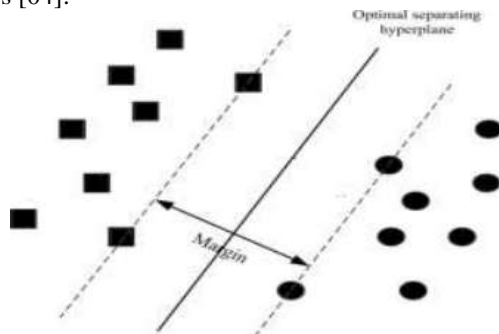


Fig 7: Support Vectors Machines principle

The principle of the SVM in the case of linearly separable class is to construct a hyper-plan that has a largest distance to the nearest data points of any class. Generally, the manipulated data are not linearly separable. In this case, the SVM use a nonlinearly kernel function named  $K$  to construct the boundary function  $f(x)$  that is defined by the following formula where  $b$  is the bias value and  $y_i$ ,  $a_i$  covers a Lagrange optimization parameters :

$$f(x) = \sum_{i=1}^M y_i \alpha_i K(x, x_i) + b$$

In our case the SVM is used as binary classifier that affects an example to one of two gender class: man or woman and

polynomials kernels were exploited in the experimentations, one classifier for feature and the final decision is calculated by the majority voting as showed in the table 01.

SVM1 Class	SVM1 Class	SVM1 Class	Final Class
W	W	W	W
W	W	M	W
W	M	W	W
W	M	M	M
M	W	W	W
M	W	M	M
M	M	W	M
M	M	M	M

Table 1: Majority venting principle

In the next section, the context and experimental results will be discussed for the proposed approach.

## VI. EXPERIMENTATION

In the experimentation, we have used for the training phase a subset of 200 frontal images from the experimental FERET [08] dataset that contain different facials variations for ethnic, age, pose, illumination and facial expression.



Fig. 8: Example of the FERET databases Images

In the testing phase, we have exploited a second subset of frontals images from the FERET dataset with a WWW dataset that we have constructed by regrouping frontals face images from the World Wide Web.



Fig 9: Examples of WWW images

The details of the used datasets in the testing phase for the three classifiers are showed in the table 02.

	<b>FERET</b>	<b>WWW</b>
<b>DataBases size</b>	800	800

Table 2: Datasets details

The accuracy of the gender classification of each classifier was calculated by the formula:

$$Accuracy = \frac{NTP}{TDZ} * 100\%$$

where *NTP* = Number of Trues Prédictions and *TDZ* = the Test Dadabase Size.

The obtained results with the FERET dataset are presented in the table 03.

<b>SVM1</b>	<b>SVM2</b>	<b>SVM3</b>	<b>Combinati on</b>
91,5%	90,75%	90,12%	92,87%

Table 3: Accuracy for gender classification with the FERT dataset

In the same, the results with the WWW dataset are presented in the following table 04.

<b>SVM1</b>	<b>SVM2</b>	<b>SVM3</b>	<b>Combination</b>
87,12%	85,62%	85,12%	89%

Table 4: Accuracy for gender classification with the WWW dataset

By analysing the obtained results, we are finding that the integral face detected by the Viola-Jones algorithm has returned the best accuracy for gender classification with the support vectors machines classifiers that was introduced by Mogadam [7] in comparison with the cases of elimination of hair part and the internal face. In addition, we observe that the used features provide a complementary decision and the majority voting inter SVMs classifiers has improved the global classification accuracy.

In this paper, a support vectors machines classifiers based approach was proposed for the gender classification from face images. The classifiers were trained with thumbnails features corresponding to different face parts. It was observed that the better individual classification accuracy is obtained with the integral Viola-Jones detected face and the majority voting combination technique has allows a classification accuracy gain.

REFERENCES

[01] A. Roayaei Ardakany, M. Nicolescu, M. Nicolescu, "An Extended Local Binary Pattern for Gender Classification", IEEE International Symposium on Multimedia, 2013.

[02] S. A. KHAN, M. NAZIR, N. RIAZ, "Optimized features selection for gender classification using optimization algorithms", Turkish Journal of Electrical Engineering & Computer Sciences, 2013.

[03] A. R. A rdakany, M. Jou la, " Gender Recognition Based on Edge Histogram", International Journal of Computer Theory and Engineering Vol. 4, No. 2, April 2012

[04] E. Mäkinen, and R. Raisamo, " An experimental comparison of gender classification methods ". Pattern Recogn. Lett. 29, 10 , pp. 1544-1556, Jul. 2008.

[05] P. Viola and M. Jones, " Rapid object detection using a boosted cascade of simple features", Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, pp. 511-51, 2001.

[06] V. Vapnik, " Support-Vector Networks ", Machine Learning, 20, 1995.

[07] B. Moghaddam and M.-H. Yang, " Gender Classification with Support Vector Machines ", Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition. 2000.

[08] P. J. Phillips, et al., " The FERET database and evaluation procedure for face-recognition algorithms", Image and Vision Computing, vol. 16, pp. 295-306, 1998.

[09] A. Bansal, R. Agarwal, R.K. Sharma, " SVM Based Gender Classification Using Iris Images", Fourth International Conference on Computational Intelligence and Communication Network, IEEE, 2012.

[10] GH. Amayeh, G. Bebis, M. Nicolescu., " Gender Classification from Hand Shape." Computer Vision and Pattern Recognition Workshops, pp. 1-7, 2008.

[11 ] L. Cao, M. Dikmen, Y. Fu and T. S. Huang, " Gender recognition from body", Proceeding of the 16th ACM International Conference on Multimedia, pp. 725- 728, 2008.

# Localisation de zones surchauffées dans les installations électriques par thermographie infrarouge

<sup>1</sup>Belkhodja Leila, <sup>2</sup>Hamdadou Djamilia

<sup>3</sup>Keddi Abdelhamid, <sup>4</sup>Fadhel Smail

<sup>1,3,4</sup>*Institut de Maintenance et de Sécurité Industrielle IMSI-Université d'Oran2 Algérie*

<sup>1,2</sup>*Laboratoire LIO, Département d'Informatique, Université d'Oran Algérie*

<sup>1</sup>Belkhodja\_leila@yahoo.fr

<sup>2</sup>dzhamdadoud@yahoo.fr

**Résumé**— Ce travail entre dans le cadre des études portant sur la maintenance préventive en milieu industriel et en particulier dans les installations électriques où le danger peut arriver rapidement et sans prévenir, engendrant avec lui des dégâts matériels et humains très graves. Prévenir ce problème revient à le contrôler à temps par élimination de causes susceptibles d'engendrer de lourds coûts économiques et humains dans une chaîne de production. Une de ces causes est l'échauffement dans les installations électriques qui a comme conséquences feux, explosions, etc... L'apport de la thermographie dans ce domaine est très bénéfique puisque cette technique permet une bonne maintenance préventive de ces installations à travers le traitement de l'image thermique. Le traitement automatique de l'image thermique se réduit en vision artificielle à une simple tâche de classification de l'image couleur et donc à un problème de reconnaissance des formes en intelligence artificielle. Nous avons mené cette tâche à l'aide d'une approche très simple en classification rigide qui est le K means.

**Mots clés**— Vision industrielle, Image thermique, Classification couleur, K means.

## I. INTRODUCTION

La vision industrielle, approche parallèle à la vision par ordinateur ou la vision artificielle, est selon UKIVA l'application des techniques de la vision artificielle à des problématiques de production afin de reproduire le plus fidèlement possible la vision humaine[1]. La maintenance préventive dans une chaîne logistique est un des objectifs de la vision industrielle. Elle s'effectue à l'aide d'inspections non-destructives sur des équipements de déceler toute détérioration ultérieure.

La vision industrielle, a réussi à prouver son efficacité dans tous les domaines où l'émergence des nouvelles technologies est remarquable. Elle est présente dans le domaine de l'automobile pour le montage de pare-brises et le contrôle d'aspect de disque et de plaquettes de freins. Elle est utilisée aussi en agroalimentaire pour mesurer le niveau du liquide lors du remplissage des bouteilles, le contrôle des logos et des marquages sur les emballages. Le marché de la vision est aussi présent dans la télésurveillance, l'informatique médicale, le contrôle financier et la biométrie. [2].

Les développements récents des caméras infrarouges, largement employées pour la prévention et la lutte des

incendies, ont laissé apparaître des systèmes fiables et rapides de vision et de contrôle en industrie. Cependant leur utilisation reste restreinte .

## II. LA THERMOGRAPHIE INFRAROUGE

La thermographie infrarouge utilisée en maintenance préventive a pour objectif d'éviter toute défaillance dans un système industriel et autre. [3] la définit selon la norme française AFNOR 09-400 comme étant une technique permettant d'obtenir au moyen d'un appareillage approprié l'image thermique d'une scène thermique dans un domaine spectral de l'infrarouge.

L'outil de base est la caméra thermique. Elle enregistre les différents rayonnements infrarouges émis par les corps et qui varient en fonction de leur température. La couleur produite par la caméra est obtenue en associant une couleur à l'intensité reçue, afin de faciliter la lecture directe de la température. Ainsi à chaque couleur de l'image correspond une température.

L'usage de la thermographie est très vaste. En industrie, par exemple Legrand et al. [7] l'utilisent pour un système de contrôle et de décision. Ils proposent un système autonome de contrôle par thermographie active infrarouge qui détecte et caractérise des occlusions de colle à l'intérieur de bouchons plastiques.

En reconnaissance faciale Yoshitomi et al [4] [5] proposent une approche d'identification faciale qui exploite les techniques de traitement de l'image thermique.

En santé et médecine la thermographie est surtout pratiquée pour surveiller la température du corps humain qui est un indicateur naturel de maladies. Dans ce contexte Bahiri et al. [8] passent en revue différentes approches sur l'analyse automatique de distribution de la température des régions d'intérêt et leur analyse statistique pour la détection d'anomalies.

En rejoignant leur idée dans le milieu industriel nous proposons une approche très simple de localisation des régions d'intérêt d'anomalies qui peuvent subvenir dans les installations électriques et sur d'autres équipements. Nous procédons par une classification rigide et floue de l'image thermique respectivement par les algorithmes k plus proches voisins et Fuzzy C Means.

### III. CLASSIFICATION DE L'IMAGE COULEUR PAR L'ALGORITHME DES K MOYENNES

En traitement d'images couleurs, la plupart des images vérifient localement des propriétés d'homogénéité, notamment en termes d'intensité lumineuse. L'image thermique correspond exactement à ce critère et les algorithmes de classification rigide permettent donc d'apporter une solution à la segmentation d'images couleurs. Cependant le choix d'une approche adéquate va dépendre des résultats et du contexte dans lequel elle est appliquée.

Le K-means est une approche basée segmentation et classification de groupes de pixels à des classes couleurs. C'est une méthode itérative non supervisée est un outil de classification classique permettant de répartir un ensemble de données en classes homogènes [6].

L'algorithme est le suivant :

- Choisir k points aléatoires (pixels) qui représentent la position moyenne des partitions  $m_1(1), \dots, m_k(1)$  initiales .
- Répéter jusqu'à convergence ou jusqu'à atteindre le nombre d'itération donné
- Assigner chaque observation x (pixel) à la partition S la plus proche selon une métrique

$$S_i^{(t)} = \{x_j: \| x_j - m_i^{(t)} \| \leq \| x_j - m_{i^*}^{(t)} \| \quad i^* = 1..k\}$$

- Mettre à jour la moyenne de chaque classe ainsi

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

La convergence est atteinte quand il n'y a plus de changement de classes.

L'inconvénient majeur de ces approches est la définition du paramètre K qui nous permet de donner exactement les différentes classes sur l'image. Pour cette raison nous allons procéder à un jeu d'essai à travers des valeurs aléatoires pour obtenir un résultat conforme visuellement.

### IV. TRAITEMENT DE L'IMAGE THERMIQUE

L'étude a tout d'abord porté sur la conversion des images du modèle RVB au modèle Lab.

L'étape qui va suivre consiste à classifier les couleurs de l'image thermique à travers l'algorithme des K plus proches voisins décrit précédemment. Le K-means représente chaque pixel par sa position dans le modèle de couleurs. Il faut spécifier le nombre de groupes ou de partitions K, le type de distance métrique (euclidienne par exemple) entre les classes et le nombre d'itérations. La fonction Kmeans définie dans la version Matlab 7.10 prend en charge cette partie. Le choix du paramètre K est empirique. Pour commencer on porte le K à 2 classes.

Après exécution du K means chaque classe est caractérisée par un niveau de gris moyen et chaque pixel de l'image appartient à une classe résultat. Avec 2 classes le pixel est soit dans la classe 1 soit dans la classe 2 qui sont labélisées et indexées.

Sur l'image thermique d'un disjoncteur (Fig. 1.) le nombre K engendre 2 classes objets et fond de l'image (voir figure 2). En analysant l'image nous remarquons que nous avons perdu une grande partie de l'information à rechercher d'où un choix non conforme du paramètre K.

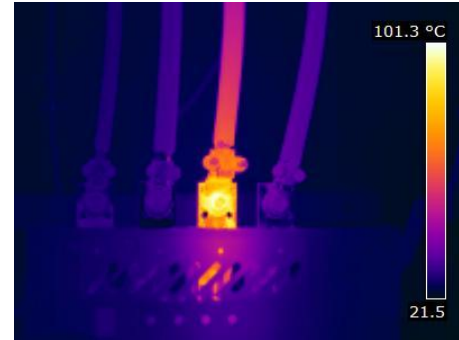


Fig. 1. Image thermique prise sur un disjoncteur

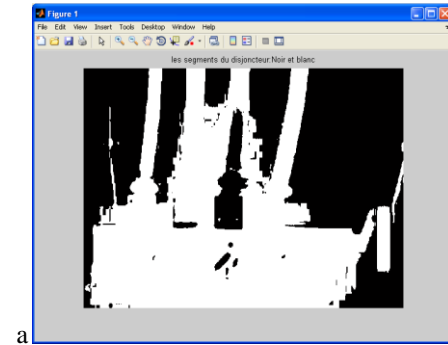


Fig. 2. Image classifiée du disjoncteur avec 2 classes

Pour remédier à ce problème nous allons augmenter le nombre de classes à 4. Remarquons que pour K=4 l'image est sous-segmentée (Fig. 3.). La zone la plus chaude est distinguée par le niveau de gris le plus clair. On remarque que le troisième câble (de gauche à droite) est caractérisé par un léger niveau de gris différent. Cependant le résultat reste imprécis.

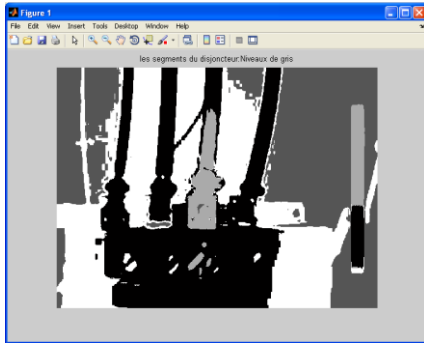
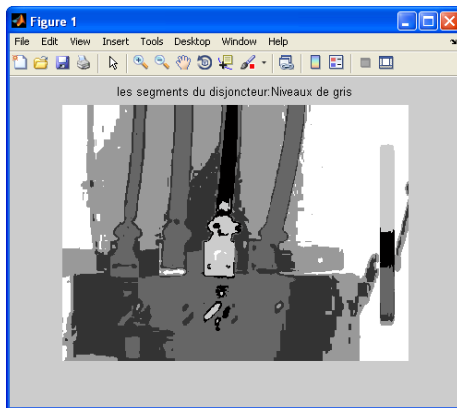
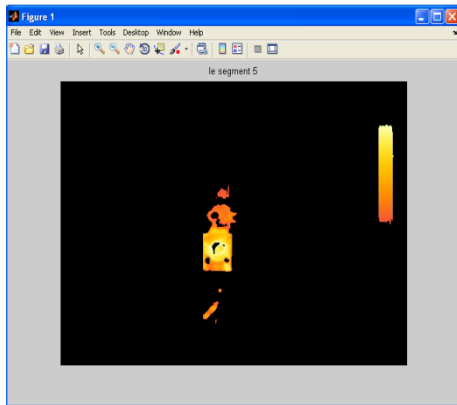


Fig. 3. Image classifiée du disjoncteur avec 4 classes

Pour un  $K=6$  nous arrivons enfin à un résultat visuellement convenable (Fig. 4.). Nous pouvons observer clairement que la partie désirée est bien déterminée par un niveau de gris spécifique (le plus clair).



(a)



(b)

Fig. 4. (a) Image classifiée du disjoncteur avec 6 classes en niveaux de gris (b) Segment isolé de la zone surchauffée

Nous isolons cette partie dans une image à part (fig 4.b), puis nous réalisons des retours arrière sur l'image originale.

C'est probablement une zone surchauffée de mauvais serrage (voir image référence Fig. 5.). Cette partie n'est pas

traitée dans ce papier, elle relève de la phase de reconnaissance (spécification de la panne).



Fig. 5. Image référence de l'image thermique en figure 1

## V. D'AUTRES RESULTATS SUR DIFFERENTS TYPES D'IMAGES

Nous avons appliqué l'algorithme K means sur d'autres types d'images thermiques de palier de moteur (Fig. 6.) et d'un transformateur (Fig. 8.a).

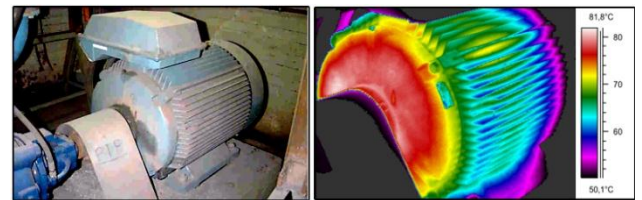
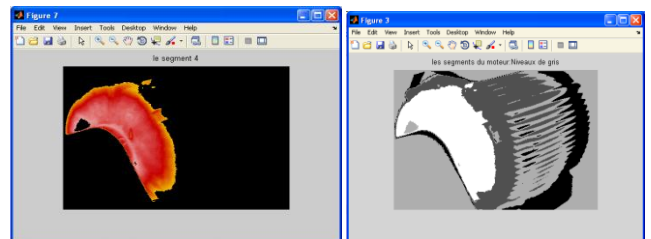


Fig. 6. Image d'un palier de moteur et son image thermique

L'image référence de la fig. 6. ne porte aucune indication à première vue. Cependant l'image thermique de ce moteur indique qu'il y a un échauffement (plus de  $80^{\circ}\text{C}$  au niveau de la carcasse).



(a)

(b)

Fig. 7. Image de la zone surchauffée (a) Image classifiée par le K-means

On observe clairement la zone surchauffée sur la fig 7.a qui est extraite après exécution de l'algorithme K-means appliqué sur l'image thermique correspondante. Le nombre de classes le plus adéquat est égal à 4

Voici un deuxième exemple sur une image thermique d'un transformateur (figure 8.b) où on constate une zone chaude sur l'image thermique (figure 8.a).

## VI. CONCLUSION

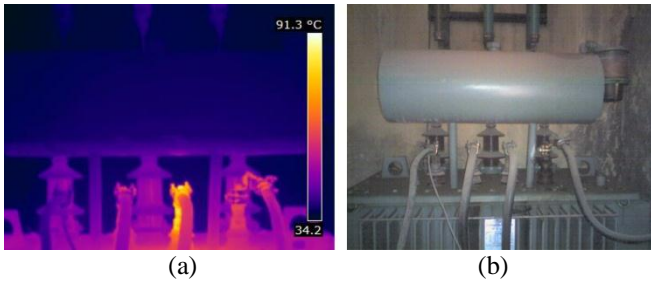


Fig. 8. (a) Image thermique d'un transformateur (b) Image originale du transformateur

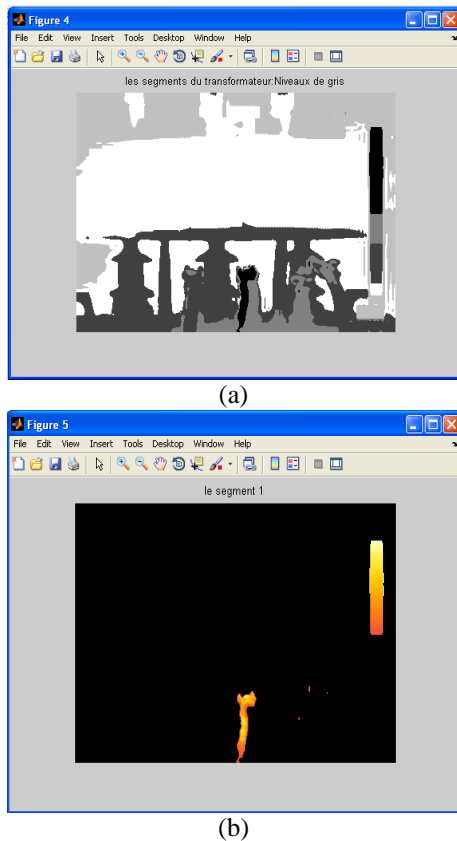


Fig. 9. (a) Image classifiée du disjoncteur avec 5 classes en niveaux de gris (b) Segment isolé de la zone surchauffée

Le nombre de classes adéquat obtenu est égal à 5 (figure 9.a). La zone surchauffée est isolée dans l'image de la figure 9.b.

L'application a été réalisée sur des images réelles d'installations électriques de la SONELGAZ-Oran (Algérie) de la SONATRACH-Oued Noumer et Hassi R'mel (Algérie) que nous remercions.

Le système de détection des parties en échauffement par thermographie infrarouge permet de localiser les zones de danger sur des images thermiques d'installations électriques et autres.

La technique consiste à appliquer un classifieur simple basé sur l'algorithme des K moyennes en vue de segmenter ce type d'images. Cette approche non supervisée, loin d'être idéal consiste à répartir un ensemble de données en k classes homogènes vérifiant localement cette propriété.

Loi d'être parfaits, nous jugeons que nos résultats sont très satisfaisants et encourageants. Ils démontrent l'intérêt des approches de segmentation et de classification dans l'automatisation de la tâche de détection des zones de danger. Cependant l'approche que nous avons utilisée présente un inconvénient important et reste semi automatique à cause du choix manuel du paramètre K. L'automatisation de cette tâche reste un enjeu futur majeur.

Aussi nous envisageons d'étendre ce travail pour localiser la cause principale du danger dans l'image thermique par l'application du flou

## REFERENCES

- [1] Braggins D., The European Vision Sourcebook 2000, UKIVA, 2000.
- [2] Robert Viseur, La vision industrielle selon VisionIndustrielle.Org, 2003
- [3] G. Gilbert, "La thermographie infrarouge: Applications", Ed. Masson, Paris, 1999.
- [4] Yoshitomi, Y. ; Fac. of Eng., Miyazaki Univ., Japan ; Miyaura, T. ; Tomita, S. ; Kimura, S. Face identification using thermal image processing, .. RO-MAN '97. Proceedings., 6th IEEE International Workshop on Robot and Human Communication 1997
- [5] Yuu Nakanishi, Yasunari Yoshitomi, Taro Asada, Masayoshi Tabuse, Facial expression recognition of a speaker using thermal image processing and reject criteria in feature vector space, Artificial Life and Robotics, Volume 19, Issue 1, pp 76-88, February 2014
- [6] J. B. MacQueen "Some Methods for classification and Analysis of Multivariate Observations", Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, 1:281-297, 1967
- [7] Anne-Claire LEGRAND, Fabrice MERIAUDEAU, Patrick GORRIA, "Système de contrôle et de décision par thermographie active infrarouge : application à la détection d'occlusion de colle à l'intérieur de bouchons plastiques" Colloque sur le Traitement du Signal et des Images, GRETSI 2001
- [8] B.B. Lahiri, S. Bagavathiappan, T. Jayakumar, John Philip, Medical applications of infrared thermography: A review, Infrared Physics & Technology, Volume 55, Issue 4, July 2012, Pages 221-235



# La Spectroscopie De Corrélation De Fluorescence Pour L'analyse Ultra-Sensible Des Systèmes Biologiques

<sup>1</sup>D.Issaad, <sup>1</sup>A.Medjahed, <sup>1,2</sup>S.Amara, <sup>1</sup>M.Bouafia,  
<sup>3</sup>N. Djaker, <sup>3</sup>M.Lamy De La Chapelle

<sup>1</sup>Laboratoire d'Optique Appliquée, Institut d'Optique et de Mécanique de Précision, Université Ferhat Abbas, Sétif-1, 19000,  
ALGERIE.

<sup>1,2</sup> Centre de Développement des Technologies Avancées, Unité de Recherche Optique et Photonique, Cité du 20 Août 1956, BP n°  
17 Baba Hassen (16303) Alger, ALGERIE.

<sup>3</sup>Laboratoire de Chimie, Structure et Propriétés des Biomatériaux et Agents Thérapeutiques,  
Université Paris 13, FRANCE.

E-mail : dihiaissaad@yahoo.fr

**Résumé** — La spectroscopie de corrélation de fluorescence la FCS, est une technique incontournable pour l'analyse dans le domaine des sciences du vivant grâce à sa forte efficacité optique. La fluorescence repose sur le principe d'absorption et d'émission de photons par les molécules. La technique FCS consiste à la mesure des fluctuations d'intensité du signal inélastique diffusée en fonction du temps, qui permettent d'accéder à la dynamique du système. Parmi les paramètres chimiques et physiques accessibles par FCS sont les concentrations locales, les coefficients de mobilité, les constantes de vitesse des processus d'association et de dissociation etc... . L'application de cette méthode qui est non invasive et simple, est utilisée par exemple pour démêler les processus dans des cellules vivantes en raison de la haute résolution spatiale. De plus, l'observation de la dynamique moléculaire locales permet d'accéder à des paramètres environnementaux tels que : la concentration locale, le pH, le temps de diffusion ou de la viscosité. Cette technique est particulièrement importante pour les chercheurs afin d'évaluer quantitativement les interactions des molécules ainsi que la dynamique de petites quantités moléculaires dans des systèmes d'intérêts biologiques et plus particulièrement en médecine.

**Mots Clés** — La microscopie confocale, la fluorescence, la spectroscopie de corrélation de fluorescence, biophotonique.

## I. INTRODUCTION

La révolution dans le domaine de la biomédicale nécessite le développement de nouveaux outils technologiques et en particulier des outils d'optique de pointe. En effet, la caractérisation et la manipulation d'objets quantiques peuvent être effectuées en utilisant des techniques expérimentales dans la gamme des fréquences optiques. Dans ce contexte, les techniques de sondage d'objets biologiques par méthodes optiques ont connu un développement considérable au cours de ses dernières années, conduisant à l'avènement d'un domaine de recherche à la frontière entre l'optique, la

photonique, la chimie et les sciences du vivant c'est la biophotonique. L'utilisation d'un rayonnement optique en biologie permet à l'heure actuelle l'étude non-invasive des propriétés du vivant à l'échelle nanométrique. En particulier, le suivi de réactions biologiques au niveau d'une molécule individuelle présente un intérêt majeur dans le domaine de l'analyse biomédicale car il pourrait permettre la détection de mutation ponctuelle de l'ADN. En dépit du développement de contrastes, le contraste de fluorescence demeure le plus répandu pour détecter des molécules individuelles grâce à sa haute sensibilité de détection et sa simplicité de mise en œuvre. Parmi les différentes méthodes d'analyse de molécules en solution existantes, la spectroscopie de corrélation de fluorescence (ou FCS pour Fluorescence Correlation Spectroscopy), introduite en 1972 par Madge, Elson et Webb, est sans doute la technique la plus efficace pour réaliser des études moléculaires dans des milieux peu concentrés avec une grande résolution spatio-temporelle.

La quantité d'informations accessible par la FCS donne accès à une large gamme de processus, dont les temps caractéristiques s'étendent de quelques nanosecondes jusqu'à la seconde. En particulier, les phénomènes de diffusion, les concentrations locales, la brillance des molécules, les vitesses d'association moléculaires et le temps de vie peuvent être exploré par la spectroscopie de corrélation de fluorescence.

## II. PRINCIPE DE LA SPECTROSCOPIE DE CORRELATION DE FLUORESCENCE (FCS)

En FCS, le paramètre important n'est pas l'intensité diffusée elle-même, mais les fluctuations relatives à l'émission de fluorescence  $\langle \delta F^2 \rangle$  essentiellement causées par les fluctuations de particules dans le volume d'observation.

L'échelle de temps caractéristique des fluctuations est le temps mis pour qu'une molécule initialement dans le volume de détection en sorte, c'est-à-dire

$$\tau = \frac{r_0^2}{4D}$$

$r_0$  Le petit axe de l'ellipsoïde de révolution qui décrit le volume (la figure 1) et  $D$  décrit le coefficient de diffusion. La caractérisation des fluctuations temporelles d'intensité de fluorescence se fait à l'aide de la fonction d'autocorrélation suivante :

$$G_{FCS} = \frac{1}{\langle N \rangle} \cdot \frac{1}{\left(1 + \frac{t}{\tau}\right) \left(1 + \frac{\omega^2 t}{\tau}\right)^{1/2}}$$

Avec :

$\omega^2 = \frac{r_0^2}{z_0^2}$  et  $\langle N \rangle$  est le nombre de particules dans le volume.

La spectroscopie de corrélation de fluorescence mesure les temps de diffusion des particules fluorescentes en solution, on peut ainsi accéder à la stœchiométrie des complexes, à la dynamique des interactions entre biomolécules et ceci même au sein d'une cellule par imagerie.

Par contre, on sait que l'intensité du signal mesuré étant proportionnelle au nombre de particules  $N$ . Si  $N$  est grand on sera peu sensible aux fluctuations de particules mais si on réduit le nombre de particules, on sera sensible à leur fluctuation de nombre. Pour cela, il faut réduire le volume d'observation. Habituellement, on réduit le volume d'observation à l'aide d'un microscope confocal, (la figure 1) est un exemple de cône de distribution d'intensité obtenu avec un tel microscope.

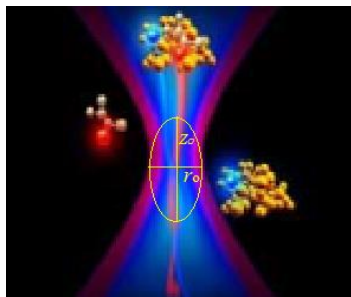


Fig. 1 Cône de distribution d'intensité d'une tache focale

### III. MONTAGE DE LA SPECTROSCOPIE DE CORRELATION DE FLUORESCENCE

Le montage expérimental est basé sur le principe d'un microscope inversé comme illustré en (figure 2). Lors des mesures FCS, l'excitation s'effectue avec un laser continu à une longueur d'onde de 633 nm. La puissance du faisceau excitateur est contrôlée par une lame demi-onde ( $\lambda/2$ ) accolée à un cube polariseur. Le faisceau est élargi à l'aide d'un dispositif afocal. Un miroir dichroïque conçu pour travailler sous une incidence de 45° degrés, réfléchit le faisceau excitateur et transmet uniquement le signal de fluorescence. Après réflexion sur le miroir dichroïque, le faisceau incident est focalisé sur l'échantillon à l'aide d'un objectif à eau de forte ouverture numérique NA=1.2 de grossissement x60 muni d'une table piézoélectrique qui va également collecter le signal de fluorescence provenant de l'échantillon. Pour les mesures FCS à 633 nm, les sondes fluorescentes utilisées sont des molécules d'Alexa Fluor en solution aqueuse. Après excitation, la fluorescence des molécules traverse la lentille de tube du microscope ensuite, filtrée par un trou confocal de 50µm placé dans le plan image de l'objectif de microscope. En sortie du trou confocal, le signal de fluorescence est divisé par un cube séparateur 50/50 et focalisé sur deux photodiodes à avalanche placées derrière les filtres. Les fluctuations du signal de fluorescence sont alors analysées par corrélation croisée des signaux de chacune des photodiodes à l'aide d'un corrélateur. L'analyse des données FCS permet de déterminer le nombre de molécules et le temps de diffusion.

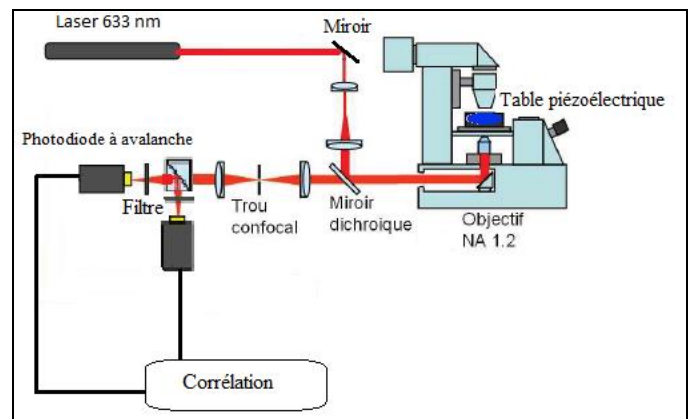


Fig. 2 Dispositif expérimental de la FCS

#### IV. RESULTATS ET DISCUSSION

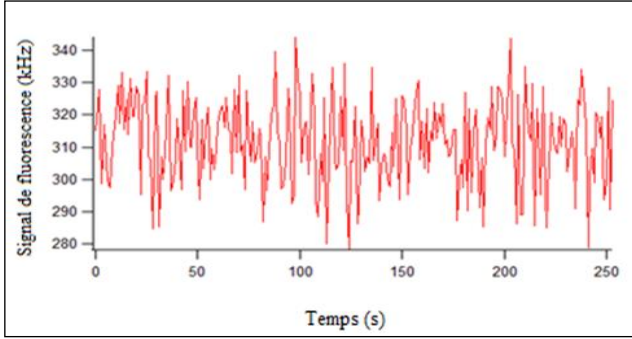


Fig. 3 Traces temporelles du signal de fluorescence correspondant à des molécules d'Alexa 647

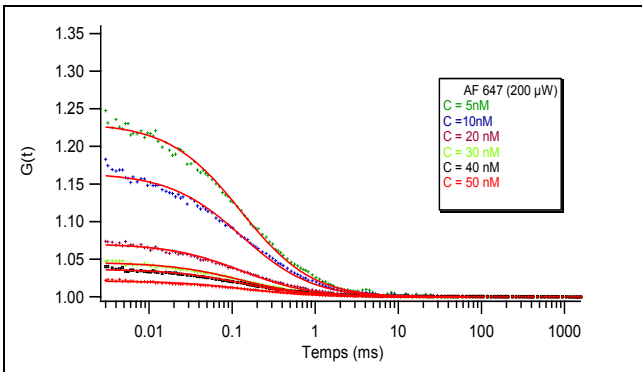


Fig. 4 Courbes de corrélation d'Alexa 647 pour différentes concentrations

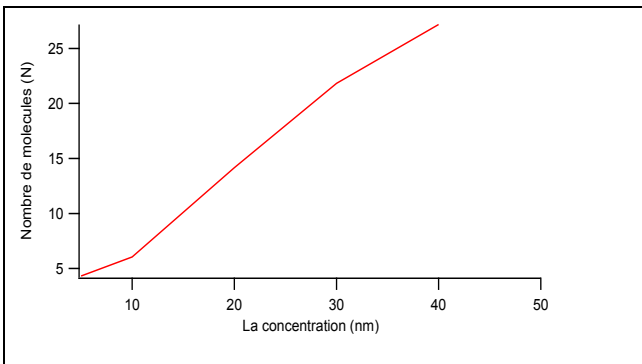


Fig. 5 Evolution du nombre de particules détectées pour différentes concentrations d'Alexa 647

Table I

Résultats Des Mesures De La FCS Pour l'Alexa 647

Concentration Théorique (nM)	5	10	20	30	40	50
N	4.32	6.07	14.16	21.83	27.17	47.21
$\tau_D$	0.12	0.13	0.13	0.13	0.13	0.14

La courbe FCS par elle-même ne représente qu'un spectre sur les phénomènes, quand la particule traverse l'espace confocale. Quand un modèle approprié est connu, FCS peut être utilisée pour obtenir une information quantitative telle que : la dynamique singulet-triplet des particules, les coefficients de diffusion et les concentrations moyennes.

L'intensité de fluorescence est fluctuante en raison du mouvement de diffusion des particules donc, les paramètres physiques extraits sont essentiellement liés au montant des fluctuations. En utilisant l'autocorrélation temporelle, les particules dans le sous-espace défini par le système optique sont en train de changer, l'analyse donne le nombre moyen de particules fluorescentes et le temps de diffusion moyen. Quand trop d'entités de particules (grandes concentrations) sont mesurées, les fluctuations globales sont faibles en comparaison au signal total donc les fluctuations ne peuvent pas être résolues cependant, dans l'autre sens (faibles concentrations), peu d'entités de particules sont mesurées, les fluctuations individuelles sont trop peu dans le temps alors la résolution est meilleur pour la FCS.

#### V. CONCLUSION

Dans cette étude, nous avons appliqué la spectroscopie de corrélation de fluorescence pour mesurer la concentration et le coefficient de diffusion des molécules fluorescentes. Un intérêt primaire pour la spectroscopie de corrélation de fluorescence est sa convenance aux applications aux solutions des concentrations nano-molaires des sondes fluorescentes. La technique est particulièrement attrayante dans les situations d'exigence de l'utilisation de petites quantités de matériaux coûteux ou rares.

Cette technique est caractérisée par sa haute sensibilité, une bonne sélectivité, la simplicité, la courte durée de l'analyse possède de grandes applications potentielles dans le diagnostic clinique, la nourriture, les analyses biologiques et biomédicales.

Cette technique peut parvenir au Diagnostic Ultra-sensible c'est à dire la détection et l'identification des montants clairsemés de molécules spécifiques de la maladie sur les cellules à l'intérieur ou dans les fluides corporels, ce qui va ne mener à un dépistage précoce (petits volumes d'échantillons, faibles concentrations et une rapidité de lecture) pour les thérapies médicamenteuses avancées.

## REFERENCES

- [1] Oleg Krichevsky, Gregoire Bonnet "Fluorescence correlation spectroscopy: the technique and its applications", *Rep. Prog. Phys* 65, 251–297, 2002.
- [2] Kirsten Bacia et Petra Schwille, "A dynamic view of cellular processes by in vivo fluorescence auto- and cross-correlation spectroscopy", *Methods* 29 74–85, 2003.
- [3] Laure WAWREZINIECK, "Confinement moléculaire et organisation de la membrane des cellules vivantes : analyse de la diffusion par spectroscopie de corrélation de fluorescence", THÈSE de Doctorat en Biologie des Eucaryotes, Université Aix-Marseille II, Septembre 2006.
- [4] David M. Jameson & Justin A. Ross & Joseph P. Albanesi., "Fluorescence fluctuation spectroscopy: ushering in a new age of enlightenment for cellular dynamics", *Biophys. Rev* DOI 10.1007/s12551-009-0013-8, 2009.
- [5] Lichuan Tang, Chaoqing Dong, Jicun Ren, "Highly sensitive homogenous immunoassay of cancer biomarker using silver nanoparticles enhanced fluorescence correlation spectroscopy", *Talanta* 81, 1560–1567, 2010.
- [6] Nibedita Pal, Sachin Dev Verma, Moirangthem Kiran Singh et Sobhan Sen, "Fluorescence Correlation Spectroscopy: An Efficient Tool for Measuring Size, Size-Distribution and Polydispersity of Microemulsion Droplets in Solution", *Anal. Chem*, 83, 7736–7744, 2011.

# Extraction des joints de grains par une approche multi-échelles

Mme. SLIMANI Kahina<sup>1</sup>  
[rebahkahina@yahoo.fr](mailto:rebahkahina@yahoo.fr)

Mr. AMEUR Soltane<sup>1</sup>  
[Ameursoltane@yahoo.com](mailto:Ameursoltane@yahoo.com)

Mme. AMEUR Zohra<sup>1</sup>  
[amzohra@yahoo.fr](mailto:amzohra@yahoo.fr)

<sup>1</sup>Laboratoire d'analyse et de modélisation des phénomènes aléatoire  
Faculté de Génie Électrique et Informatique (U.M.M.T.O.)

**Résumé.** Une méthode de segmentation et d'analyse d'une image représentant la structure interne d'un matériau est développée dans ce papier. Elle est utilisée pour extraire la totalité des joints de grains, et notamment les microstructures les plus fines, ce qui rend possible la caractérisation complète de la structure. L'objectif visé est d'utiliser plusieurs échelles pour affiner le résultat. Le détecteur multi échelles de Canny consiste à utiliser une transformation en ondelette, qui est obtenue par projection orthogonal de l'image sur les espaces affines de la base d'ondelette à plusieurs échelles.

**Mots-clés :** segmentation, détecteur de Canny, multi-échelles, ondelette.

## I. INTRODUCTION

Dans le domaine de la métallographie, l'analyse d'images est appliquée pour la caractérisation géométrique de la microstructure des matériaux. En effet, la connaissance de la structure interne de ces matériaux permet d'établir des relations entre les paramètres morphologiques et les propriétés physiques (résistance, dureté, etc..) [1] [2] [3].

Le but de cette présentation est la segmentation de la micro structure d'un acier par une analyse multi-échelles qui consiste à décomposer l'image sur une gamme très étendue d'échelles, opération que nous pouvons comparer à une cartographie. A chaque échelle, l'image est remplacée par l'approximation la plus adéquate que l'on puisse y tracer. En allant des échelles les plus grossières vers les plus fines, nous accédons à des représentations de plus en plus précises de l'image donnée.

L'analyse s'effectue en calculant ce qui diffère d'une échelle à l'autre, c'est à dire les détails à une résolution donnée.

## II. DESCRIPTION D'UNE MICROSTRUCTURE

Le type de structure qui fait l'objet de cette étude est l'acier qui trouve son application dans de nombreux domaines.

Les frontières des grains qui sont appelés joints de grains [4] [5] sont des zones de discontinuité dans l'orientation de la structure cristalline. Dans l'étude que nous menons, il est important que les grains soient vus bien individualisés à partir de segmentation en contour.

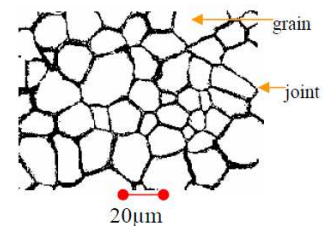


Figure.1. exemple d'une micro structure (grossissement 500)

## III. ANALYSE MULTI RESOLUTION

L'algorithme d'analyse multi résolution a été proposé par S. Mallat. Cet algorithme a pour but d'extraire les caractéristiques d'un signal en l'analysant à diverses échelles. La décomposition est réalisée par un banc de filtres (passe bas et passe haut) suivi d'une décimation selon les lignes puis les colonnes [6].

L'inconvénient de l'algorithme de Mallat est cette décimation, nous ne pouvons pas faire une correspondance pixel à pixel entre les différentes résolutions.

### A. Algorithme à trous

Dans ce type de transformation la taille des sous bandes de détail reste invariante quelque soit l'échelle. L'utilisation de l'algorithme à trous crée une redondance d'information qui permet de maintenir l'invariance en translation à tous les niveaux de décomposition. De plus, il permet une bonne localisation spatiale des fréquences plus basses [7] [8] [9]. Mais son inconvénient est le sur échantionnage, les maxima locaux se diffuse à travers les échelles.

Pour pallier les inconvénients des algorithmes cités précédemment, nous allons adopter l'algorithme par convolution en cascade B-spline qui analyse l'image sur des échelles assez fines et qui sont proches les unes des autres.

### B. Algorithme par convolution en cascade B-spline

L'algorithme par convolution en cascade B-spline [10] peut être interprété comme un banc de filtrage dont le filtre passe-haut est un filtre de dérivation  $D_x$  et  $D_y$ , et le filtre passe-bas est le masque B-spline  $\theta$ .

A chaque étape, la composante basse-fréquence est convoluée d'une part par le filtre passe-bas  $\theta$  :

$$f * \overbrace{\theta * \theta * \dots * \theta}^{n \text{ convolutions}} \quad (1)$$

D'autre part par un filtre passe-haut de dérivation horizontale  $Dx$  :

$$w^1 f(x, y, n) = f * \overbrace{\theta * \theta * \dots * \theta}^{n \text{ convolutions}} * Dx \quad (2)$$

Et enfin par un filtre de dérivation verticale  $Dy$  :

$$W^2 f(x, y, n) = f * \overbrace{\theta * \theta * \dots * \theta}^{n \text{ convolutions}} * Dy \quad (3)$$

Pour effectuer la décomposition en gradient multi-échelle, Mallat et Zhong [11] on construit une ondelette appartenant à la famille des Splines. Cette ondelette est associée à une décomposition discrète qui approxime la dérivée première d'une gaussienne, ce qui permet d'avoir certaines propriétés importantes :

- Compacité : puisqu'on effectue des convolutions, le nombre de coefficients  $h$  est forcément fini, et la fonction d'échelle est compacte.
- Régularité : c'est une condition essentielle pour éviter les artefacts liés à l'irrégularité de la fonction d'échelle.

Le masque utilisé est:

$$\theta = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (4)$$

#### IV. LE DETECTEUR MULTI-ECHELLES DE CANNY

Pour travailler sur des fonctions bidimensionnelles, et en particulier des images, le détecteur de Canny, consiste à utiliser non plus une, mais deux transformées en ondelettes  $\psi^1$  et  $\psi^2$  orientées (les dérivées suivant  $x$  et  $y$  d'un noyau de lissage).

Soit  $\psi$  une ondelette bidimensionnelle dont les deux composantes sont les dérivées selon  $x$  et selon  $y$  d'une fonction B-spline  $\theta(x, y)$

$$\psi^1 = \frac{\partial \theta}{\partial x} \quad (5)$$

$$\psi^2 = \frac{\partial \theta}{\partial y} \quad (6)$$

Avec  $\vec{\psi} = (\psi^1, \psi^2)$

Avec  $\theta$  un noyau de lissage que est positif et isotrope, à variables séparées ; il vérifiera:

$$\iint_{R^2} \theta = 1 \quad (7)$$

$\theta(x, y)$  est isotrope où  $\theta(x, y) = \theta_1(x) \cdot \theta_2(y)$

On définit ainsi deux transformées en ondelettes, l'une détecte les singularités verticales et l'autre les singularités horizontales [12], dont les valeurs à l'échelle  $a \geq 0$  et au point  $(u, v)$  de  $R^2$  sont:

$$W^1 f(x, y, a) = \iint_{R^2} \frac{1}{a} f(x, y) \psi^1 \left( \frac{x-u}{a}, \frac{y-v}{a} \right) dx dy = f * \check{\psi}^1(x, y) \quad (8)$$

$$W^2 f(x, y, a) = \iint_{R^2} \frac{1}{a} f(x, y) \psi^2 \left( \frac{x-u}{a}, \frac{y-v}{a} \right) dx dy = f * \check{\psi}^2(x, y) \quad (9)$$

Avec :

$$\psi^p(x, y) = \frac{1}{a} \psi^p \left( \frac{x}{a}, \frac{y}{a} \right) \text{ et}$$

$$\check{\psi}^p(x, y) = \psi^p(-x, -y) \text{ pour } p = 1, 2.$$

Plutôt que de représenter  $\vec{w}f(x, y, a)$  par ses deux composantes  $W^1 f(x, y, a)$  et  $W^2 f(x, y, a)$  il sera donc plus pratique, d'utiliser son module et son orientation [13]

$$Mf(x, y, a) = \sqrt{Mf|W^1 f(x, y, a)|^2 + Mf|W^2 f(x, y, a)|^2} \quad (10)$$

Et

$$Af(x, y, a) = \tan^{-1} \left( \frac{W^2 f(x, y, a)}{W^1 f(x, y, a)} \right) \text{ si } W^1 f(x, y, a) \geq 0 \quad (11)$$

$$Af(x, y, a) = \pi - \tan^{-1} \left( \frac{W^2 f(x, y, a)}{W^1 f(x, y, a)} \right) \text{ si } W^1 f(x, y, a) < 0 \quad (12)$$

$Mf(x, y, a)$  est le module en  $(x, y)$  du gradient de l'image  $f$  convoluée par un noyau de lissage à l'échelle  $a$ .

$Af(x, y, a)$  est l'orientation en  $(x, y)$  du gradient de l'image lissée par ce même noyau à la même échelle.

Pour décider si  $Mf(x, y, a)$  est un maximum local, on le compare au module en ses deux voisins dans les directions  $Af(x, y, a)$  et  $-Af(x, y, a)$  : appelés les  $Mf(x_1, y_1, a)$  et  $Mf(x_2, y_2, a)$ .

Un maximum local de  $Mf(x, y, a)$  sera en fait un maximum local dans la direction  $Af(x, y, a)$  et sera défini en pratique comme suit :

Soit  $\varepsilon > 0$  un seuil fixé.  $Mf(x, y, a)$  est un maximum local si :

$$Mf(x, y, a) - Mf(x_1, y_1, a) > \varepsilon \text{ et } Mf(x, y, a) \geq Mf(x_2, y_2, a) \quad (13)$$

Ou

$$Mf(x, y, a) - Mf(x_2, y_2, a) > \varepsilon \text{ et } Mf(x, y, a) \geq Mf(x_1, y_1, a) \quad (14)$$

$\varepsilon > 0$  sert à ce que de petites oscillations ne génèrent pas de maximum local. Ces maxima locaux sont détectés à chaque échelle : soit  $\{a\}$  une échelle fixée. Si  $Mf(x, y, a)$  n'est pas un maximum local, on le met à 0 ainsi que  $Af(x, y, a)$ , Sinon il conserve sa valeur.

Une fois les modules maximaux sont calculés à toutes les échelles, il faut les chaînés à travers ces échelles [14]. Nous faisons une comparaison carte à carte afin de voir si, pour un pixel  $(x, y, a)$  correspondant à un module max, il existe, dans l'échelle inférieure, un pixel  $(x, y, a-1)$  et ses 8 voisins qui pourraient être lié à ce dernier. Pour cela, nous sélectionnons alors parmi ces 9 pixels celui dont la direction du gradient est proche de  $Af(x, y, a)$ . Le point sélectionné caractérise un contour. Chacune de ces comparaisons faites deux à deux, jusqu'à l'échelle la plus fine.

## V. APPLICATION DU DETECTEUR MULTI-ÉCHELLES DE CANNY

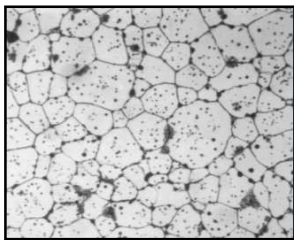


Figure.2. image initiale.

### A. Cartes des modules du gradient

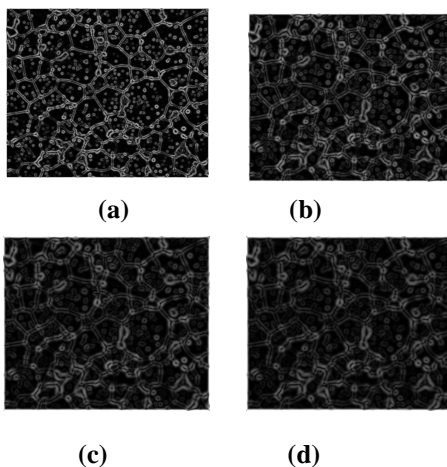


Figure.3. module du gradient à plusieurs échelles.

Modules du gradient : (a) à l'échelle 1, (b) à l'échelle 10, (c) à l'échelle 15, (d) à l'échelle 20.

### D. Maximas locaux du gradient

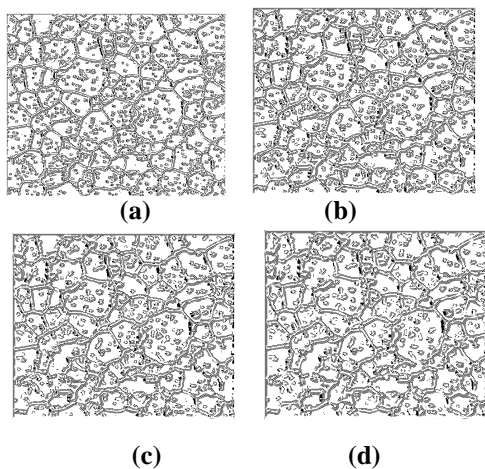


Figure.4. maximas locaux à plusieurs échelles.

Maximas locaux : (a) à l'échelle 1, (b) à l'échelle 10, (c) à l'échelle 15, (d) à l'échelle 20.

### B. Résultat obtenu après chaînage des maximas locaux du gradient

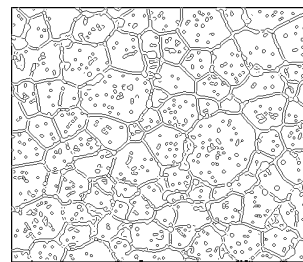


Figure.5. Résultat après chaînage

## VI. INTERPRETATION DES RESULTATS

A partir des singularités horizontales et verticales, nous obtenons les cartes des modules du gradient (figures « 3 »). L'analyse de ces dernières montre qu'aux petites échelles nous arrivons à détecter les détails les plus fins, et en allant vers les échelles grossières, seuls les détails importants sont présents.

Les cartes des maximas-locaux (figures « 4 ») sont obtenues en combinant les cartes des orientations et celles des modules du gradient. Nous constatons alors que sur les échelles fines des maximas locaux nous apercevons tous les détails, même les non significatifs et cela dû à la détection des pics du bruit. Sur les échelles intermédiaires nous apercevons que seuls les contours significatifs et les frontières entre les structures de l'image sont détectées. A partir des échelles grossières les détails disparaissent. Après chaînage (figures « 5 ») nous retrouvons les contours les plus communs entre les échelles, par conséquent les contours les plus significatifs.

## VII. CONCLUSION

L'application de cette méthode multi échelles à des images de microstructure a conduit à des résultats très significatifs. En effet, la segmentation utilisée a conduit à une structure dans laquelle les grains sont correctement séparés. Ces résultats montrent bien que la méthode utilisée est adaptée à ce type de traitement par rapport aux méthodes qui utilisent une seule échelle.

## VIII. REFERENCES

- [1] J.L. Chermant, Caractérisation des poudres et des céramiques, éditions Hermès, Paris, 1992.
- [2] R.T. Dehoff and F. N. Rhines, Quantitative microscopy, McGraw-Hill, New York, 1972.
- [3] Michael F. Ashby and David R.H. Jones, *Matériaux*, éditions Dunod, France, 1991.
- [4] S. Journaux, P. Gouton, M. Paindavoine, G. Thauvin, 2001 : « Extraction des joints de grains par ondelettes directionnelles et morphologie mathématique : application

à l'évaluation de l'endommagement par fluage ». Revue de Métallurgie-CIT/Science et Génie des Matériaux. pp.485-498.

[5] S.Amer medjkane et S.Ameur ,2005 : «Traitement et Analyse Morphologique de la Structure Interne des Matériaux par la Théorie de la Morphologie Mathématique », SETIT 2005, 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications March 27-31, 2005, Tunisia.

[5] Victor Racine, 2003 : « Traitement des données générées par vidéo-microscopie 4D (3D+t) pour l'analyse de l'implication des protéines Rab6A et Rab6A dans le transport vésiculaire ». Thèse de DEA, Institut Curie. France.

[6] Hela Mahersia Sahbani, Kamel Hamrouni, Noureddine Ellouze, 2007 : « Identification de texture par fusion d'attributs multi-résolutionnels et invariants à la rotation ». Ecole Nationale d'Ingénieurs de Tunis (ENIT), 4th International Conférence: Sciences of Electronic, Technologies of Information and Télécommunications, Tunisie.

[7] Bijaoui, Jean-Luc Starck, Fionn Murtagh, 1994 : « Restauration des images multi-échelles par l'algorithme à trous », Traitement du Signal. Vol 11, N° 3. pp.230-243.

[8] Le Cadet, O. (2004) : Méthode d'ondelettes pour la segmentation d'image. Application à l'imagerie médicale et au tatouage d'images. PhD thesis, specialty Mathematics. National Polytechnic Institut. Grenoble. France.

[9] Stephane Mallat, and Sifen Zhong, 1992: "Characterization of Simals from Multiscale Edges". IEEE, Transactions on Patiern analysis and machine intelligence, vol.14, N° 7, pp.674-693.

[10] Grégoire Mercier, Stéphane Derrode, Wojciech Pieczynski, 2004 : « Segmentation multi échelle de nappes d'hydrocarbure ». Traitement du signal 2004, vol 21, N° 4. pp.329-344.

[11] Jie Hou, Jin-Hua Ye, Sha-Sha Li, 2007: "Application of Canny Combining and Wavelet Transform in The Bound of Step-Structure Edge Detection". IEEE. Proceedings of the 2007, International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China.

[12] Guo Fude Yang Yahui Ning Tao Chen Bin Guo Liejin, 2008: "Edge Detection Based on Wavelet Analysis with Gaussian Filter". IEEE, Congress on Image and Signal Processing. pp.724-728.

[13] Enguerran Grandchamp, 1998 : « Outils d'aide à la segmentation d'image : classification de ruptures par décomposition en ondelettes », INP-ENSEEIH.



# L'efficacité de la transformée de Radon vis-à-vis du bruit de quantification : Etude comparative

Cherifi Mehdi<sup>#</sup>, Lahdir Mourad<sup>#</sup>, Ameer Soltane<sup>#</sup>

<sup>#</sup>Laboratoire d'Analyse et de Modélisation des Phénomènes Aléatoires LAMPA,

Département d'Electronique, Faculté de Génie Electrique FEGI, Université Mouloud Mammeri

Tizi-Ouzou, Algérie

<sup>1</sup>cherifi.m85@gmail.com

<sup>2</sup>mlahdir@yahoo.fr

<sup>3</sup>ameursoltane@yahoo.com

**Résumé**—La compression d'images a pour but de réduire le nombre de bits nécessaire pour représenter une image. La transformée de Radon est devenue un outil très intéressant dans le domaine de traitement d'images. Sa robustesse envers des bruits comme le bruit blanc a poussé les chercheurs à réaliser des méthodes de détection des objets dans des images bruitées. La transformée en cosinus discrète a montré son efficacité dans la compaction d'énergie de l'image à compresser dans un nombre de coefficients réduit. Elle fait partie de nombreuses normes internationales y compris JPEG et MPEG. Nous présentons dans cet article une comparaison entre la méthode basée sur la transformée de Radon et celle basée sur la transformée en cosinus discrète (DCT). Cette comparaison est réalisée pour un niveau de quantification élevé et égal à 20 dans le cas de la DCT. Des critères d'évaluation comme le PSNR et le taux de compression ont été utilisés pour quantifier la différence entre les deux méthodes.

**Mots clés**- Compression d'images; Transformée en cosinus discrète (DCT); Transformée de Radon; Processus d'intégration; Bruit de quantification.

## I. INTRODUCTION

La compression d'images consiste à réduire le volume de données nécessaires à la représentation d'une image. L'exploitation de la redondance spatiale et spectrale et les faiblesses du système psychovisuel ont permis de développer des méthodes qui conduisent à la réduction de la quantité de données nécessaires pour la représentation de l'image. Ce qui conduit à éliminer tout problème de transmission ou d'archivage. Nous distinguons deux grandes classes de compression : la compression sans pertes et avec pertes.

La compression sans pertes permet de reconstruire après décompression, une image identique à l'originale mais avec un faible taux de compression. Parmi les codeurs les plus utilisés, nous citons le codage Huffman [1], le codage arithmétique, le codage Golomb-Rice [2], [3], le codage Tunstall [4] et le codage RLC (Run Length Coding) [5]. Nous distinguons aussi le codage prédictif dans le cas d'utilisation des prédicteurs simples sans introduction d'un filtrage comme le codage LPC (Linear Predictif Coding) [6].

Ces codeurs se distinguent de la compression avec pertes qui introduit une dégradation irréversible sur l'image originale mais qui autorise une compression bien plus importante que

celle obtenue par des méthodes sans pertes. Ces méthodes sont basées généralement sur la quantification des blocs de coefficients issus d'une transformée. Parmi ces transformées nous distinguons : La transformée de Karhunen-Loève (KLT), la DCT [7], [8], la Transformée Sinus Discrète (DST), la transformée de Radon [7], [9], [10], [11], [12], la transformée en fractal [13], la Transformée de Walsh Hadamard discrète (DWHT), La transformée en ondelettes discrète (DWT) [13], [14], [15], [16],, la transformée Mojette [10], [17]. Nous distinguons aussi le codage prédictif avec pertes comme le DPCM (Differential Pulse Code Modulation) [18].

Nous montrons dans cet article la puissance de la transformée de Radon par le moyen de la méthode de compression d'images basée sur la transformée de Radon et la DCT développée par S. A. Pradeep and R. Manavalan [7].

La transformée de Radon est devenue un outil très important qui a séduit beaucoup de chercheurs dans le domaine de la compression d'images vu sa robustesse aux différents bruits comme le bruit blanc et le bruit de quantification par rapport aux autres types de transformations comme la transformée de Fourier et les ondelettes [19]. Elle est notamment utilisée dans la détection des objets dans des images bruitées comme dans le cas de sillages de mobile marins dans les images radar [20]. La DCT a été créée par N. Ahmed, T. Natarajan et K. Rao en 1974 [8]. La capacité de compression de la DCT est très proche de celle de la KLT. Elle est aussi presque égale à la KLT dans sa capacité de compaction d'énergie. Contrairement à la KLT, la DCT est indépendante de l'image en question, c'est-à-dire, le noyau de sa matrice est fixé à une taille donnée. Donc, aucune information au côté du récepteur à propos de la taille des blocs n'est nécessaire pour la reconstruction de l'image originale.

Le reste de cet article est organisé comme suit : La section 2 est consacrée aux approches de compression d'images utilisées dans la comparaison. La section 3 donne les différents résultats des critères d'évaluation utilisés dans la comparaison ainsi que leurs discussions. Nous terminons avec la section 4 qui résume l'ensemble de l'article avec une conclusion.

## II. LES APPROCHES DE COMPRESSION D'IMAGES

Nous allons présenter dans cette section les différentes méthodes que nous avons utilisées dans notre comparaison.

### A. La compression d'images basée sur la transformée de Radon.

Cette méthode a été réalisée par S. A. Pradeep et R. Manavalan [7]. Cette approche exploite la transformée de Radon pour compresser les images. Cette méthode commence par l'application d'une transformée de Radon de l'image originale afin d'obtenir des points Radon. L'étape suivante consiste à coder les points Radon en utilisant une quantification des coefficients issues de la transformée en cosinus. Ces étapes dans l'ordre inverse permettent de reconstruire l'image originale. Dans leurs travaux, la méthode a été appliquée à des images tests et les résultats expérimentaux ont montré que cette méthode réalise un taux de compression élevé avec un PSNR (Peak Signal to Noise Ratio) supérieur à celui obtenu par d'autres méthodes comme DWT bi-orthogonale. Dans notre travail, nous utilisons le même schéma de la méthode, mais avec un niveau de quantification élevé pour montrer la puissance et la stabilité de la transformée de Radon vis-à-vis du bruit de quantification par rapport à la méthode de base de la compression avec DCT.

La transformée de Radon permet de représenter une image dans le domaine de Radon comme une collection de projections tout le long des différentes directions pour chaque angle donné [9]. En d'autres termes, la transformée de Radon mappe une ligne en un point dans le domaine de Radon ; une ligne particulière en un point particulier. L'une des propriétés les plus importantes de la transformée de Radon est qu'elle est inversible d'où la possibilité de la reconstruction de la fonction projetée à partir de la connaissance de ses intégrations le long d'hyperplans de son espace. La transformée de Radon est appliqué à une image pour une plage de  $\theta \in [0, 2\pi]$  pour recueillir la totalité de ses pour toutes les directions spécifiques dans le domaine de l'axe x et y. Toutes les projections de l'image I est donnée par l'équation 1.

$$\hat{I}_\theta(\rho) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(\vec{r}) \delta(\rho - x \cos \theta - y \sin \theta) d\vec{r} \quad (1)$$

Avec:

I: Image d'entrée.

$\vec{r}$ : Vecteur de position de composantes  $(x_1, x_2, \dots, x_n)$ .

$\rho = \vec{\rho} \cdot \vec{r}$ : Un hyperplan, avec  $\vec{\rho}$  un vecteur unitaire.

$\delta$ : La fonction de Dirac

Le passage entre le domaine de l'image et le domaine de Radon est assuré par la matrice du passage suivante :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}}_{\text{Matrice du passage}} \begin{pmatrix} \rho \\ s \end{pmatrix}$$

La représentation graphique de la transformée de Radon donne un sinogramme qui représente la superposition de toutes les sinusoïdes. La Fig. 1 donne un exemple de la représentation de la transformée de Radon de l'image Lena.

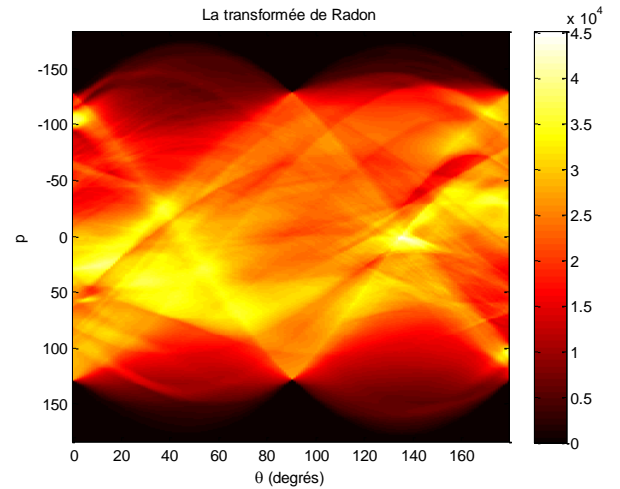


Fig. 1 Le sinogramme de l'image test Lena.

Nous allons expliquer dans ce qui vient le principe de la DCT.

### B. La compression d'images basée sur la DCT

La Transformée Cosinus Discrète (DCT) tire son nom du fait que les lignes de la matrice de la transformée de taille  $N \times N$  sont obtenues en fonction du cosinus comme le montre l'équation 2 [21] :

$$C(i, j) = \begin{cases} \sqrt{\frac{1}{N}} & i=0, j=0, 1, \dots, N-1 \\ \sqrt{\frac{2}{N}} \cos\left(\frac{(2j+1)i\pi}{2N}\right) & i=0, 1, \dots, N, j=0, 1, \dots, N-1 \end{cases} \quad (2)$$

Avec  $C(i, j)$  représente les coefficients de la DCT à la position  $(i, j)$ .

L'un des inconvénients de la compression d'image en utilisant la DCT est l'effet mosaïque ou la pixellisation. Lors d'une quantification forte, la division en bloc devient visible car un bloc entier se trouve codé avec la même valeur (peu de coefficients non nuls). Nous allons utiliser cette technique de compression pour notre comparaison avec un niveau de quantification élevé pour quantifier la différence entre les deux méthodes utilisées.

## III. RESULTATS ET DISCUSSIONS

Nous avons utilisé dans cette étude comparative comme le montre la Fig. 2 une base de données contenant quatre images testes de résolution  $512 \times 512$  avec 8 bits par pixel très utilisées dans l'évaluation des méthodes de compression d'images. Ces images comportent des textures de complexité différente sont : Lena, Barbara, Mandrill et Fingerprint.

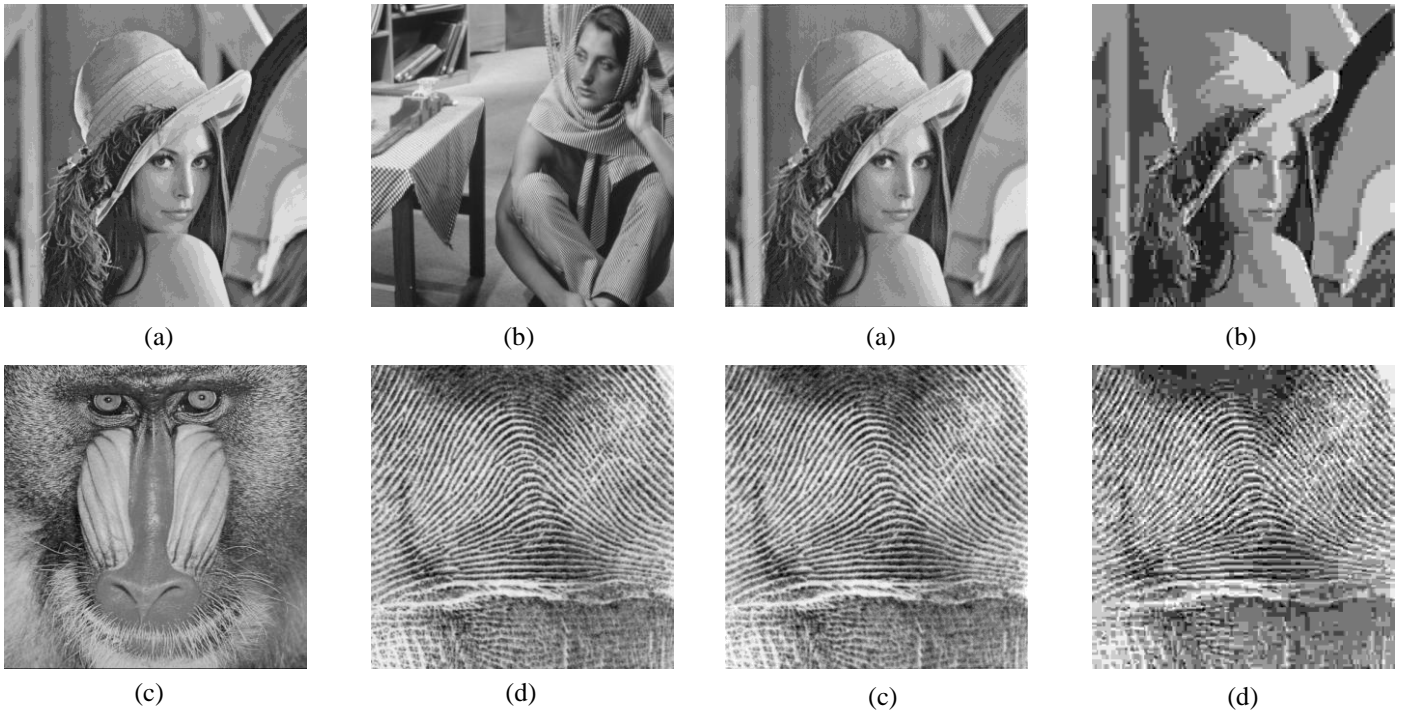


Fig. 2 Les images de la base de données utilisée.  
 (a) Lena, (b) Barbara, (c) Mandrill, (d) Fingerprint

Pour évaluer les deux méthodes et quantifier la différence entre elles, Nous avons utilisé le PSNR et le taux de compression. Le PSNR permet de mesurer la distorsion introduite par l'opération de compression. Le PSNR est donné par l'équation suivante.

$$\text{PSNR}(I_0, I_c) = 20 \log_{10} \left( \frac{2^r - 1}{\sqrt{\text{MSE}(I_0, I_c)}} \right) \quad (3)$$

Avec  $I_0$  et  $I_c$  représentent respectivement l'image originale et l'image reconstruite de taille  $M \times N$  et  $r$  représente la résolution numérique de l'image qui est 8 dans notre cas.

Le MSE (Mean Square Error) représente l'erreur quadratique moyenne qui est définie comme.

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_0(i,j) - I_c(i,j))^2. \quad (4)$$

Le taux de compression permet d'évaluer l'efficacité et la puissance de la méthode de compression d'images choisie. Le taux de compression est donné par l'équation suivante.

$$\text{RC} = \frac{\text{Nombre de bits dans l'image originale}}{\text{Nombre de bits dans l'image compressée}} \quad (5)$$

Le taux de compression en pourcentage  $T_c$  est représenté comme suit.

$$T_c = \left( 1 - \frac{1}{\text{RC}} \right) \times 100 \quad (6)$$

La Fig. 3 donne la reconstruction des images de la base de données pour les deux méthodes utilisées.

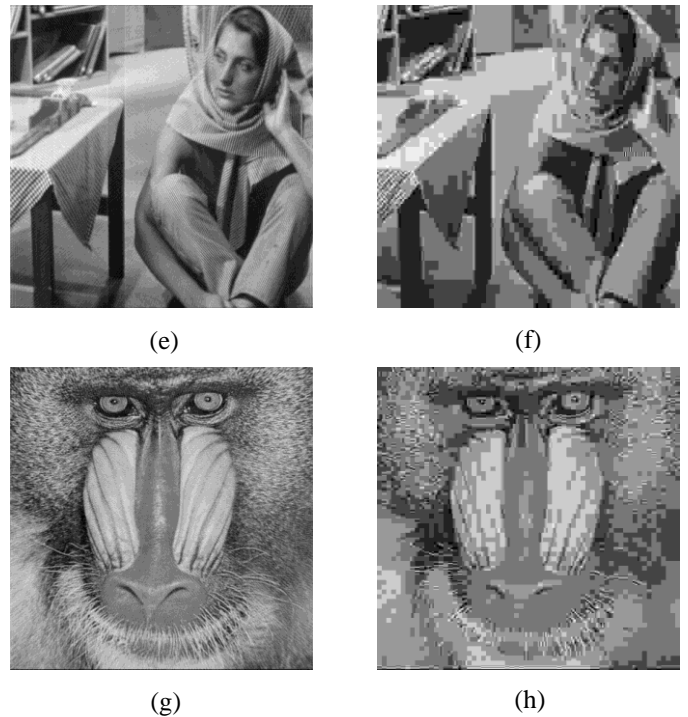


Fig. 3 Reconstruction des images de la base de données pour les deux méthodes.

- (a) Lena avec la méthode Radon,
- (b) Lena avec la méthode DCT,
- (c) Fingerprint avec la méthode Radon,
- (d) Fingerprint avec la méthode DCT
- (e) Barbara avec la méthode Radon.
- (f) Barbara avec la méthode DCT.
- (g) Mandrill avec la méthode Radon.
- (h) Mandrill avec la méthode DCT.

Nous remarquons d'après la Fig. 3 que les images reconstruites par la méthode Radon sont de qualité meilleure que celles reconstruites avec la méthode DCT. Nous remarquons aussi que l'effet de pixellisation est imperceptible dans les images reconstruites par la méthode basée sur la transformée de Radon malgré le niveau de quantification est important. Cela est assuré, comme le montre la Fig. 4 et la Fig. 5, que la transformée de Radon atténue les fluctuations d'intensité par le processus d'intégration [19].

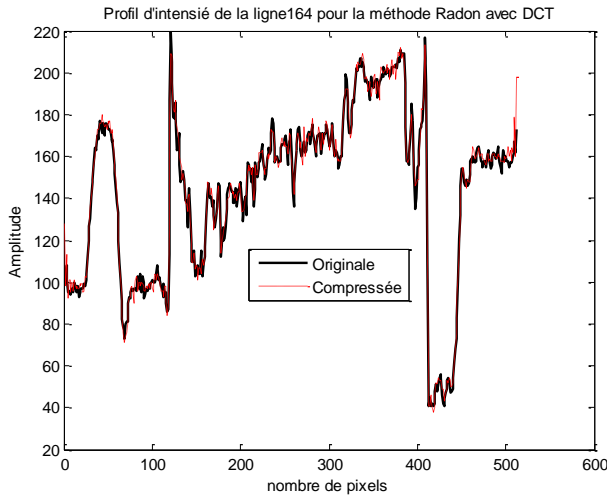


Fig. 4 Profil d'intensité le long de la ligne 164 de l'image Lena originale et reconstruite par la méthode Radon .

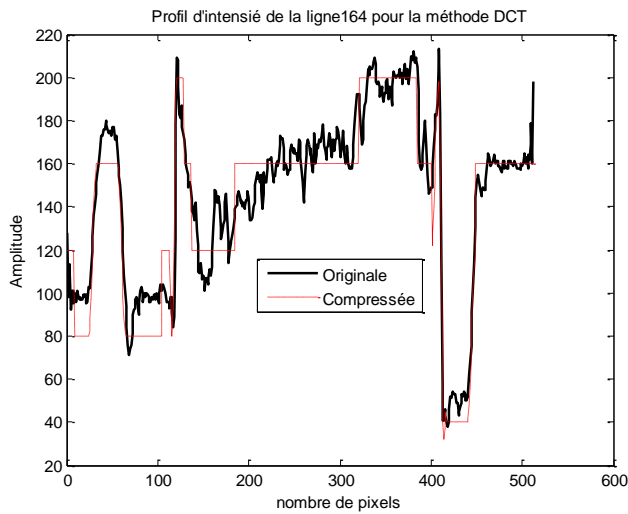


Fig. 5 Profil d'intensité le long de la ligne 164 de l'image Lena originale et reconstruite par la méthode DCT.

Nous remarquons d'après la Fig. 4, que le profil d'intensité de l'image Lena originale et presque le même avec celui de l'image reconstruite dans le cas de la méthode Radon. Par contre, dans le cas de la Fig. 5, les fluctuations d'intensité sont remarquable ce qui justifie l'effet mosaïque de l'image reconstruite avec la méthode DCT. Donc les valeurs des pixels sont différentes et les bords ne sont pas nets.

L. M. Murphy a montré que la transformée de Radon atténue les fluctuations par le processus d'intégration [19]. Ce processus vient du fait que la transformée de Radon inverse est

obtenue en utilisant l'algorithme de la rétroprojection des projections filtrées FBP (Filter-BackProject). La FBP permet de filtrer chaque point Radon (projection) par le filtre Ram-Lak [22]. La projection filtrée  $\tilde{I}_\theta(p)$  est donnée par l'intégrale suivante:

$$\tilde{I}_\theta(p) = \int_{-\infty}^{+\infty} F_\theta(v) |v| e^{i2\pi vp} dv \quad (7)$$

Avec  $F_\theta(v)$  représente la transformée de Fourier 1-D de la projection  $\hat{I}_\theta(p)$ . Les étapes de la FBP sont résumées comme suit :

- Pour chaque angle de projection  $\theta$ , calculer la transformée de Fourier 1-D de la projection  $\hat{I}_\theta(p)$  pour obtenir  $F_\theta(v)$ .
- Multiplier  $F_\theta(v)$  par  $|v|$  (Filtre Ram-Lak).
- Pour chaque  $\theta$ , calculer la transformée de Fourier inverse de  $F_\theta(p)|v|$  pour obtenir la projection filtrée  $\tilde{I}_\theta(p)$  par l'application de l'équation 7.
- Rétro-projeter le sinogramme filtré  $\tilde{I}_\theta(p)$  pour récupérer l'image de départ  $I$ . Le sinogramme filtré est donné par l'équation suivante :

$$I = \int_{-\infty}^{\infty} \hat{I}_\theta(p) (x \cos \theta + y \sin \theta) d\theta \quad (8)$$

Le filtre Ram-Lak est défini comme suit :

$$H(v) = \begin{cases} |v| & \text{Si } -v_{\max} \leq v \leq v_{\max} \\ 0 & \text{Sinon} \end{cases} \quad (9)$$

Le filtre Ram-Lak est représenté par la Fig. 6.

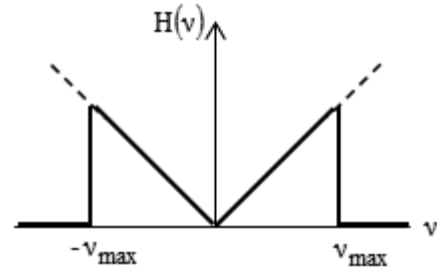


Fig. 6 Le filtre Ram-lak

Les différents résultats des critères d'évaluation sans listés dans le tableau 1 et représentés comme graphes par les Fig. 7 et Fig. 8.

TABLE I  
Résultats du PSNR et du taux de compression

Images tests	Radon avec DCT		DCT	
	PSNR (db)	Tc (%)	PSNR (db)	Tc (%)
Lena	34.02	56.77	24.02	98.78
Barbara	27.95	50.96	22.98	98.72
Mandrill	24.74	39.37	19.58	98.70
Fingerprint	30.79	36.60	19.82	97.36

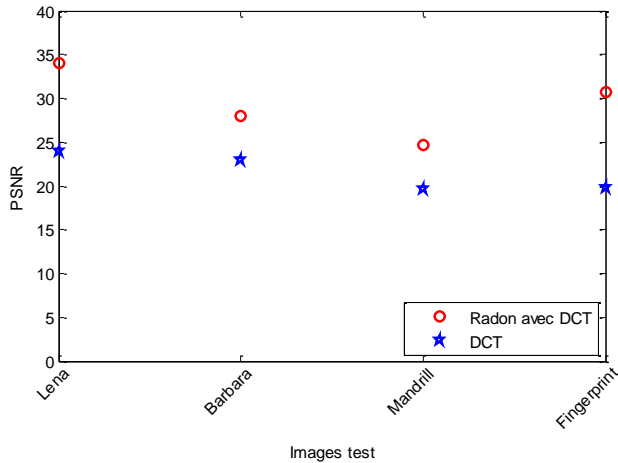


Fig. 7 Valeurs du PSNR (dB) des images reconstruites pour les deux méthodes.

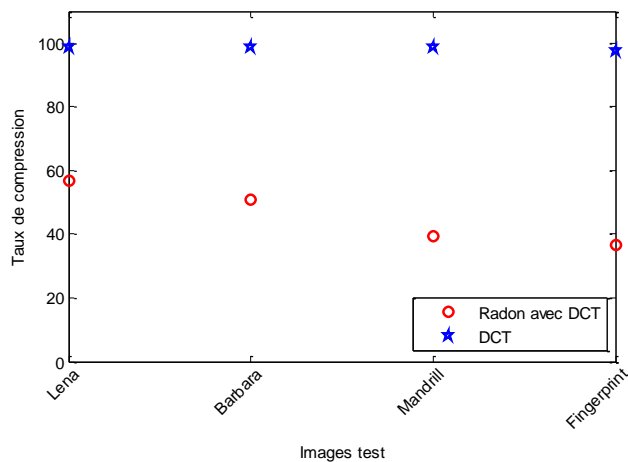


Fig. 8 Valeurs du taux de compression (%) des images reconstruites pour les deux méthodes.

Nous observons d'après les Fig. 7 et Fig. 8 que malgré la forte quantification, la méthode de compression d'images basée sur la transformée de Radon permet d'assurer un bon compromis débit-distorsion. Alors que la méthode de

compression basée sur la DCT génère des images de qualité inférieure malgré le taux élevé de la compression. Ces résultats confirment la puissance de la transformée de Radon dans l'atténuation des différents bruits et son utilisation dans la compression et la détection des objets dans des images bruitées.

Une autre remarque importante que nous pouvons tirer à partir des résultats obtenus est que la valeur du PSNR est maximale pour l'image Lena, par contre elle est minimale pour l'image Mandrill. Ce résultat est justifié, comme le montre le sinogramme de chaque image par la forte texture de l'image Mandrill par rapport à celle de l'image Lena. Les deux sinogrammes sont représentés par les Fig. 9 et 10.

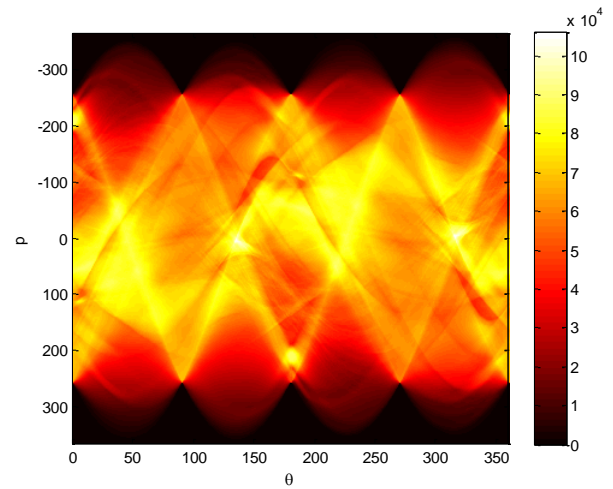


Fig. 9 Sinogramme de l'image Lena.

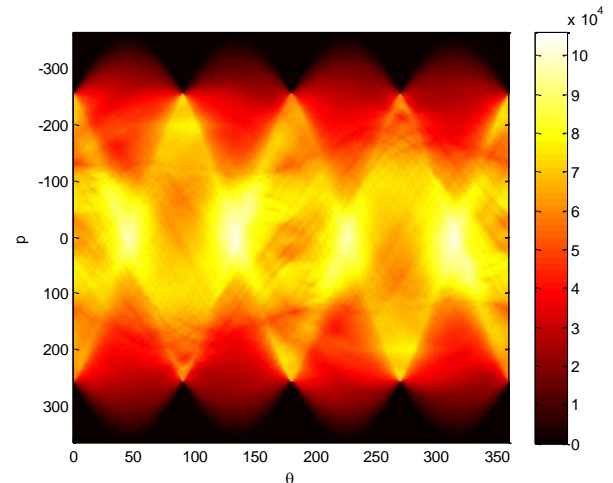


Fig. 10 Sinogramme de l'image Mandrill.

D'après ces deux sinogrammes, Nous remarquons que la perte des détails fins sera élevée pour l'image Mandrill par rapport aux autres images de la base de données lors de la compression.

Nous pouvons confirmer cette remarque en visualisant les deux sinogrammes en 3-D à partir de la Fig. 11 et 12.

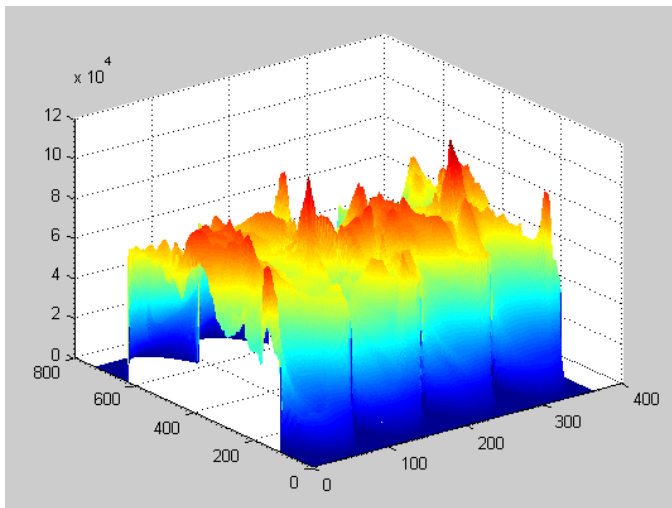


Fig. 11 Sinogramme en 3-D de l'image Lena.

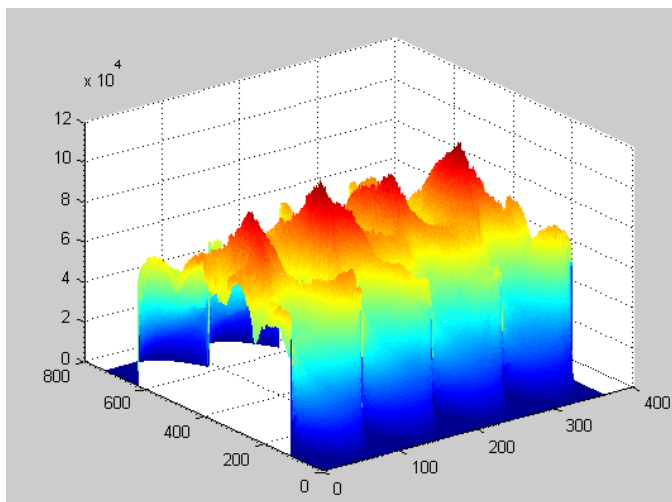


Fig. 12 Sinogramme en 3-D de l'image Mandrill.

La Fig. 12 montre clairement la forte concentration de l'énergie présentant dans les quatre pics représentés par la Fig. 10. Nous remarquons aussi d'après la Fig. 11, la faible intensité des pics présentant dans le sinogramme de l'image Lena. Ces deux remarques montrent clairement la forte texture des objets présentant dans l'image Mandrill ce qui permet de récupérer lors du calcul de la transformée de Radon des point Radon d'intensité supérieure à ceux récupérés pour l'image Lena.

#### IV. CONCLUSION

Dans cet article, nous avons comparé deux méthodes de compression d'images qui sont : la compression d'images basée sur la transformée de Radon et la méthode de compression d'images basée sur la DCT. Cette comparaison a été réalisée pour un niveau de quantification élevé égal à 20 ce qui nous a permis de vérifier la puissance de la transformée de Radon vis-à-vis du bruit de quantification. Nous avons remarqué d'après cet article que malgré la forte quantification,

la méthode de compression d'images basée sur la transformée de Radon reste un bon choix pour l'archivage. Le point le plus important dans cette comparaison est la vérification du processus d'intégration dans le cas de la forte quantification.

#### REMERCIEMENTS

Nous remercions S. A. Pradeep et R. Manavalan d'avoir accepté d'utiliser le schéma de leur méthode pour des fins de recherche.

#### REFERENCES

- [1] D. A. Huffman, "A method for the construction of minimum redundancy codes", IRE Proc. Vol. 40, No.9, pp. 1098-1101, 1952. DOI: 10.1109/JRPROC.1952.273898.
- [2] R. F. Rice and R. Plaunt, "Adaptive Variable-Length Coding for Efficient Compression of Spacecraft Television Data," IEEE Transactions on Communications, vol. 16(9), pp. 889-897, 1971. DOI: 10.1109/TCOM.1971.1090789.
- [3] S. M. Bacha, B. C. Jinaga "An Optimum Novel Technique Based on Golomb-Rice Coding for Lossless Image Compression of Digital Images " International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.6, No.5, pp.291-304, 2013.
- [4] B. Parker, "Synthesis of noiseless compression codes", thèse de Doctorat, Georgia Institute of Technology, p. 108, 1967.
- [5] P. Devaki, Dr. G. Raghavendra "Lossless reconstruction of secret image using threshold secret sharing and transformation" International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.3, pp.111-119, May 2012.
- [6] A. Devi, M. G. Mini "Gray scale image compression based on wavelet transform and linear prediction" The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.1, pp.47-62, February 2012.
- [7] S. A. Pradeep, R. Manavalan "Image compression using Radon transform with DCT : Performance analysis" International Journal of Scientific Engineering and Technology Vol. 2, No. 8, pp : 759-765 2013.
- [8] N. Ahmed, T. Natarajan and K. Rao, "Discrete Cosine Transform", IEEE Computers Transactions, vol.23, no.1, pp.90-94, 1974. DOI: 10.1109/T-C.1974.223784.
- [9] J. Radon, "On the determination of functions from their integrals along certain manifolds", translation of Radon's 1917 paper by R. Lohner, The Radon transform and some of its applications, Annexe A, pp. 204-217, John Wiley & Sons, 1983.
- [10] A. Kingston, F. Atrousseau, "Lossless Image Compression via Predictive Coding of Discrete Radon Projections", Signal Processing Image Communication, vol. 23, no. 4, pp. 313-324, 2008.
- [11] Z. Haddad, A. Beghdadi, A. Serir and A. Mokraoui, "Fingerprint Identification using Radon Transform, Image Processing Theory, Tools and Applications", IPTA 2008 IEEE, pp. 1-7, 2008, E-ISBN: 978-1-4244-3322-3. DOI: 10.1109/IPTA.2008.4743781.
- [12] S. A. Angadi, S. Gour, G. Bhajantri, "Offline Signature Recognition System Using Radon Transform", IEEE Signal and Image Processing (ICSIP), pp. 56-61, 2014. DOI: 10.1109/ICSIP.2014.13.
- [13] M. Lahdir, S. Ameer, A. Adane, "Algorithme non itératif basé sur les ondelettes biorthogonales et les fractales pour la compression d'images satellitaires, Télé-détection, Vol. 6, n. 4, pp. 345-360, 2007.
- [14] J. M. Shapiro "Embedded image coding using zerotrees of wavelet coefficients" IEEE transactions on signal processing Vol. 41 No. 12 December 1993.
- [15] M. Lahdir, A. Nait-ali, S. Ameer, "Fast Encoding-Decoding of 3D Hyperspectral Images Using a Non-Supervised Multimodal Compression Scheme, Journal of Signal and Information Processing, Vol. 2, n. 04, pp. 316-321, 2011. DOI: 10.4236/jsip.2011.24045
- [16] A. Said, W. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees". IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 3, pp. 243-250, 1996. DOI: 10.1109/76.499834.

- [17] J. Guédon “ Psychovisual image coding via an exact discrete Radon transform “ Proc. Visual Communications & Image Processing (VCIP), Lance T. Wu editor, Vol. 2501, pp.562-572, Taipei, Taiwan ,May 1995.
- [18] R. Tyagi, D. K. Sharma “ Digital image compression comparisons using DPCM and DPCM with LMS algorithm” International Journal of Computer Applications & Information Technology Vol. 1, No. 2, pp. 65-71, September 2012 .
- [19] L. M. Murphy, “ Linear feature detection and enhancement in noisy images via the Radon transform “, Pattern Recognition Letters Vol. 4, No. 4, pp. 279-284, 1986.
- [20] P. Courmontagne “An improvement of ship wake detection based on the radon transform” Signal Processing Vol. 85, No. 8, pp. 1634–1654, 2005.
- [21] K. Sayood, “ Introduction to data compression “, 4<sup>th</sup> Edition, Morgan Kaufmann, p.743, 2012.
- [22] G. N. Ramachandran and A. V. Lakshminarayanan, “Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms”, . Proceedings of the National Academy of Sciences. vol. 68, no. 9, pp. 2236–2240, 1971.

# Système de Diagnostic de l'Infarctus du Myocarde par le Réseau de Neurones LVQ et la Carte SOM-LVQ

Djazia Bendani<sup>1</sup>, Zidelkhir Chalabi<sup>2</sup>, Nasreddine Berrached<sup>3</sup>

<sup>1,2,3</sup>Laboratoire de Recherche en Systèmes Intelligents. Département d'Electronique.  
Université des Sciences et de la Technologie d'Oran. Mohamed Boudiaf.

Bir El Djir Oran BP1505. 31000- Algérie

bendani\_d@yahoo.fr  
hzidelkhir@yahoo.fr  
neberrached@netscape.net

**Résumé**—Le présent article concerne la détection semi-automatique de pathologie myocardique: l'infarctus du myocarde (IDM), dans le cadre d'une aide au diagnostic. Nous mettons en compétition deux classifieurs neuronaux: LVQ et SOM-LVQ. Les vecteurs d'entrée de ces réseaux sont les paramètres caractéristiques de l'infarctus du myocarde dans ses différentes phases (ischémie, lésion et nécrose), repérés sur le signal ECG filtré et extrait par la décomposition en ondelettes du chapeau mexicain. Ces paramètres constituent la base d'apprentissage et de généralisation. Les signaux exploités dans cette étude proviennent de la base de données «ST-T European Database». Notre objectif est d'améliorer la détection par rapport à d'autres travaux cités en référence et en comparant les classifieurs entre eux. Les résultats obtenus sont très satisfaisants avec des taux de reconnaissance  $Tr$  moyens supérieurs à 93% et des  $Tr$  de 100% pour la nécrose.

**Mots clés**—Détection; IDM; Ondelettes; LVQ; SOM-LVQ.

## I. INTRODUCTION

Le signal ECG est la représentation de l'activité électrique du muscle cardiaque, enregistrée à la surface du corps. Le tracé ECG est formé d'une série d'ondes désignées par les lettres P, Q, R, S, T et U. Du point de vue traitement du signal, l'ECG se présente comme un signal aléatoire composite, structuré, de nature segmentale. Il est aléatoire par le déroulement spatiotemporel non déterministe des processus électrophysiologiques [1]. Les maladies cardiaques dues aux problèmes électriques ou anatomiques chroniques, produisent des changements temporaires ou permanents, complètement ou partiellement dans la forme d'onde de l'ECG de base [2]. L'analyse de l'ECG comprend la mesure des amplitudes et des durées ainsi que l'examen morphologique des ondes (P, QRS, T), des intervalles RR, PR, QT et le segment ST (Fig. 1).

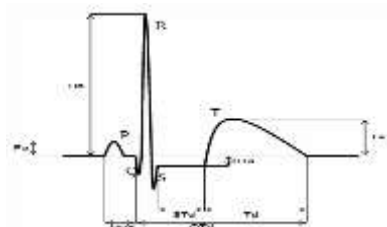


Fig. 1. Paramètres essentiels pour la détection d'un battement

Notre objectif est la mise en évidence de l'infarctus du myocarde (IDM). L'IDM est défini comme une diminution de l'apport en oxygène aux cellules myocardiques [3], responsable de la mort de ces cellules et donc de la destruction d'une partie du muscle. Il est caractérisé par un changement soudain du complexe ST-T. Dans la forme typique d'un IDM transmural [3] trois types d'anomalies électriques apparaissent: ischémie, lésion, et nécrose.

La procédure que nous proposons ici tient compte des avantages des outils d'analyse non-stationnaire par la transformée en ondelettes (TO) et des outils d'analyse non linéaire par les réseaux de neurones dans le but d'améliorer les performances de classification. La méthode des ondelettes est devenue une technique, par excellence, très efficace pour extraire les paramètres qui caractérisent l'IDM dans ses différentes phases. Dans l'article [4] Senhadji a largement démontré l'apport de la décomposition en ondelettes du signal ECG. Elle permet de produire les instants de début et de fin de chaque onde (P, Q, R, S et T) ainsi que les intervalles RR et QT. Dans [5], la TO est vue comme un outil pour décrire des changements locaux dans les oscillations du rythme cardiaque.

Deux modèles de réseaux neuronaux destinés à l'identification de l'IDM sont mis en œuvre: le réseau de Kohonen LVQ (Learning Vector Quantization) et la carte SOM-LVQ. Cette analyse non linéaire semble prometteuse pour fournir des indicateurs pertinents à l'identification de la pathologie.

Beaucoup de travaux ont été élaborés pour la détection des anomalies ischémiques par les réseaux neuronaux [2], [6], [7], [8]. Dans l'article [9], le taux de reconnaissance pour la détection d'épisodes ischémiques par la carte SOM est de 83.33%, le vecteur de paramètres contient la déviation du segment ST, de l'onde T et l'amplitude des ondes Q, R, et S.

Ces paramètres sont mesurés battement par battement sur deux dérivations de l'ECG. La sensibilité de détection de l'ischémie myocardique par l'algorithme de la rétropropagation du gradient est de 88.62% selon [10], de même un taux de 88.6% a été marqué en employant le LVQ dans [11], l'élévation des segments ST sur les 12 dérivations de l'ECG a été le paramètre d'entrée de ce classifieur dans le but de localiser le siège de l'IDM (inférieur, antérieur).



La classification semi-automatique est d'un intérêt majeur surtout en surveillance en soin intensif ou encore dans le cas de dépistage systématique. Les classifieurs sont, ici, dédiés à reconnaître deux types de signaux : ECG normal et ECG anormal représentant l'IDM.

Ce papier est organisé comme suit : La section 2 définit l'étape d'analyse, la section 3 les outils de classification utilisés. La section 4 concerne l'application et les résultats obtenus. Les commentaires sont traités en 5. Nous concluons par des remarques dans la section 6.

## II. EXTRACTION DES PARAMÈTRES

### A. Objectif

L'objectif est l'extraction des paramètres permettant de diagnostiquer l'IDM: une description compacte du complexe ST-T composé principalement du segment ST et de l'onde T. Ces indicateurs seront issus d'une technique d'analyse non-stationnaire.

La transformée en ondelettes est l'outil privilégié pour cette tâche. Elle consiste à décomposer un battement au moyen d'une ondelette en un certain nombre de niveaux de détail et de retenir l'extrema global à chaque niveau de la décomposition pour construire les vecteurs caractérisant les battements qui seront par la suite les vecteurs d'entrée aux classifieurs. Pour ce faire nous avons appliqué l'analyse en ondelette du chapeau mexicain (Sombrero) (Voir Fig. 2). Son succès est dû à son adaptabilité aux données. Le signal ECG est non-stationnaire, l'application de la TO est très adaptée puisqu'elle possède un caractère de localisation temporelle permettant de donner des informations précises [4], [12].

### B. Transformée en Ondelette Continue

La transformée en Ondelette continue (T.C.O) permet de décomposer un signal selon une série de signaux appelée les ondelettes  $\psi(t)$ , déduites d'une ondelette mère  $\psi(t)$  par des opérations de translation et de dilatation. Les coefficients d'ondelettes  $C$  quantifient la ressemblance du signal à analyser  $f(t)$  avec l'ondelette  $\psi$  [13]. Cette grandeur est le produit scalaire  $\langle f, \psi_{a,b} \rangle$ , opérateur de projection classique :

$$C_{a,b} = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(t) \overline{\psi_{a,b}(t)} dt \quad (1)$$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2)$$

$$C_{a,b} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (3)$$

Avec :  $a, b \in \mathbb{R}, a \neq 0$

Où  $a$ : facteur d'échelle,  $b$ : paramètre de translation, et  $b/a$  est le pas de translation à l'échelle  $a$ .

L'ondelette du Chapeau mexicain de la Fig. 2 est une ondelette continue qui s'écrit selon la formule mathématique suivante :

$$\psi(x) = \left( \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} \right) (1-x^2) e^{-\frac{x^2}{2}} \quad (4)$$

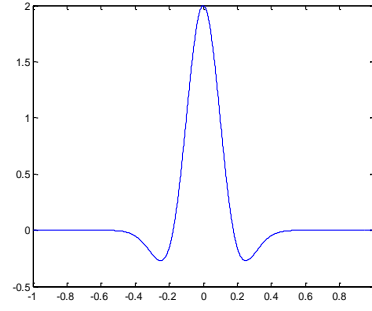


Fig. 2. Forme générale de l'ondelette du chapeau mexicain

## III. CLASSIFICATION

Le choix du classifieur est très important, il constitue l'élément de décision dans un système de reconnaissance de formes. L'approche connexionniste neuronale conserve l'idée du calcul récursif, parallèle et rapide pendant la phase d'apprentissage. Lorsque les classes se présentent avec des surfaces de séparation très complexes, les réseaux supervisés LVQ et SOM-LVQ sont plus aptes en raison de leur apprentissage local par la VQ. C'est de la construction du dictionnaire que la QV tient des spécificités de l'application donnée permet donc d'exploiter les informations a priori. Un aperçu sur les deux algorithmes de classification automatique d'apprentissage supervisé sera présenté ci-dessous.

### A. Le Réseau LVQ

Le réseau LVQ, appelé Learning Vector Quantization, développé par Kohonen, est un réseau à trois couches dont le fonctionnement est basé sur l'apprentissage supervisé qui consiste à réajuster les vecteurs de poids de manière à minimiser le taux global d'erreurs de classification. Dans notre travail la méthode LVQ1 a été adoptée, son principe général est le suivant : [6], [9], [14]-[15]

1. Fixer le seuil pour l'erreur quadratique minimale (EQM), le nombre maximal d'itérations et le nombre de neurones de la couche cachée. Les poids de chaque neurone de la couche cachée sont initialisés à des valeurs aléatoires.
2. Calcul de la distance (distance euclidienne) entre le vecteur à classer et les poids de chaque neurone de la couche cachée. La distance euclidienne entre ces deux vecteurs est définie par :

$$D_E(X, W) = \sqrt{\sum_{i=1}^n (x_i - w_i)^2} \quad (5)$$

Avec  $X = [x_1, x_2, \dots, x_n]$  Vecteur d'entrée,

$W = [w_1, w_2, \dots, w_n]$  Vecteur poids.

3. Le neurone le plus proche du vecteur d'entrée est déclaré "vainqueur" et sa sortie passe à 1, tandis que les sorties des autres neurones restent à 0. Seuls les poids de ce neurone sont modifiés par la règle d'apprentissage suivante :

$$W_i(t+1) = W_i(t) \quad (6)$$

Si le neurone n'est pas vainqueur

$$W_i(t+1) = W_i(t) + \alpha(t)(x - w_i) \quad (7)$$

S'il est vainqueur et est associé à la classe de  $x$

$$W_i(t+1) = W_i(t) - \alpha(t)(x - W_i) \quad (8)$$

S'il est vainqueur mais n'est pas associé à la classe de  $x$ .

$\alpha$ , le coefficient d'apprentissage, varie selon la condition suivante  $0 < \alpha(t) < 1$  et décroît avec le temps.

Arrêter les itérations lorsque le seuil de l'EQM ou le nombre maximal autorisé d'itérations est atteint.

4. Pour chaque groupe de neurones de la couche cachée correspondant à la même classe réaliser un OU logique de leurs sorties au moyen du neurone de sortie associé à celle-ci.

### B. La carte SOM-LVQ

Le réseau SOM-LVQ est une combinaison entre de la carte SOM de Kohonen et son algorithme modifié. La carte SOM est un ensemble structuré d'unités de traitement (neurones), qui sont disposées en une seule couche de laquelle émane une topologie définie par une notion de voisinage des cellules [15].

Le principe est identique que le LVQ sauf que les poids de neurones de la carte se calculent selon la formule "(9)"

$$W_j(t+1) = W_j(t) + \alpha(t)\beta_{i,c}(t)[X_i(t) - W_j(t)] \quad (9)$$

avec  $\beta_{i,c}(t)$  : une fonction décrivant le rayon du voisinage, active associée au neurone élu  $c$ .

## IV. SIMULATION ET RÉSULTATS

La simulation comprend la sélection des signaux ECG, l'analyse des battements par la TO, ensuite la classification par les deux réseaux et l'évaluation des performances.

### A. Base de Données

Les enregistrements sont issus de ST-T European Database, ils représentent l'IDM dans ses différentes phases définies et reportées dans le TABLEAU I. Chacun des dix-neuf enregistrements a une durée de deux heures avec une fréquence d'échantillonnage  $f_e = 250 \text{ Hz}$ . La base d'apprentissage est composée de 8771 battements dont 4307 normaux et 4464 anormaux, celle du test contient 8787 battements avec 4314 battements normaux et 4473 anormaux.

TABLEAU I  
DESCRIPTION DES ENREGISTREMENTS UTILISÉS

Enregistrement	Type d'anomalie
e0113 - e0123 - e0133 - e0151 - e0163	Ischémie sous-endocardique (IE)
e0147 - e0161-e0303 - e0409	Lésion sous-endocardique (LE)
e0121 - e0125 - e0509	Ischémie sous-épicaire (IP)
e0105 - e0129 - e0601	Lésion sous-épicaire (LP)
e0203 - e0403 - e0615 - e0817	L'onde Q de nécrose (Q)

### B. Extraction des Paramètres

Les paramètres représentatifs et pertinents constituant la base d'apprentissage et de test sont choisis avec soin pour accroître les performances de classification en battements normaux et anormaux. En effet la localisation de l'IDM joue

un rôle crucial dans la dynamique de la fréquence cardiaque et de l'activité du système nerveux autonome [12].

Le choix de l'ondelette dépend surtout de l'application, l'ondelette du chapeau mexicain sa transformée est continue, et dotée d'une symétrie de traitement droite-gauche par rapport à la position des ondes à détecter, nous a permis selon différentes décompositions de prélever:

1. La fréquence cardiaque.
2. Durée et amplitude des ondes (Q, R et S).
3. Durée, amplitude et surface de l'onde T.
4. Durée, amplitude et surface et pente par rapport à la ligne isoélectrique du segment ST.
5. L'intervalle QT.

D'après notre expertise le complexe QRS s'extrait par le coefficient 4, par contre la localisation de l'onde T et du segment ST nécessite des coefficients différents selon la nature du signal et du type d'anomalie. Le TABLEAU II reporte, en exemple, les coefficients relatifs à l'onde T.

TABLEAU II  
COEFFICIENTS D'ONDELETTES UTILISÉS POUR LA LOCALISATION DE L'ONDE T

Enregistrement	Coefficients d'ondelette
e0113 - e0163 - e0123 - e0129 - e0133 - e0151 - e0601	20
e0147 - e0161- e0303 - e0409	10
e0121 - e0125 - e0509	10
e0105	16
e0203 - e0403 - e0615 - e0817	9

### C. Classification

L'architecture de notre réseau LVQ est définie comme suit :

1. Une couche d'entrée de 8 neurones chacun d'eux correspond respectivement à la fréquence cardiaque, durée de l'onde T, durée du segment ST, amplitude de l'onde T, amplitude du segment ST, surface de l'onde T, surface du segment ST et enfin la pente du segment ST.
2. Une couche cachée de 6 neurones.
3. Une couche de sortie de 2 neurones.

L'apprentissage du réseau LVQ est proprement supervisé dont la phase d'apprentissage va s'arrêter sous l'une de ces conditions :

1. L'erreur quadratique moyenne est inférieure à  $10^{-10}$ .
2. Le nombre d'itérations maximum qui est égal à 500 est atteint.

L'apprentissage du réseau SOM-LVQ est hybride puisqu'il est non-supervisé pour la couche cachée (la carte) et supervisé pour la couche de sortie. En suivant les recommandations de Kohonen, et après plusieurs essais, l'ajustement de la carte SOM a été réalisé en passant par les étapes suivantes :

1. Initialement la base d'apprentissage a été normalisée et les vecteurs poids synaptiques initialisés aléatoirement.
2. Les paramètres d'apprentissage choisis : voisinage hexagonal, la taille de la carte fixée à  $11 \times 8$  neurones, l'algorithme par défaut est le batch algorithme.

Enfin l'algorithme LVQ1 a été appliqué, pour toutes les cartes développées avec le batch algorithme. Après plusieurs essais on est arrivé à un coefficient d'apprentissage égal à 0.01 et un nombre d'itérations entre 200 et 2000 a permis de stabiliser le taux de reconnaissance.

Les figures 4, 5 et 6 représentent respectivement la qualité de séparation et de reconnaissances des battements normaux et pathologiques ainsi que la distribution des paramètres de chaque cycle sur la carte.

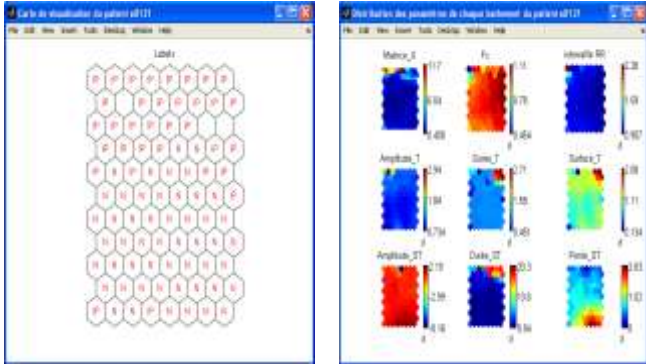


Fig. 3. LA CARTE SOM DE L'ANOMALIE IP : N CYCLES NORMAUX, IP ANORMAUX

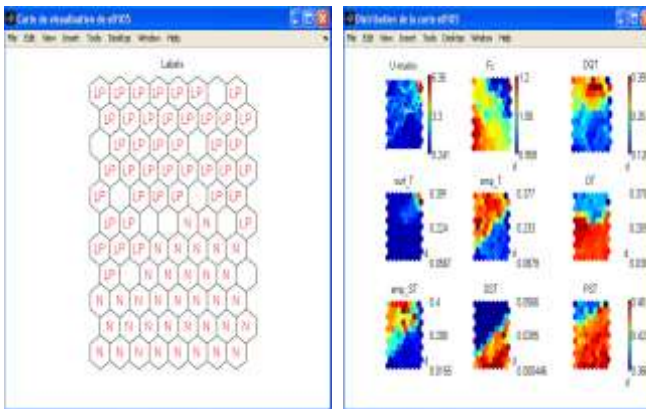


Fig. 4. LA CARTE SOM DE L'ANOMALIE LP : N CYCLES NORMAUX, LP ANORMAUX

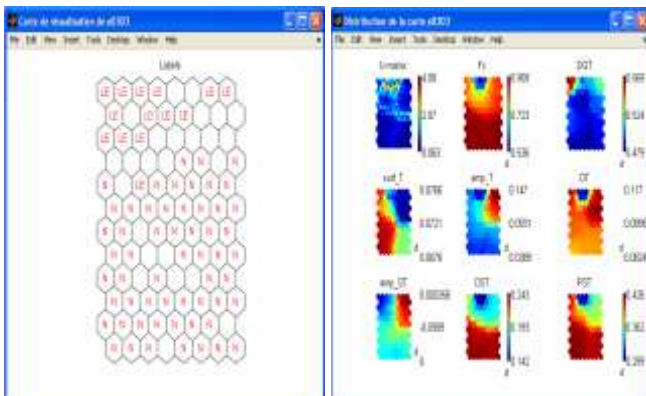


Fig. 5. LA CARTE SOM DE L'ANOMALIE LE : N CYCLES NORMAUX, LE ANORMAUX

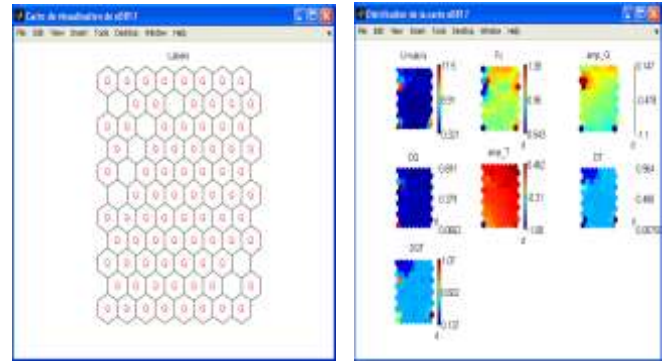


Fig. 6. LA CARTE SOM DE L'ANOMALIE Q : Q CYCLES ANORMAUX

Bien sûr après l'entraînement des réseaux, il faut les tester afin de vérifier leurs performances. En phase de validation, on applique trois lois statistiques (prédictibilité, sensibilité et taux de reconnaissance).

1. La prédictibilité (Pr) :

$$Pr = \frac{TP}{(TP + FP)} \quad (10)$$

2. La sensibilité (Se) :

$$Se = \frac{TP}{(TP + FN)} \quad (11)$$

3. Le taux de reconnaissance (Tr) :

$$Tr = \frac{(TP + TN)}{(TN + FP + FN + TP)} \quad (12)$$

Avec :

- TP (Vrai Positif) : représente un battement anormal détecté comme anormal.
- FP (Faux Positif) : représente un battement anormal détecté comme normal.
- FN (Faux Négatif) : représente un battement normal détecté comme anormal.
- TN (Vrai Négatif) : représente un battement normal détecté comme normal.

Les Tableaux III, IV et V montrent respectivement le taux de reconnaissance par rapport aux différentes anomalies ainsi que les taux de prédictibilité, sensibilité et de reconnaissance des réseaux LVQ et SOM-LVQ

TABLEAU III  
TAUX DE RECONNAISSANCE PAR ANOMALIE DE CHAQUE RÉSEAU

Pathologie	Tr (%)	
	LVQ	SOM-LVQ
<b>Ischémie sous-endocardique</b>	96,26	66,72
<b>Lésion sous-endocardique</b>	94,87	64,07
<b>Ischémie sous-épicaudique</b>	92,79	88,97
<b>Lésion sous-épicaudique</b>	98,62	59,47
<b>Onde Q de nécrose</b>	100	98,36

TABLEAU IV  
RÉSULTATS DU RÉSEAU LVQ

Patients	LVQ		
	Pr (%)	Se (%)	Tr (%)
e0105	97,77	100	98,99
e0113	96,84	100	98,88
e0121	42,47	89,95	78,37
e0123	96,82	100	98,83
e0125	100	100	100
e0129	81,42	100	97,29
e0133	31,42	91,66	94,39
e0147	92,59	100	99,34
e0151	89,03	100	96,65
e0161	0	Nan	97,42
e0163	57,14	100	92,53
e0203	100	100	100
e0303	0	Nan	86,53
e0403	100	100	100
e0409	100	72,22	96,18
e0509	100	100	100
e0601	99,37	100	99,58
e0615	100	100	100
e0817	100	100	100
<b>La moyenne</b>	<b>78,17</b>	<b>87,04</b>	<b>96,58</b>

TABLEAU V  
RÉSULTATS DU RÉSEAU SOM-LVQ

Patients	SOM-LVQ		
	Pr (%)	Se (%)	Tr (%)
e0105	95,55	96,62	96,48
e0113	76,84	100	91,97
e0121	78,13	81,53	86,2
e0123	85,71	81,81	87,71
e0125	90,29	100	90,29
e0129	82,85	93,54	96,67
e0133	0	0	97,13
e0147	96,51	95,36	96,22
e0151	76,27	75	83,84
e0161	0	0	94,04
e0163	94,80	84,88	96,15
e0203	100	100	100
e0303	98,24	59,57	90,64
e0403	100	100	100
e0409	61,53	44,44	88,54
e0509	98,5	98,5	98,87
e0601	90,68	96,68	91,76
e0615	98,15	100	98,15
e0817	98,36	100	98,36
<b>La moyenne</b>	<b>80,13</b>	<b>79,36</b>	<b>93,83</b>

## V. INTERPRÉTATION DES RÉSULTATS

Des Tableaux IV et V on remarque que les réseaux utilisés ont montré de bons résultats de classification puisque les taux moyens de classification ont dépassé les 93%.

Les taux de classification des patients e0125 (ischémie sous-épicaudique), e0203, e403, e0615 et e0817 (onde Q de

nécrose) sont très élevés, s'approchant du 100%, cela s'interprète par le fait que ces derniers possèdent que des battements anormaux le long de leurs enregistrements, les paramètres définissant ces ischémies sont pertinents. Le patient e0509 a présenté des taux très importants car il possède l'onde T inversée le paramètre plus discriminant entre le cycle normal et anormal.

On remarque aussi que les réseaux LVQ présentent un taux de prédictibilité nul pour les patients e0161 et e0303 de même pour la carte SOM-LVQ dans la cas e0133 cela s'explique par l'handicap des réseaux à la détection des cycles anormaux, contrairement au e0409 où tous ses battements ischémiques ont été localisés (Pr = 100%), pourtant les trois patients font partie de la même classe (lésion sous-endocardique) et cela est dû au paramètre de la fréquence cardiaque qui est plus important que celui des patients e0161 et e0303, on remarque aussi que dans certains cas que le taux de sensibilité est nul cela est dû à l'incapacité de détection non seulement des cycles anormaux là où ils devraient se trouver mais en plus d'avoir considéré les cycles normaux comme anormaux (conversion totale).

Pour les patients e0123, e0151 et e0163 les taux Tr des deux réseaux sont presque équivalents et spécialement celui de prédictibilité malgré que l'allure du cycle anormal se rapproche du normal.

Pour la détection de l'onde Q de nécrose et de l'ischémie sous-épicaudique les trois réseaux ont répondu parfaitement. Mais il est bien clair que le réseau LVQ est meilleur que la carte SOM-LVQ, notons les patients e0121, e01303 et e0161 ont présenté des taux de sensibilité et de prédictibilité nuls signifiant que le SOM-LVQ n'a pas été efficace pour la détection des battements anormaux et ceci s'explique que l'algorithme est hybride non supervisé puis supervisé et aussi marquons que certains battements normaux ont des caractéristiques quasiment identiques à ceux des battements anormaux chose qui a compliqué la distinction pour la carte SOM-LVQ.

Notons bien que la performance des réseaux compte surtout sur les taux de prédictibilité et de sensibilité que sur celui de classification car ce dernier calcule le taux global des vrais positifs et des vrais négatifs par rapport au nombre total des cycles, cela veut dire que lorsque les cycles normaux se détectent correctement cela influe sur le taux de reconnaissance final surtout que le nombre de battements normaux est toujours élevé que celui des anormaux, contrairement aux taux de prédictibilité et de sensibilité qui dépendent fortement des cycles anormaux, car une confusion de classe peut mener à de faibles résultats, donc la performance des réseaux se mesure surtout selon la sensibilité et la prédictibilité.

## VI. CONCLUSION

Cette étude entre dans le cadre de la classification semi-automatique pour une aide prédictive au diagnostic de l'IDM en pratique médicale quotidienne à la lecture et d'analyse des signaux ECG. Beaucoup d'algorithmes ont été proposés ces dernières années pour la détection et la classification des

battements cardiaques, d'autres sont proposés pour l'amélioration, des techniques de traitement et des méthodes d'extraction de paramètres. Notre estimons que l'objectif est atteint vu les bons résultats obtenus. Nous avons opté à mettre en œuvre deux classificateurs pour la reconnaissance automatique de l'IDM, ils s'adaptent bien avec ce type de problème, puisqu'ils ont amélioré les performances. Nos résultats très satisfaisants pour le but recherché et les conditions utilisées grâce aux précautions prises dans le choix des indicateurs discriminants des pathologies et de l'approche adoptée. On peut dire que l'apport des réseaux de neurones et des ondelettes était de taille sur le plan de précision, de vitesse de calcul et d'architecture puisque nous avons eu des Tr de 100%. Le choix de l'ondelette a influé sur nos résultats, car les paramètres extraits ont joué un rôle important dans la distinction de chaque anomalie. L'ondelette du chapeau mexicain permet de bien déterminer les extrémités de chaque onde du signal ECG, elle possède des propriétés telle que sa forme qui ressemble à celle de l'onde électrocardiographique, sa transformée qui est continue, sa symétrie qui permet de localiser les frontières, sa reconstruction exacte et ses bonnes propriétés de localisation temporelle.

Dans l'article [11] utilisant les 12 dérivations où seulement l'élévation du segment ST a été prise comme paramètre d'entrée du réseau, le taux de reconnaissance est de 88.6%. Dans un traitement d'image du tissu myocardique par la TO [16] dans le but d'identifier un cancer du myocarde, le taux obtenu est de 96% en utilisant la statistique du test-t. Les auteurs laissent entendre un avenir propice dans cette voie.

La base de données utilisée dans notre cas n'est pas à 12 dérivations par conséquent nous ne pouvons pas comparer nos résultats à ceux des travaux étudiés. Notre base d'apprentissage est plus variée du point de vue valeurs pertinentes : 8 paramètres d'entrée pour chaque enregistrement ECG mais avec une dérivation chacun.

Notons que nos classificateurs peuvent être élargis à d'autres classes telles que la reconnaissance d'autres anomalies ischémiques ou pathologies cardiaques autres que l'IDM.

## BIBLIOGRAPHIE

- [1] Dr R. Lepage, "Détection et analyse de l'onde P d'un électrocardiogramme: application au dépistage de la fibrillation auriculaire", *Thèse de Doctorat*, Université de Bretagne Occidentale.2003.
- [2] R. Silipo and C. Marchesi, "Artificial Neural Network for Automatic ECG Analysis", *IEEE Transactions on Signal Processing*, Vol. n°5, May 1998.
- [3] Dr K. Perlemuter, Pr.G. Montalescot, Pr.J.P. Bassand, "Infarctus du myocarde".
- [4] L. Senhadji, "Approche multirésolution pour l'analyse des signaux non stationnaires", *Thèse de Doctorat*, Université de Rennes 1.
- [5] S. Kuklin , A.Dzizinku. & all, "Continuous Wavelet Analysis: A new method for studying nonstationary oscillations in the cardiac rhythm", *Somer : Human Physiology*, volume 32, number 1,January 2006,pp. 116-126(6); Publisher : Maik, Nau
- [6] S. Palreddy, J. Tompkins, Y Hen Hu, "Customization of ECG beat classifiers developed using SOM and LVQ", *IEEE*, 1995.
- [7] S. Papadimitriou, S. Mavroudi, and all, "Ischemia detection with a self-organizing map supplemented by supervised learning", *IEEE Transactions on Neural Network*, vol 12, NO3, May 2001.

- [8] Y. Hen Hu, S. Palreddy, J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach", *IEEE Transactions on Biomedical Engineering*, vol 44, NO 9, September 1997.
- [9] E.A Fernandez, P. Presedo and S. Barro, "An ECG Ischemic Detection System Based on Self-Organizing Map and a Sigmoid Function Pre-Processing Stage", *LNAI 1620*, pp.207- 216, 1999.
- [10] T. Stamkopoulos, M. Strintzis, & all, "One Lead Ischemia Using a New Backpropagation Algorithm and the European ST-T Database", *IEEE Computers in Cardiology*, 1992.
- [11] L. Reinhardt, T. Dimitrios, J. Montonen, "Localisation of myocardial infraction based on Learning Vector Quantization Networks applied to ST elevations of the 12 lead ECG", *Annals of Noninvasive Electrocardiology*, vol 2. N° 4, p 331-337, October 1997.
- [12] L.G. Gamero, J. Vilo and F. Palasios, "Wavelet Transform analysis of heart rate variability during myocardial ischemia", *Journal: Medical and Biological Engineering and Computing* 2001.
- [13] P.G. Lemarié, "Ondelettes à localisation exponentielle", *Journal de Math. Pures et Appl*, vol 67, p.227-236, 1988.
- [14] T. Kohonen, "The self organizing map". 2<sup>nd</sup> ed., Soringer-verlag, Berlin 1997.
- [15] T. Kohonen, "The Self organising map". *Proceedings of the IEEE*, vol 60, N° 3, April 1990.
- [16] A. Mojsilovic, M. V. Popovic and All, "Wavelet Image Extension for Analysis and Classification of Infarcted Myocardial Tissue". *IEEE Transactions on Biomedical Engineering*, vol. 44, N° 9, Sept 1997.

# MACHINE LEARNING FOR MALWARE DETECTION USING API CALLS

Mohamed Brahimi<sup>1</sup>, Abdelouahab Moussaoui<sup>2</sup>

<sup>1</sup> Computer science department, Mohamed Khider University, BISKRA 7000, Algeria  
[m\\_brahimi@esi.dz](mailto:m_brahimi@esi.dz)

<sup>2</sup> Computer science department, Ferhat Abbas University, SETIF 19000, Algeria  
[moussaoui.abdel@gmail.com](mailto:moussaoui.abdel@gmail.com)

**Abstract**—since the emergence of the first malware detection system based on the machine learning model proposed by schultz in 2001, several studies have demonstrated the efficiency of the machine learning tools in the fight against malware. These systems exploit the available data on previous attacks in order to achieve a more intelligent detection method. The majority of these systems is based on supervised learning. This limits their adaptations in dynamic environments (computer viruses environment) because they are not flexible after the training phase. Moreover, the training of a model requires a large number of labeled programs as training set. In this paper, we propose a malware detection system based on an incremental process of data mining which will be able to optimize the number of training examples while reducing the cost of labeling. A considerable improvement is made in the process of detecting malware through our system based on a new active and incremental learning architecture.

**Index Terms**— Malware detection, machine learning, supervised learning, active learning, incremental learning, support vector machine.

## I. INTRODUCTION

There are two ways to build a malware detector: manually or generating it automatically. In the first approach, we need to know what to detect and how to detect [1]. Whereas, the second one represents the data mining based approach, it doesn't need anything except data and the desired objectives. Because, malware detection is classification problem (malware or benign), the majority of these detectors is based on supervised learning.

The most part of the computer viruses systems based on supervised learning (VDS-DM: Viruses Detection System based on Data Mining) [2,3,4,5,6,7] have chosen a linear generic architecture (Fig. 1).

Using this architecture, after one training phase and testing phase, the resulting VDS-DM has a fixed character, thus limits its efficiency in a dynamic environment (the environment of the malware).

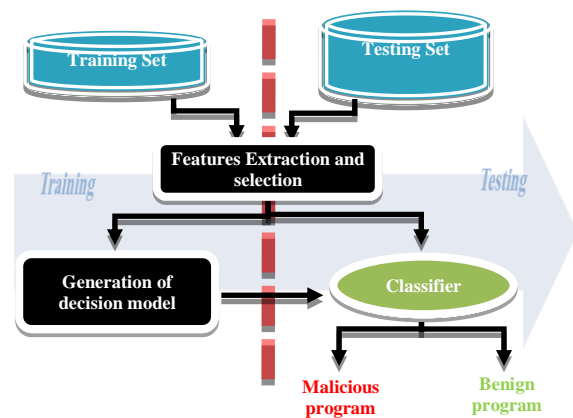


Fig.1. Supervised learning for malware detection.

Major limitations are summarized in the two following points:

- ✓ The non-evolutionary character of the VDS-DM: elaborating a VDS-DM by this approach hinders the acquisition of new labeled examples (analyzed programs) and thus the interaction of a VDS-DM with a security expert is infeasible. Thereby, the VDS-DM is deprived by the capacity of evolution, in contradiction with viruses' field which has a daily fast evolution [8,9].
- ✓ The necessity for a large number of labeled programs: the VDS-DM approach needs, as many as possible, labeled programs used as training examples. It is known that programs labeling is a precious procedure, because it needs an expert intervention, the availability of the resources to analyze and a considerable time for analysis [10,11].

The above mentioned limitations have encouraged us to improve the VDS-DM approach. Our goal is to make the VDS-DM evolutionary and also to optimize the number of examples to label and therefore to deal with the second limitation imposed by labeling. In the rest of article, we expose our methodology in order to building this evolutionary VDS-DM.

## II. FEATURES EXTRACTION AND SELECTION

### A. Feature extraction

In our work, we have primarily relied on features of program automatically generated from a set of labeled programs. We use n-gram extracted from the string of API system calls of the program at runtime (dynamic analysis). To do so, sliding window of size N is passed over the sequence of API calls [12] of each program in the training set. The content of this window will be used each time to produce several features with different lengths.

### B. Features selection

If all the features extracted from the training set are used, the number of these features risks to become too big and thus increase the computing time and degrade the accuracy. Moreover, the program features appearing in the majority of the programs are not able to make the difference because the most of the programs contains these types of features [7,13]. In our study, a feature is selected if its information gain (IG) exceeds a given threshold. The information gain (IG) of a program feature represents its degree of discrimination (between malicious and benign). IG is calculated using the following formula [14,13].

$$IG(X) = \sum_{x \in \{0,1\}} \sum_{C \in \{C_v, C_b\}} P(X=x \cap C=c) \cdot \log_2 \frac{P(X=x \cap C=c)}{P(X=x) \cdot P(C=c)} \quad (1)$$

Where, X is variable indicates the existence of feature and C indicates the class (C<sub>v</sub>:malware, C<sub>b</sub>:benign).

## III. PROGRAMS REPRESENTATION

After selection of features set, programs are represented as vectors of features. This vectors represents the projection of program in the space of  $n$  features:  $\{0,1\}^n$ . The value 1 (0) in vector indicates the presence (absence) of a feature in sequence of API system calls generated at runtime. So far, we can represent any program like vector. This representation, which is called *Vector space model*, is compatible with machine learning algorithms used in the rest of article.

## IV. ACTIVE AND INCREMENTAL LEARNING

Our main contribution is to make the VDS-DM incremental on the basis of the incremental and active learning, which offers:

- ✓ The possibility to update the VDS-DM with the acquisition of new labeled programs and especially new malwares.
- ✓ The minimization of labeling cost by reducing the number of training examples (Examples must be labeled).

In supervised learning, the entire training set must be available from the beginning. However, in incremental learning can produce a decision model from a limited number of training examples, while leaving the possibility of acquiring new labeled examples [15,16]. Also, in supervised learning, the learner (model M) is a passive entity that does not participate in training. On the contrary, in active learning the learner can influence the choice of training examples by selecting examples which improves the model accuracy [17].

We try to take advantage of the two learning techniques mentioned above (incremental and active), that is why our proposed algorithm is noted active and incremental learning. We based on the support vector machine (SVM) in order to produce the classification model M. Therefore, the active/incremental learning should be implemented on SVM.

The (Fig. 2) is overview of the proposed VDS-DM. The algorithm execution in an infinite loop allows immediate acquisition of examples. Therefore, the algorithm starts at each reception of examples incrementing the model M (Fig. 2). In the following sections we detail the three algorithms: Training of SVM, Incremental Selection and Active Selection.

---

### Algorithm 1 Active and incremental Learning.

---

```

1:  $M_0 \leftarrow \text{TrainingSVM}(E_0)$  ▷ initialisation of learning.
2:  $i \leftarrow 1$ 
3: while True do ▷ Active and incremental Learning loop.
4:   if (New set NE of unlabeled examples is available) then
5:      $E_1 \leftarrow \text{IncrementalSelection}(E_{(i-1)}, M_{(i-1)})$ 
6:      $E_2^* \leftarrow \text{ActiveSelection}(NE, M_{(i-1)})$ 
7:      $E_2 \leftarrow \text{LabellingBySecurityExpert}(E_2^*)$ 
8:      $E_i \leftarrow E_1 \cup E_2$ 
9:      $M_i \leftarrow \text{TrainingSVM}(E_i)$ 
10:     $i \leftarrow i + 1$ 
11:   end if
12: end while

```

---

Fig. 2.Active and Incremental Learning

### A. Training SVM

This is a standard training algorithm of SVM which takes as input a set of labeled examples and outputs a classification model M. Support Vector Machine (SVM) is a binary classification model. The basic principle of SVM is to find a hyperplane that best separates the two classes. However, where the examples are not separable using hyperplane, the kernel trick is used to transform the problem in other space and find the hyperplane in that new space.

### B. Incremental Selection

This is a selection function for the examples already labeled in the previous iterations. Not all the old labeled examples are taken into account; however, the representatives who

have an influence on the formation of the model are selected.

In our approach, support vectors are used to be included in training set of the following iteration[15]. Additionally, the examples of revision (examples misclassified in previous iteration) and examples situated in the margin of SVM are included. In summary, the old already labeled examples selected for incrementing VDS-DM are:

$$y^* \left( \sum_{i=1}^l \alpha_i y_i K(x_i, x) + b \right) \leq 1, \dots, (2)$$

Where  $x, y$  are respectively features vector and label,  $K(*, *)$  is used Kernel,  $\{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$  are support vectors.

### C. Active Selection

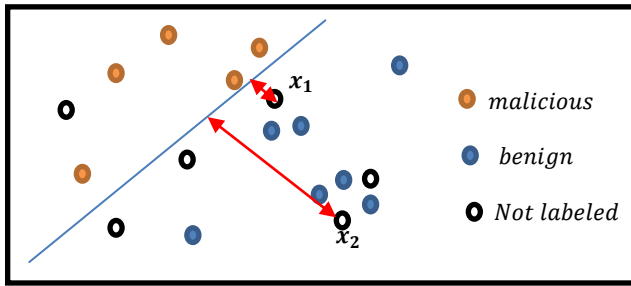


Fig. 3. Simple margin

Its objective is to ask an external entity (security expert) to label a minimum of examples without losing the accuracy of the resulting model. The participation of the model by selecting relevant examples in the learning process represents the major advantage of using active learning [16]. In the work of Tong et al. [16], it was shown that the examples that are close to the decision model (model of the previous iteration) strongly contribute to the formation of a SVM model for the next iteration. In this study we have used the method “Simple margin” based on the distance of the example to the SVM model.

Practically, this distance is calculated through the model as follows [17]:

$$\text{Distance}(x, M) = \left| \sum_{i=1}^l \alpha_i y_i K(x_i, x) + b \right|, \dots, (3)$$

Where  $M$  represents the classification model of the previous iteration. For instance in (Fig. 3), example  $x_1$  is preferable to  $x_2$  in order to be selected for labeling because it is near to boundary decision.

### D. Formation of the incremented model

After the formation of the two sets of labeled examples (Results sets of active and incremental selection) they are

merged into a single training set  $E_i$  of iteration  $i$ . Afterwards; the algorithm of training SVM is invoked in order to produce the incremented model  $M$  (Fig. 2). For the task of malware detection, the model is used until the forming of the new one. Otherwise, incrementing and detection (using the model) can be conducted in parallel.

## V. IMPLEMENTATION AND EXPERIMENTAL RESULTS

In this part we focused on the description of our system implementation by motivating our technical choices. Thus, we present a series of tests together with experimental results.

### A. Implementation and testing environment

All experiments were realized and performed on a computer Intel (R) Core (TM) i7-2670QM CPU @ 2.20GHz 2.20 GHz having 8 GB RAM.

The overall architecture of our implementation is illustrated in (Fig. 4). There are three phases of implementations.

In the first phase, the files which have the extension (\*.EXE) are executed in parallel to produce a detailed

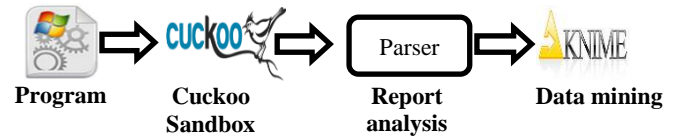


Fig. 4. Implementation of VDS-DM

behavioral report for each file. For automating this phase, we have used Cuckoo Sandbox and eight virtual machines in Virtual Box. Then in the second phase, these reports are analyzed by a parser to extract a features vector for each report by storing all vectors in a table saved in the (\*.CSV) file. Finally, in the last phase, we implement a data mining system able to classify the files into two classes, namely non-malicious or malicious. As data mining platform, we opted for the Knime benchmark due to its wealth of functionalities available in the form of nodes (> 900 nodes). In addition, a comparative study stated that Knime is the best among different famous data mining platforms [18]

### B. Data set of test

The tests were performed using the data set on the site (<http://vxheaven.org/>) used in previous works [14,19,20]. In our study, we have removed the files that were not executed because of the errors. The following table summarizes the dataset used in our test scenario.



<i>Programs</i>	<i>Number</i>
Malicious	1268
Benign	915
Total	2181

### C. Test methodology

In our test scenario, we have modified the cross validation in order to estimate the performance of our system. In our modified version, we added a T parameter which represents the rate of selected examples. Such as:

$$L = T * \text{number of examples from } (k - 1) \text{ groups}$$

In this case the model is trained by using L examples selected from the (k-1) groups. Once the model is ready, the remaining K<sup>th</sup> group is used to evaluate the model. In all our tests, we consider the following sampling strategies:

1. Random Sampling: The examples are selected randomly without any examples selection algorithm ( $0 < T < 1$ ).
2. Sampling based on our method: the examples are selected according to our algorithm ( $0 < T < 1$ ).
3. Standard supervised learning: all examples are selected in order to train the model ( $T=1$ ).

Finally, we have used (K=5) to calculate the following metrics:

<i>Accuracy</i>	<i>Number of programs classified correctly</i> <i>Total number of programs</i>
<i>True positive</i>	$\frac{\text{Number of detected malware}}{\text{Total number of programs}}$
<i>False positive (false alarms)</i>	$\frac{\text{Number of benign files classified malware}}{\text{Total number of programs}}$

### D. Experimental results

#### a. Comparison with random sampling.

We compare our algorithm with random sampling in order to show its usefulness. Indeed, if an algorithm does not improve the results compared to the random selection of examples, it cannot be used as a sampling strategy [21].

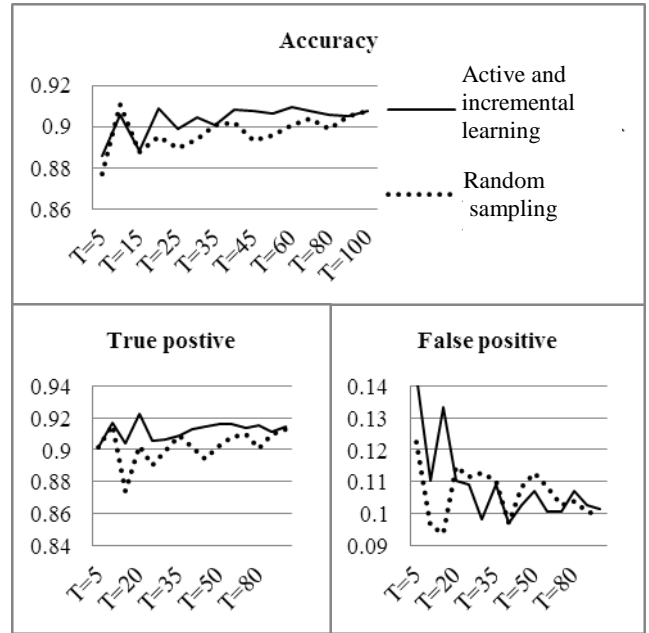


Fig. 5. Comparison between active and incremental sampling and random sampling.

#### ✓ Discussion.

Noticing the results illustrated in (Fig. 5), it is clear that the incremental/active sampling gives the best results. Indeed, the true positive (number of detected viruses) constantly exceeds that of the random learning. However, when the number of examples is not large enough ( $T < 20\%$ ) false positive is relatively elevated. Similarly, the accuracy is influenced by the false positive when the number of examples is small. The three curves converge before ( $T = 100$ ) and this strengthens the hypothesis that the reduction of samples does not seriously affect the results.

In addition, an incremental/active sampling strategy converges faster than a random strategy. However, in order to avoid false alarms a minimum number of examples is required. In other words, the examples selection based on the SVM margin can reduce the number of labeled examples without degrading the performance of the VDS-DM.

#### b. Influence of the number of training examples (N)

In this section, our goal is to empirically measure the impact of the examples number on the performance of the VDS-DM. To do so, we take the standard supervised learning (use all the available training examples) as a reference.

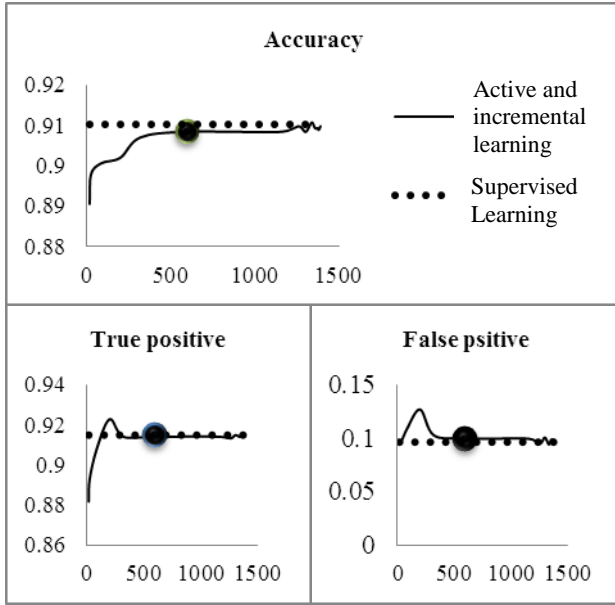


Fig. 6. Influence de N sur la performance de VDS-DM

✓ *Discussion.*

According to (Fig. 6) the three curves converge with a small number of samples (about 500). However, a very small number of examples of ( $N < 500$ ) may degrade performance. To summarize, we can achieve the standard supervised learning outcomes ( $N = 1500$  labeled examples) with a reduced number of samples ( $N = 500$ ). This reduction in 1000 labeled examples saves a lot of resources used in the labeling process.

*c. Influence of the distance R to the plane of SVM*

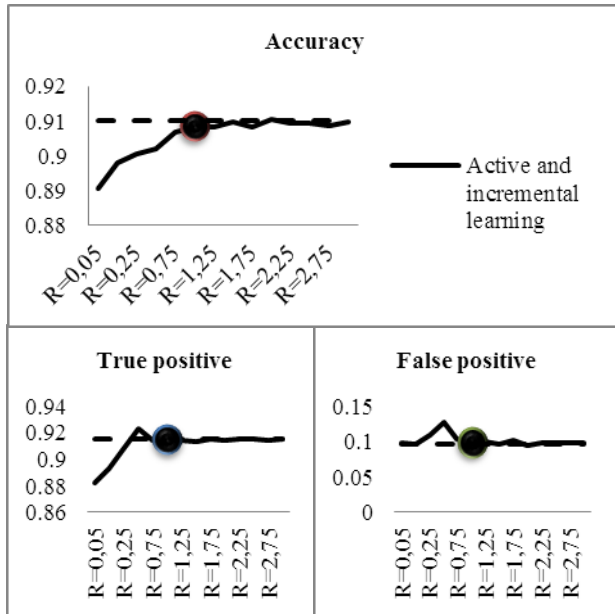


Fig. 7. Influence of the threshold R.

Our method is based on the distance to the SVM model. To calculate this distance of an example ( $x$ ) to a SVM model ( $M$ ) the formula (3) is used. We set an acceptance threshold  $R$  and we select the examples which have values of distances less than this threshold.

In this section we want to study empirically the influence of  $R$  on the performance of VDS-DM which led us to choose the optimal value of this parameter.

✓ *Discussion.*

In (Fig. 7.), the two curves converge while increasing  $R$ . When ( $R < 1$ ) the performance of the VDS-DM is degraded compared to the one of the supervised learning. However, when threshold exceeding ( $R = 1$ ) the VDS-DM exhibits a performance almost identical to that of the supervised learning. Knowing that the examples that verify  $Distance(x, M) \leq 1$  are the examples located in the margin of the SVM model. Therefore, the examples of margin, when we ask a security expert to label, improve the performance of the system better than the other examples that are outside margin.

VI. CONCLUSION AND FUTURE WORKS

Our investigations in malware detection based on data mining have enabled us to make contributions which are summarized in two points. In the first, we are interested in addressing the rapid evolution of computer malwares. Indeed, the explosion in the number of computer malwares is a real challenge for the detection techniques based on data mining. To address this problem, we have invested in other types of learning such as incremental and active learning. In the second point, we have treated the optimization of the interaction between the security expert and the malwares detection system. This optimization is manifested in the reduced number of programs analyzed by the expert (labeled program) for training the system to classify program.

In the future studies, we will try to include other types of features to enhance program representation. In other words, the combination of several visions can improve the accuracy of detecting malwares. In addition, the use of our approach in real-world scenarios or with a data set larger than ours can provide more concrete test results.

VII. REFERENCES

- [1] Marcus A. Maloof , *Machine Learning and Data Mining for Computer Security: Methods and Applications (Advanced Information and Knowledge Processing)*, Springer London Ltd, Ed., 2006.
- [2] Matthew G. Schultz, Eleazar Eskin, and Erez Zadok , "Data Mining Methods for Detection of New Malicious Executables," in *Proceedings of the 2001 IEEE Symposium on Security and Privacy*, Oakland, CA, 2001, pp. 38-49.

- [3] Jau-Hwang Wang and Peter S. Deng, "Virus detection using data mining techniques," in *Security Technology, 2003. Proceedings. IEEE 37th Annual 2003 International Carnahan Conference on*, 2003, pp. 71-76.
- [4] Jeremy Z. Kolter and Marcus A. Maloof, "Learning to detect malicious executables in the wild," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, Seattle, WA, USA, 2004a, pp. 470-478.
- [5] Asaf Shabtai, Robert Moskovitch, Yuval Elovici, and Chanan Glezer, "Detection of malicious code by applying machine learning classifiers on static features: A state-of-the-art survey," *Information Security Technical Report*, vol. 14, no. 1, pp. 16-29, February 2009.
- [6] Ashkan Sami et al., "Malware detection based on mining API calls," in *Proceedings of the 2010 ACM Symposium on Applied Computing*, New York, NY, USA, 2010, pp. 1020-1025.
- [7] Veelasha Moonsamy, Ronghua Tian, and Lynn Batten, "Feature reduction to speed up malware classification," in *Proceedings of the 16th Nordic conference on Information Security Technology for Applications*, vol. 7161, Tallinn, Estonia, 2012, pp. 176-188.
- [8] Jeremy Z. Kolter and Marcus A. Maloof, "Learning to detect malicious executables in the wild," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, Seattle, WA, USA, 2004, pp. 470-478.
- [9] Nir Nissim, Robert Moskovitch, Lior Rokach, and Yuval Elovici, "Novel active learning methods for enhanced PC malware detection in windows OS," *Expert Systems with Applications*, no. 0, March 2014.
- [10] Robert Moskovitch, Nir Nissim, and Yuval Elovici, "Malicious Code Detection Using Active Learning," *Privacy, Security, and Trust in KDD*, vol. 5456, pp. 74-91, 2009c.
- [11] Igor Santos, Javier Nieves, and Pablo G. Bringas, "Semi-supervised Learning for Unknown Malware Detection," in *International Symposium on Distributed Computing and Artificial Intelligence*, vol. 91, 2011, pp. 415-422.
- [12] Zahra Salehi, Mahboobeh Ghiasi, and Ashkan Sami, "A miner for malware detection based on API function calls and their arguments," in *Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on*, Shiraz, Fars, 2012, pp. 563-568.
- [13] Qingshan Jiang, Xinxing Zhao, and Kai Huang, "A feature selection method for malware detection," in *International Conference on Information and Automation*, Shenzhen, China, 2011, pp. 890-895.
- [14] Wei Wang, Pengtao Zhang, Ying Tan, and Xingui He, "A hierarchical artificial immune model for virus detection," in *International Conference on Computational Intelligence and Security*, 2009, pp. 1-5.
- [15] Stefan Ruping, "Incremental Learning with Support Vector Machines," in *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, 2001, pp. 641-642.
- [16] Simon Tong and Daphne Koller, "Support vector machine active learning with applications to text classification," *J. Mach. Learn. Res.*, vol. 2, pp. 45-66, March 2002.
- [17] Bin-Bin Peng, Zheng-Xing Sun, and Xiao-Gan Xu, "system, SVM-based incremental active learning for user adaptation for online graphics recognition," in *Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on*, vol. 3, 2002, pp. 1379-1386.
- [18] Xiaojun Chen, Yunming Ye, Graham Williams, and Xiaofei Xu, "A Survey of Open Source Data Mining Systems," in *Emerging Technologies in Knowledge Discovery and Data Mining.*: Springer Berlin Heidelberg, 2007, vol. 4819, pp. 3-14.
- [19] Wei Wang, Pengtao Zhang, and Ying Tan, "An Immune Concentration Based Virus Detection Approach Using Particle Swarm Optimization," *LNCS*, vol. 1, no. 6145, pp. 347-354, 2010.
- [20] Rui Chao and Ying Tan, "A Virus Detection System Based on Artificial Immune System," in *International Conference on Computational Intelligence and Security*, 2009, pp. 6-10.
- [21] Ran Wang and Sam Kwong, "Sample selection based on maximum entropy for support vector machines," in *Machine Learning and Cybernetics (ICMLC)*, vol. 3, Qingdao, 2010, pp. 1390-1395.

# Identification Automatique d'Objets Microscopiques de Formes Complexes. Application aux Grains de Pollen

Ouiza Boukendour<sup>1</sup>, Fatiha Houdène<sup>2</sup>, F. Boudjema<sup>1</sup> Moussa Diaf<sup>1</sup>

1. Département Automatique, FGEI, Université M. Mammeri, Tizi-Ouzou,

2. Département Maintenance, FSI, Université M. Bouguara, Boumerdès,

[ouizaboukendour@yahoo.fr](mailto:ouizaboukendour@yahoo.fr), [boudjemaadj@yahoo.fr](mailto:boudjemaadj@yahoo.fr)

, [mousmoeses@gmail.com](mailto:mousmoeses@gmail.com)

**Résumé.** Dans ce papier nous nous intéressons à l'identification automatique d'objets microscopiques de formes complexes et plus particulièrement les grains de pollen et ce, en utilisant, d'abord, des descripteurs des formes; de texture et leur combinaison. Nous utilisons, ensuite, les moments invariants de Zernike. Après l'acquisition d'images de six cent échantillons de six différentes espèces de grains de pollen, nous avons procédé aux opérations de prétraitement, de normalisation et de sélection des descripteurs en utilisant l'analyse en composantes principales et une méthode mixte de classification automatique. Les résultats de l'identification sont présentés dans plusieurs tableaux montrant les différentes confusions et les taux de reconnaissance. Une étude comparative a montré que l'utilisation des moments de Zernike a donné des résultats très satisfaisants.

**Mots clés :** Reconnaissance de formes, pollen, moments de Zernike, paramètres de forme, paramètres de texture, combinaison de paramètres, classification automatique.

## I. INTRODUCTION

Certains objets présentant des formes complexes sont difficiles à reconnaître. En effet, pour cette catégorie d'objets, l'application classique des méthodes habituellement utilisées peuvent s'avérer insuffisantes. De plus, le choix des descripteurs ainsi que des méthodes utilisées doit conduire à un système de reconnaissance performant. Dans ce papier nous nous intéressons à l'identification de grains de pollen dont les images sont effectivement complexes. Notons que, dans ce cadre et dans des objectifs divers, quelques projets d'envergure continentale ont été financés pour, par exemple, l'étude rétrospective des pollens fossiles en tant qu'indices permettant de reconstituer les paléoenvironnements et paléoclimats d'un passé ainsi que l'évolution des stocks de carbone sur notre planète, l'aéro-palynologie pour analyser la présence dans l'air de différents types de pollens dans un objectif médical, la méliissopalynologie pour l'analyse pollinique d'un miel pour déterminer sa pureté ou son origine florale et géographique etc.[1],[2],[3],[4],[5]. Cependant, d'après les spécialistes du pollen que nous avons contactés, les logiciels réalisés ne donnent pas de bons résultats pour le même pollen évoluant dans des milieux climatiques différents. En effet, un logiciel identifiant une espèce de pollen dans un pays donné ne sera pas automatiquement fonctionnel pour cette même espèce évoluant dans un autre pays qui présente des conditions environnementales différentes. Ceci, nous a amenés à nous intéresser à l'identification automatique de grains de pollen en collaboration avec des chercheurs en agronomie spécialistes de ce domaine. Ainsi, dans ce présent

papier, nous proposons une nouvelle méthode basée sur d'autres descripteurs et une autre méthode de classification automatique en vue d'améliorer la qualité des résultats que nous déjà obtenus [6]. Rappelons brièvement que, en reconnaissance de formes, l'objectif est d'associer une étiquette à un objet à partir d'une information extraite de l'image de cet objet. De manière générale, le processus de reconnaissance s'effectue en quatre phases. La première se résume en la localisation de l'objet et l'acquisition de son image à l'aide du capteur approprié dont le manque de qualité nécessiterait un prétraitement de cette image. En seconde phase, il s'agit de la représentation des informations acquises et l'extraction de primitives ou de paramètres pertinents et leur normalisation en utilisant, par exemple, l'Analyse en Composante Principales (ACP), leur réduction pour rendre les données plus classifiables. Ces paramètres qui peuvent être de forme, de couleur ou de texture sont représentés dans un espace multidimensionnel. Il s'agit, ensuite, du partitionnement de cet espace de représentation en différentes classes. Ceci constitue la troisième ou phase d'apprentissage. L'assignation d'un objet à la classe appropriée ou décision constitue la phase finale du processus de reconnaissance. Pour l'identification des grains de pollen, nous avons suivi ces différentes étapes. Ainsi, dans la prochaine section, nous présentons les espèces de pollens auxquels nous nous sommes intéressés. Dans la section 3, nous montrons les résultats obtenus en utilisant les paramètres de formes, de texture, les paramètres de forme et texture combinés. En section 4, l'identification automatique est effectuée en introduisant les moments de Zernike [7]. Une étude comparative les résultats obtenus est présentée dans la section 5. Nous concluons dans la dernière section.

## I. ACQUISITION D'IMAGES ET PRETRAITEMENT

Nous avons prélevé du pollen de six espèces de plantes cueillies dans la région de Kabylie. Ces plantes sont le pied-de-chèvre (oxalis pes-capera), le chardon laiteux (galactite tomentosa), la mauve (lavaera), le chardon d'Espagne (scolymus hispanicus), la sulla (hedysarum) et le mélinet (cerinthe major). Pour l'acquisition des images de ces pollens, nous avons utilisé le microscope électronique à balayage environnemental (Philips XL 30 ESEM) à filament de tungstène du laboratoire de microscopie de l'université Mouloud Mammeri, Tizi-Ouzou. Ainsi 100 images en niveaux de gris de chacune des six espèces ont été acquises. La figure 1 montre un exemple de ces images. A ces images qui

différent en formes et en textures et positionnées selon une orientation quelconque, nous avons appliqué le filtre médian, la binarisation selon la méthode d'Otsu [8] et le suivi de contours (fig.2).

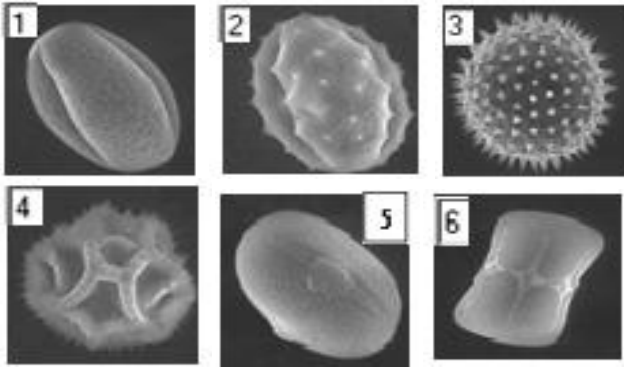


Fig.1 Images de pollen des six espèces de fleurs avec un grossissement de 1000 pour 1,2,4,5,6 et de 500 pour la 3: 1. Oxalis pes-caprea, 2. Galactite tomentosa, 3.Lavaera, 4.Scolymus hispanicus, 5.Hedysarum, 6.Cerinthe major

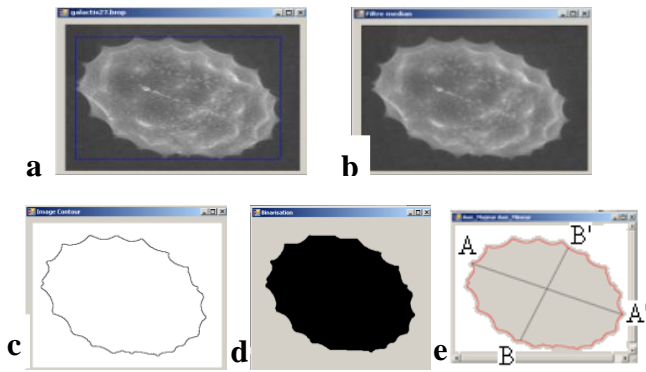


Fig. 2 Prétraitement ;a. Images initiale, b. Filtre médian, c. Binarisation, d. Contour, e. image paramètres calculés

### 3. IDENTIFICATION A BASE DES PARAMETRES DE FORME ET DE TEXTURE

Les images acquises étant en niveaux de gris, seuls les paramètres de formes et de texture sont pris en considération. Les paramètres de formes calculés dans notre application sont, en nombre de pixels, le périmètre P, la surface S, la compacité C qui est le rapport entre  $4\pi S$  et le carré de P, l'axe majeur AA' et l'axe mineur BB' (Fig.1.e). Les paramètres de textures que nous avons calculés sont l'Homogénéité (H), l'Homogénéité Locale (HL), l'Entropie (Ent), le Contraste (Cont), la Corrélation (Corr), l'Uniformité (Unif) et la Directivité (Dir). Le calcul de ces paramètres est basé sur la matrice de cooccurrence des niveaux de gris en utilisant les expressions suivantes :

$$MC_i(a,b) = \begin{cases} 1 & \text{si } MC_i(a,b) \neq 0 \\ 0 & \text{si non} \end{cases} \quad (1)$$

$$m_x = \frac{1}{N_C} \sum_a \sum_b a MC_i(a,b) \quad (2)$$

$$Dir = \frac{1}{N_C} \sum_a MC_i(a,a) \quad (3)$$

$$H = \frac{1}{N_C^2} \sum_a \sum_b (MC_i(a,b))^2 \quad (4)$$

$$HL = \frac{1}{N_C} \sum_a \sum_b \frac{1}{1+(a+b)^2} MC_i(a,b) \quad (5)$$

$$Ent = 1 - \frac{1}{N_C \ln(N_C)} \sum_a \sum_b MC_i(a,b) \ln(MC_i(a,b)) \quad (6)$$

$$Cont = \frac{1}{N_C(L-1)^2} \sum_{k=0}^{L-1} k^2 \sum_{|a-b|=k} MC_i(a,b) \quad (7)$$

$$Corr = \frac{1}{N_C \cdot \sigma_x \cdot \sigma_y} \left| \sum_a \sum_b (a - m_x)(b - m_y) MC_i(a,b) \right| \quad (8)$$

$$Unif = \frac{1}{N_C^2} \sum_a MC_i^2(a,a) \quad (9)$$

$$Dir = \frac{1}{N_C} \sum_a MC_i(a,a) \quad (10)$$

$$Corr = \frac{1}{N_C \cdot \sigma_x \cdot \sigma_y} \left| \sum_a \sum_b (a - m_x)(b - m_y) MC_i(a,b) \right| \quad (11)$$

$$(Unif = \frac{1}{N_C^2} \sum_a MC_i^2(a,a) \dots \quad (12)$$

$$Dir = \frac{1}{N_C} \sum_a MC_i(a,a) \dots \quad (13)$$

Ainsi, pour réduire le nombre de paramètres, nous avons appliqué l'Analyse en Composantes Principales (ACP) qui nous a permis de ne conserver que les composantes parmi les descripteurs qui apportent plus de 95% de l'information. Aux tableaux résultant, nous avons appliqué une classification automatique pour former les classes en présence. Notons qu'il est possible, pour nous, de former ces classes, nous-mêmes, puisque nous connaissons chacune des espèces des pollens qui les forment. Cependant, de manière générale, l'utilisation d'une méthode non supervisée de classification est mieux indiquée puisque, dans le cas particulier du pollen, des espèces peuvent être mélangées à d'autres lors de la cueillette des fleurs. Pour procéder à cette opération de manière automatique, nous pouvons choisir une méthode non supervisée et non paramétrique parmi la panoplie de méthodes

de classification automatique qui existent. Dans notre cas, nous avons appliqué la methode des k-means [9],[10].

Après cette opération de classification automatique, nous avons procédé à la phase de reconnaissance où il s'agit d'affecter une forme inconnue à la classe appropriée. Cette phase de reconnaissance d'un grain inconnu à l'une des six classes est effectuée en calculant la distance euclidienne entre le vecteur paramètres du grain inconnu et le vecteur moyenne de chacune des classes. Le grain est affecté à la classe qui correspond à la distance la plus petite. Pour évaluer les performances de la methode utilisée, nous avons, tout d'abord, utilisé les éléments de la base d'apprentissage. Le taux de reconnaissance est de 100 %. Nous avons, ensuite, utilisé les 266 images restantes de la base de données. Les résultats de reconnaissance sont donnés dans les tableaux 1 et 2, 3 et 4, 5 et 6. Ces tableaux représentent les matrices de confusion et les taux de reconnaissance. Notons que les taux de reconnaissance (TR) et de rejet (TJ) pour chacune des espèces de pollen sont évalués de la façon suivante :

$$T_R = \frac{\text{nombre d'entités reconnues}}{\text{nombre total d'entités présentées au classifieur}} \quad (14)$$

$$T_J = \frac{\text{nombre d'entités rejetées}}{\text{Nombre total d'entités présentées au classifieur}} \quad (15)$$

TABLEAU 1 MATRICE DE CONFUSION, PARAMETRES DE FORME

	1	2	3	4	5	6	R	Total
1	38	0	0	0	1	0	7	46
2	0	40	0	3	2	0	2	47
3	0	2	35	0	0	1	3	40
4	0	0	0	33	5	0	5	46
5	0	1	0	8	34	0	2	45
6	0	0	0	0	0	40	2	42

TABLEAU 2 TAUX DE RECONNAISSANCE, PARAMETRES DE FORME

TR%	TJ%	TJ Total(%)
82.6	15.21	
85.10	4.25	5.21
87.5	0.07	
71.73	10.86	
75.55	0.44	
95.23	4.76	

TABLEAU 3 MATRICE DE CONFUSION, PARAMETRES DE TEXTURE

	1	2	3	4	5	6	R	Total
1	41	0	0	0	0	0	5	46
2	0	38	2	3	0	3	4	47
3	0	2	30	0	0	1	7	40
4	0	0	0	40	5	0	3	46
5	0	0	0	8	39	0	2	45
6	0	0	0	0	0	40	2	42

TABLEAU 4 TAUX DE RECONNAISSANCE, PARAMETRES DE TEXTURE

TR%	TJ%	TR Total(%)	TJ Total(%)
89.13	10.86		
80.85	8.5	85.59	8.83
75.00	17.5		
86.95	6.52		
86.66	4.44		
85	4.00		

TABLEAU 5 MATRICE DE CONFUSION, PARAMETRES COMBINES

	1	2	3	4	5	6	R	Total
1	43	0	0	0	0	0	3	46
2	0	42	0	1	0	0	4	47
3	0	0	40	0	0	0	0	40
4	0	0	0	41	2	0	3	46
5	0	0	0	4	40	0	1	45
6	0	0	0	0	0	40	2	42

TABLEAU 6 TAUX DE RECONNAISSANCE, PARAMETRES COMBINES

TR%	TJ%	TR Total(%)	TJ Total(%)
93.47	6.52		
89.36	8.51	92.67	4.71
100	0		
89.13	6.52		
88.88	2.22		
95.23	4.76		

D'après les tableaux ci-dessus, nous remarquons que, en utilisant les paramètres de forme seuls, le taux de reconnaissance global obtenu est de 82.95% alors que celui de rejet est de 5,21%. En utilisant les paramètres de texture seuls, le taux de reconnaissance global obtenu est de 85,59% alors que celui de rejet est de 8,83%. Par rapport à la première approche utilisant les paramètres de forme, nous avons enregistré une amélioration de 6.64%. En combinant les paramètres de forme et de texture, le taux de reconnaissance est élevé à 92.67% et le taux de rejet est redescendu à 4.71%. L'apprentissage étant une étape importante dans tout processus de reconnaissance de formes, en augmentant le nombre d'échantillons de la base d'apprentissage, nous avons constaté une nette amélioration du taux global de reconnaissance ainsi qu'une nette régression de confusion entre les espèces dans les trois cas. Dans ce cas, le taux de reconnaissance global a été élevé 86,62% dans le cas de l'utilisation des paramètres de forme seuls alors qu'il était à 82,95%. Toujours en augmentant le nombre d'échantillons de la base d'apprentissage, dans le cas de l'utilisation des paramètres de texture seuls, le taux de reconnaissance global a

été élevé à 88,39% alors qu'il était à 85,59%. En utilisant les paramètres de forme et de texture réunis, le taux de reconnaissance global a été élevé à 94,49% alors qu'il était à 92,67% et une nette diminution des confusions entre espèces est observée. Ainsi, nous notons que, en augmentant la taille de la base d'apprentissage, le taux de reconnaissance s'améliore et la confusion et le taux de rejet diminuent dans tous les cas.

#### 4. IDENTIFICATION A BASE DES MOMENTS DE ZERNIKE

Introduit pour la première fois en vision par ordinateur par Teague [11], les moments de Zernike qui sont des descripteurs de forme ont prouvé leur supériorité sur les autres descripteurs comme ceux de Fourier ou de Hu [12]. Jusqu'à présent, de nombreux chercheurs se penchent encore sur ces descripteurs en vu d'introduire des améliorations dans la précision et pour optimiser leur temps de calcul [13],[14]. Ainsi, les fonctions complexes de Zernike constituent un ensemble de fonctions orthogonales définies sur le disque unité. Elles possèdent trois propriétés principales que sont l'orthogonalité, l'invariance en rotation et en translation et au changement d'échelle. Tout en préservant l'information contenue dans les images, ils réduisent aussi le bruit et la redondance dans les vecteurs des caractéristiques. Les moments de Zernike d'une image sont donc construits par projection sur ces fonctions. Ils sont orthogonaux complexes et définis sur un disque de rayon unité et basés sur une représentation polaire de la forme. Ces moments sont construits à partir de polynômes complexes  $V_{mn}$  qui forment un ensemble orthogonal complet défini sur le disque unité. Ainsi, la base polynomiale de Zernike, d'ordre  $n$  et de répétition  $m$ , est donnée par :

$$A_{mn} = \frac{m+1}{\pi} \int_x \int_y I(x,y) [V_{mn}(x,y)] dx dy \quad (16)$$

Où  $m$  et  $n$  définissent l'ordre du moment,  $V_{mn}$  est un polynôme complexe (Zernike 1934) dit polynômes de Zernike  $I(x, y)$  est le niveau de gris d'un pixel de l'image  $I$ . les polynômes de Zernike sont exprimés en coordonnées polaires

$$V_{mn}(x, y) = W(r; \theta) = R_{mn}(r) e^{-jn\theta} \quad (17)$$

Méthode de reconnaissance	Taux de Reconnaissance
Reconnaissance basée sur La forme	<b>82.95</b>
Reconnaissance basée sur la texture	<b>85.95</b>
Reconnaissance basée sur la forme + texture	<b>94.49</b>
Reconnaissance basé sur les moments de Zernike	<b>94.95</b>

Où  $R_{mn}$  forme un polynôme radial orthogonal

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s \frac{(m-s)!}{s! \left(\frac{m+|n|}{2}-s\right)! \left(\frac{m-|n|}{2}-s\right)!} r^{m-2s} \quad (18)$$

$R_{mn}(r)$  est e gale à 0 dans le cas ou  $m$  et  $n$  n'ont pas la même parité.

Pour identifier les grains de pollen, nous avons calculé les moments de Zernike sur leurs différentes images définies pour l'apprentissage et pour la reconnaissance comme dans le cas des descripteurs précédents. En effet, les moments géométriques ne sont pas des paramètres abstraits mais plutôt des paramètres explicites ou explicables suivant une signification géométrique de l'image. Chaque moment de l'image représente une caractérisation géométrique bien définie tels que la surface, le centre de masse, les symétries horizontale et verticale. Ainsi, un grain de pollen, par sa forme particulière et son contenu géométriquement riche par les détails des éléments qu'il contient (les ports, ornementation, pore, sillon, etc.), pourrait avoir une certaine prépondérance au niveau du vecteur Zernike. Le tableau 7 nous montre un exemple de valeurs des moments de Zernike. Le taux de reconnaissance global obtenu est de 94.95, bien supérieur à ceux obtenus avec les autres descripteurs comme le montre le tableau 8.

Tab9 : matrice de confusion, Moment de Zernike

	1	2	3	4	5	6	total
1	4	1	1	3	2	0	46
2	0	4	0	3	2	0	47
3	0	3	4	0	0	0	40
4	0	0	0	4	1	0	46
5	0	0	0	6	4	0	45
6	0	0	0	0	0	4	42

Tab14 : Taux

de reconnaissance

T <sub>rg</sub> (%)	TR(%)
95.65	
85.10	<b>94.95</b>
100	
97.82	
91.11	
100	

Fig3 Comparaison de toutes les méthodes utilisés

#### 5. CONCLUSION

Si dans certains domaines d'applications la reconnaissance de formes donne des résultats avec une très grande précision, dans d'autres où les objets à reconnaître présentent une grande variabilité et des formes complexes, le taux de reconnaissance est nettement réduit. Dans de tels cas, des taux de rejet qui correspondent à la décision de ne pas classer la forme sont souvent introduits. Dans notre application, le rejet a été

introduit pour éviter des confusions dans la prise de décision. Par ailleurs, pour avoir l'assurance d'une amélioration de la qualité de classification et du taux de reconnaissance plus élevés, il faut fusionner les paramètres de forme et de texture et utiliser une base de test plus grande. Le taux d'erreur le plus important est dû à la confusion entre les classes. Cela peut s'expliquer par la similarité structurelle élevée de ces deux classes et il est difficile de les distinguer même visuellement au microscope sur un certain nombre d'images. D'après les matrices de confusion, nous pouvons observer que nous obtenons de meilleurs résultats avec l'augmentation de nombre d'exemple de la base d'apprentissage. Ainsi, si le taux global de reconnaissance a été évalué à 94.49%.en utilisant les paramètres de forme combinés et augmentant la base d'apprentissage, ce taux de reconnaissance a été élevé à 94,95% en utilisant les descripteurs de Zernike, ce qui prouve l'efficacité de ces moments.

#### 6. REFERENCES

1. Pollen, projet européen lancé en 2006, <http://www.emse.fr/larotonde/pollen/Livret-pollen.pdf>
2. Tonkov, E. Marinova, Pollen and plant macrofossil analyses of radiocarbon dated mid-Holocene profiles from two subalpine lakes in the Rila Mountains, Bulgaria, Quaternary research 24.121-130(1985),University St. Kliment Ohridski, 8 Dragan Tsankov bd., Sofia1164, Bulgaria (2004).
3. Calderon M, Brandt T, Treatment of grass pollen allergy: focus on a standardized grass allergen extract – GRAZAX, Ther Clin Risk Manag.2008 Dec;4(6):1255-603. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
4. A. Maurizio, J. Louveax, Les méthodes de la terminologie en melissopalynologie, revue en paleobotane et palynologie, Rosenweg 9,3097 Liebefeld (Suisse), 3(1967)291-295) (1966).
5. National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov>.
6. Ouiza Boukendour, Takfarinas Chelli, Moussa Diaf, Zohra Zaabot, Reconnaissance automatique du pollen de trois espèces de pins à base de paramètres de formes en utilisant des réseaux bayésiens, Journées d'Etudes Algéro -Françaises de Doctorants en Signal-Images & Applications (JEAFFD'2012 », E.N. Polytechnique d'Alger, (2012).
7. Chong-Yaw Wee et Raveendran Paramesran. On the computational aspects of Zernike moments. Image Vision Comput., 25(6) :967–980 ( 2007).
8. N. Otsu, “A threshold selection method for grey level histograms”, IEEE, Trans. on Syst., Man and Cyber., vol. SMC-9, no.1, pp. 62-66, 1979.
9. Diaf M., Boudjemaa F. et Hammouche K., Optimisation par le Recuit Simulé, d'une Méthode Mixte de Classification Automatique, COSI'05 (Colloque sur l'Optimisation et les Systèmes d'Information), Béjaïa, 2005
10. M. Diaf and K.K Chopra, A Mixed Method for Automatic Classification, Journal of Studies in Informatics and Control. vol.3, n°1, pp.33-42( 1994)
11. M.R. Teague. Image analysis via the general theory of moments. J. J. Optical Soc. Am., 70(8) :920–930, August 1980.
12. L. Kotoulas et I. Andreadis. Real-time computation of Zernike moments. IEEE Transactions on Circuits and Systems for Video Technology, 15(6) :801–809 June (2005)
13. G. Badrinath, Naresh Kachhi, and Phalguni Gupta. Verification system robust to occlusion using low-order Zernike moments of palmprint subimages. Telecommunication Systems, 47 275–290 ( 2011).
14. P. Esther Rani and R. Shanmuga Lakshmi. Palmprint recognition system using zernike moments feature extraction. In Vinu V Das and R. Vijaykumar, editors, Information and Communication Technologies, volume 101 of Communications in Computer and Information Science, pages 449–454. Springer Berlin Heidelberg ( 2010).



# Réseaux Bayésiens appliqués à une fusion de données multi-capteurs en vue d'un contrôle intelligent d'un microclimat

Mazari Chahrazad Boussanou Tassadit, Diaf Moussa

Laboratoire Vision Artificielle et Automatique de Systèmes (LVAAS)

Département Automatique, FGEI, UMMTO

[ma.charazad@hotmail.com](mailto:ma.charazad@hotmail.com), [hbtassadit@hotmail.fr](mailto:hbtassadit@hotmail.fr), [diaf\\_moussa@yahoo.fr](mailto:diaf_moussa@yahoo.fr)

**Abstract**— Ce papier présente une méthode de fusion de données multi-capteurs pour un contrôle automatique du microclimat d'une serre agricole répondant le mieux aux exigences de l'olivier. La méthode proposée utilise les réseaux bayésiens pour fusionner des mesures issues des capteurs de température ambiante et du sol, de rayonnement et d'humidité dont l'objectif est de fournir aux modules de commande, des informations sûres et complètes concernant l'état du microclimat afin que l'intervention des actionneurs du genre ventilateurs, chauffages, humidificateurs et autres, soit rapide et efficace. Notons que la base de données réelle n'étant pas disponible, nous avons créé une base de données simulées. Des résultats de simulation sont présentés dans la dernière section pour montrer l'efficacité de l'approche adoptée même si les différentes caractéristiques sont supposées indépendantes entre elles.

**Keywords**— Réseaux bayésiens, microclimat, classification automatique, décision, apprentissage, Optimisation, serre agricole, fusion de données, multicapteurs.

## I. INTRODUCTION

La problématique de la fusion de données multi-capteurs peut prendre des chemins différents selon le contexte applicatif. Nous la retrouvons, en particulier, dans le domaine militaire, médical, la robotique etc. Quoique ces domaines sont différents, le principe de la fusion multi-capteurs reste approximativement le même pour la perception de l'environnement à partir de données issues de plusieurs sources [1][2] et [3] en vue d'un éventuel contrôle, de l'établissement d'un diagnostic ou autre. L'utilisation de plusieurs sources simultanément est importante car une seule source permet, dans la majorité des cas, de n'avoir qu'une décision partielle sur l'environnement analysé.

\*Dans ce papier, il s'agit de proposer une solution au problème de contrôle vu comme un problème de reconnaissance de formes, plus précisément, de classification, où les classes correspondent aux différents états du climat de la serre, et les formes représentent l'ensemble des observations ou mesures multi-capteurs des paramètres liés aux caractéristiques. Au niveau du système, les différents paramètres climatiques utilisés dans l'analyse du microclimat, tels que la température (°C), l'humidité (%) et le rayonnement (Lux) sont transformés en données numériques à partir des capteurs, et transmis vers une station de contrôle qui assure l'acquisition, le traitement, et l'analyse des données. Latechnique de reconnaissance utilisée au niveau du système

de décision effectuée après chaque acquisition, la classification et la séparation des données en 54 classes différentes. Une suite d'acquisitions pourrait être envisagée plusieurs fois par jour, sous des conditions prédéfinies. Un module d'apprentissage, supervisé par un expert, permet de collecter de manière continue les paramètres relatifs aux différents états du microclimat, pour la mise en œuvre d'une base de connaissance complète.

## II. RESEAUX BAYESIENS

Les réseaux bayésiens sont nés du besoin d'introduire la notion d'incertitudes des mesures en utilisant les probabilités.

Ce sont des modèles qui marient la théorie des graphes à celle des probabilités. Ils sont basés sur l'apprentissage et utilisés pour l'aide à la décision, le diagnostic ou contrôle de systèmes complexes etc.

L'intérêt particulier des réseaux bayésiens est de tenir compte simultanément, des connaissances *a priori* d'experts représentées par le graphe et de l'information contenue dans les données.

Un réseau bayésien est représenté par un graphe dirigé sans cycles, composé de nœuds reliés par des arcs représentant une relation de cause (parents) à effet (enfants) (Fig.1).

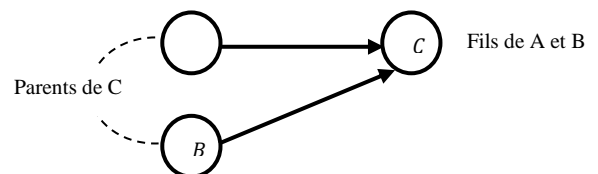


Fig.1 Réseau bayésien simple

Il est, par définition, un graphe orienté acyclique  $G = (V, E, \Omega)$  où  $V$  est l'ensemble des nœuds représentant les variables du domaine considéré,  $E$ , l'ensemble des arcs représentant les relations qui existent entre les différentes variables et  $\Omega = P(x/p(x))$ , l'ensemble des distributions de probabilités de chaque nœud  $x$  conditionnellement à ses parents immédiats  $p(x)$  dans le réseau  $G$ .

Notons que si tous les nœuds du réseau sont deux à deux conditionnellement indépendants de leurs parents, il serait possible, alors, de factoriser la probabilité jointe par le produit

des distributions conditionnelles sur l'ensemble des nœuds  $V$  comme suit :

$$P(V) = \prod_{x \in V} P(x/p(x)) \quad (1)$$

où  $P(V)$  est la probabilité jointe globale et  $P(x/p(x))$  la probabilité conditionnelle locale.

Dans le cas d'un problème de classification, la structure du réseau est simplifiée par une hypothèse dite l'hypothèse naïve de Bayes qui considère que les caractéristiques  $x_i$  sont deux à deux indépendantes conditionnellement à la variable classe  $C_j$ . Pour ce genre de réseaux, on retrouve deux niveaux de nœuds, le premier niveau est le niveau du nœud parent « classe » considéré comme quantité non observée et qu'il faut déterminer. Le deuxième est celui des nœuds enfants « caractéristiques », considérés comme étant des variables observées caractérisant l'élément étudié.

Les deux composantes essentielles pour arriver à utiliser un réseau bayésien sont l'apprentissage et l'inférence. Le problème d'apprentissage consiste à faire l'apprentissage des paramètres [4][5] et celui de la structure [6]. Si par définition l'apprentissage est le processus permettant la construction d'un modèle, l'inférence est l'opération qui permet de passer de la proposition à la conclusion à partir de ce modèle, c'est-à-dire, classer un nouvel exemple

### III. SIMULATION ET RESULTATS

Il s'agit dans cette partie de travail, d'appliquer les réseaux bayésiens dans un contexte de fusion de données pour surveiller un microclimat pour la culture de l'olivier à l'intérieur d'une serre agricole.

#### A. création de la base de données simulée

Pour un système de classification basé sur l'apprentissage, la composante la plus importante est la base de données qui permettra la construction du modèle de classification afin de distinguer les différents états du microclimat. Dans cette base de données, chaque vecteur descripteur  $X$  doit avoir sa classe d'appartenance  $C_i$ . Les données réelles n'étant pas disponibles, nous avons constitué une base de données simulée composée d'un certain nombre de cas fictifs de microclimat en se basant sur les valeurs optimales des caractéristiques climatiques pour un olivier sous serre fournies par un expert de la région d'Alger et qui sont résumées dans le tableau I.

TABLEAU I  
VALEURS OPTIMALES DES CARACTERISTIQUES CLIMATIQUES POUR UN OLIVIER SOUS SERRE.

Propriétés	Température ambiante $T_a$ (°C)	Température du sol $T_s$ (°C)	Humidité ambiante $H_a$ (%)	Rayonnement $R_a$ (Lux)
Valeurs des paramètres	$22 \leq T_a \leq 24$	$18 \leq T_s \leq 22$	$90 \leq H_a \leq 100$	$100 \leq R_a \leq 200$

La structure de cette base respecte les paramètres thermiques, hygrométriques et lumineux connus du microclimat adéquat pour un olivier sous serre et qui seront fournis par le programme établi. L'avantage d'avoir une base de données simulée est de pouvoir intégrer tous les cas possibles sans aucune difficulté, ce qui permet d'avoir une base la plus complète possible. Pour cela, nous avons organisé notre travail en deux fichiers. Le premier fichier appelé « *base\_donnees.m* » pour la base de données qui sera, par la

suite, divisée en deux ensembles, le premier pour l'entraînement du classificateur, le second pour le test et la validation du système de contrôle ainsi réalisé. Le deuxième fichier que nous avons nommé « *reseau\_bayésien.m* » est celui qui contient le code de notre application. Les quatre paramètres sont générés à l'aide de l'instruction « *rand* », en respectant les limites de chacun d'eux selon son adéquation ou non pour la culture en question. Ainsi, les différentes classes correspondent à tous les cas possibles de l'état du microclimat, selon l'état de chaque paramètre. Dans le cas où le microclimat est bon, cela implique, directement, que les quatre paramètres ( $T_a$ ,  $T_s$ ,  $H_u$  et  $R_a$ ) sont bons. On en comptabilise, alors, la première classe, telle que pour laquelle  $C_1 = 1$  correspond à « *microclimat-adéquat* ». Dans le cas où le climat de la serre est inadéquat, il faut savoir quel est ou quels sont les paramètres qui ne sont pas réglés, c'est-à-dire, un, deux, trois ou les quatre à la fois, tout en sachant que chacun d'eux peut être inférieur ou supérieur par rapport aux valeurs adéquates pour la plante. Le tableau II montre ces différents cas.

TABLEAU II  
LES DIFFERENTS ETATS DES PARAMETRES CLIMATIQUES

Les états / Les paramètres	Etat inadéquat inférieur	Etat adéquat	Etat inadéquat supérieur
La température ambiante $T_a$ (°C)	[-20,22]	[22,24]	[24,70]
La température du sol $T_s$ (°C)	[-20,18]	[18,24]	[22,70]
L'humidité ambiante $H_a$ (%)	[0,90]	[90,100]	-
Le rayonnement $R_a$ (Lux)	[0,100]	[100,200]	[200,1000]

L'avantage de connaître le paramètre et/ou les paramètres déréglés est de pouvoir intervenir efficacement et rapidement. Pour dénombrer tous les cas possibles, nous avons fait appel à l'analyse combinatoire. Comme l'ordre, dans notre cas, n'importe pas, il s'agit donc de combinaisons d'un nombre de  $p$  éléments pris parmi  $n$  pour chaque paramètre tel que :

$$C_n^p = \frac{n!}{p!(n-p)!} \quad (2)$$

Pour les trois paramètres  $T_a$ ,  $T_s$  et  $R_a$ , nous prenons un seul état parmi trois et pour l'humidité  $H_a$ , nous prenons un état parmi 2 simultanément (Tableau II).

En tous, nous avons obtenu  $N = 54$  cas différents calculé de la manière suivante :

$$N = C_3^1 \times C_2^1 \times C_3^1 \times C_3^1 = 3 \times 2 \times 3 \times 3 = 54 \text{ classes (états du microclimat)}$$

Ainsi, cette règle peut être généralisée sous la forme de l'expression 3 suivant le nombre de capteurs et les états de chaque paramètre utilisés dans une application donnée.

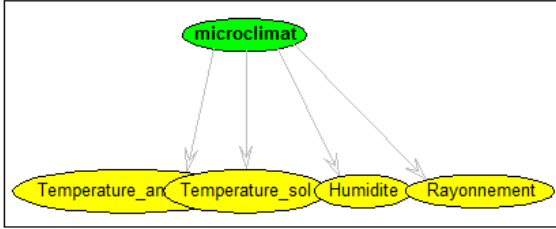
$$N = \prod_n (C_k^1)^n \quad (3)$$

où  $N$  est le nombre de tous les cas possibles,  $n$ , le nombre de paramètres utilisés dans l'application et  $k$ , le nombre d'états de chaque paramètre.

#### B. création du réseau bayésien

Le réseau bayésien a été créé par l'instruction *mk\_bnet* de *Matlab*. Il est constitué de cinq nœuds. Le nœud

$classe=microclimat$  est reconnu à partir des autres variables représentant les caractéristiques (Fig.2). La structure du réseau bayésien est représentée sous sa forme la plus simple avec une indépendance totale des caractéristiques conditionnellement au nœud microclimat. Les caractéristiques choisies sont les paramètres climatiques. Ces paramètres sont considérés comme des variables continues et les éventuels états du microclimat sont représentés par la variable



microclimat qui est considérée comme une variable discrète.

Fig.2 Le réseau Bayésien naïf des paramètres de la serre

Pour arriver à prendre une décision sur la nature du microclimat parmi les 54 cas possibles, notre algorithme passe par les quatre étapes utilisées lors de la fusion de données multi-capteurs[7], réalisées par le biais d'une méthode de classification basée sur les réseaux bayésiens. Notons que nous effectuons seulement l'apprentissage des paramètres car la structure est celle du réseau bayésien naïf.

- Modélisation

Cette étape est fondée sur l'utilisation des probabilités conditionnelles  $P(X/C_j)$  pour la représentation de toutes les variables dans un cadre commun. Le modèle proposé est un modèle où deux types de lois de probabilités sont mélangés, la loi discrète et la loi continue gaussienne. Après l'observation des deux types de variables, nous avons constaté que les différents états du microclimat sont considérés comme de simples étiquettes des classes dont les valeurs sont prises dans l'ensemble des entiers telle que  $C_j \in \{1,2,3, \dots, 54\}$ . L'observation de l'évolution des paramètres climatiques à partir de bases de données météorologiques provenant d'internet [8] nous a conduits à considérer que ces variables peuvent être modélisées par des densités de type gaussien.

- Estimation des paramètres

A partir du moment où l'étape de modélisation fait intervenir des distributions de probabilités, cette étape consiste à effectuer leurs calculs. Ces paramètres sont estimés, en utilisant, les algorithmes d'apprentissage de paramètres sans données manquantes qui est implémenté dans la BNT (Bayesian Network Toolbox). Prenons les notations suivantes pour les cinq variables du réseau bayésien.

$$x_1 = Ta, x_2 = Ts, x_3 = Ha, x_4 = Ra, C_j = microclimat$$

Les paramètres à calculer sont de deux types, à savoir : Le nœud  $C_j$  qui est un nœud parent et sans parents. La probabilité à calculer est  $P(C_j)$ . Ce terme est donné par la fréquence d'apparition de chaque classe dans la base de données puisqu'il s'agit d'une variable discrète. Dans notre cas toutes les classes sont considérées équiprobables.

$$P(C_1) = P(C_2) = \dots = P(C_{54}) = \frac{N(C_j)}{N} = \frac{N(C_j)}{\sum_{j=1}^{54} N(C_j)}, (\sum_{j=1}^{54} P(C_j) = 1) \quad (4)$$

où  $N(C_j)$  est le nombre d'observations appartenant à la classe  $C_j$  dans la base d'apprentissage et  $N$  la taille de la base.

Les nœuds  $x_i$  ( $i = \overline{1,4}$ ) sont des nœuds enfants. Les probabilités calculées sont les probabilités conditionnelles de chaque variable  $x_i$  par rapport à son parent  $C_j$ . Elles sont calculées en utilisant la loi normale  $N$ , telle que  $P(x_i/C_j) = N(x_i, \mu_{i,j}, var_{i,j})$  où  $\mu_{i,j}$  et  $var_{i,j}$  sont la moyenne et la variance du paramètre  $x_i$  dans la classe  $C_j$ .

Après avoir calculé les moyennes et les variances de tous les paramètres de chaque classe à partir de la base de données, le réseau est maintenant quantifié. Ainsi, calculons les probabilités conditionnelles des nouvelles observations dans chaque classe.

- Combinaison

Afin d'estimer la probabilité globale qu'une nouvelle observation appartienne à une classe  $C_j$ , il faut, dans un premier temps, combiner les différentes décisions multi-capteurs représentées par les probabilités estimées dans l'étape précédente en utilisant la règle de Bayes :

$$P(C_j/x_1, x_2, x_3, x_4) = \frac{P(C_j) \times P(x_1, x_2, x_3, x_4/C_j)}{P(x_1, x_2, x_3, x_4)} \quad (5)$$

où  $P(C_j/x_1, x_2, x_3, x_4)$  est la probabilité *a posteriori* de la classe  $C_j$  conditionnellement aux paramètres  $x_1, x_2, x_3, x_4$ ,  $P(C_j)$ , la probabilité *a priori* de la variable classe  $C_j$ ,  $P(x_1, x_2, x_3, x_4/C_j)$ , la vraisemblance des paramètres  $x_1, x_2, x_3, x_4$  sachant la classe  $C_j$ ,  $P(x_1, x_2, x_3, x_4)$ , la probabilité *a priori* des paramètres  $x_1, x_2, x_3, x_4$ , appelé évidence ou encore constante de normalisation.

Le terme  $P(x_1, x_2, x_3, x_4)$  dépend seulement des caractéristiques qui sont données. C'est un terme constant, facile à calculer et il est souvent négligé lors de la comparaison des différentes quantités  $P(C_j/x_1, x_2, x_3, x_4)$ , puisqu'il s'agit d'un terme positif. Il est donné par l'expression (6).

$$P(x_1, x_2, x_3, x_4) = \sum_{j=1}^{54} P(x_1, x_2, x_3, x_4/C_j) \times P(C_j) = \sum_{j=1}^{54} P(x_1/C_j) \times P(x_2/C_j) \times P(x_3/C_j) \times P(x_4/C_j) \times P(C_j) \quad (6)$$

Cependant, le terme qui nous intéresse le plus est  $P(x_1, x_2, x_3, x_4/C_j)$ . Il est soumis à la loi de probabilités de plusieurs variables avec une forte indépendance des caractéristiques, appelée « hypothèse d'indépendance naïve de Bayes » et qui nous permet de factoriser le numérateur comme suit :

$$P(x_1, x_2, x_3, x_4/C_j) \times P(C_j) = P(x_1/C_j)P(x_2/C_j)P(x_3/C_j)P(x_4/C_j) \times P(C_j) \quad (7)$$

$$P(C_j/x_1, x_2, x_3, x_4) = \frac{P(C_j) \times P(x_1/C_j) \times P(x_2/C_j) \times P(x_3/C_j) \times P(x_4/C_j)}{P(x_1, x_2, x_3, x_4)} \quad (8)$$

Cette équation montre que le type de combinaison s'exprime sous la forme d'un produit, c'est, donc, une combinaison

conjonctive. Elle est de plus associative. Sous l'hypothèse d'indépendance, elle devient commutative et associative.

- Décision

Le classificateur bayésien naïf est un modèle probabiliste basé, comme son nom l'indique, sur la règle de Bayes, qui dit qu'à tout élément  $X = (x_1, x_2, x_3, x_4)$ , on associe la classe  $C_j$  correspondant au maximum *a posteriori*, c'est-à-dire :

$$C_{MAP} = \operatorname{argmax}_{C_j} P(C_j/x_1, x_2, x_3, x_4) \propto \operatorname{argmax}_{C_j} P(C_j) \times P(x_1/C_j) \times P(x_2/C_j) \times P(x_3/C_j) \times P(x_4/C_j) \quad (9)$$

Comme exemple, prenons le cas de deux classes seulement  $C_1 = \text{microclimat adéquat}$  et  $C_2 = \text{microclimat inadéquat}$  (TABLEAU III).

TABLEAU III  
EXEMPLE DE DEUX CLASSES DU MICROCLIMAT

Etats du microclimat	La température ambiante (Ta)	La température du sol (Ts)	L'humidité ambiante (Ha)	Le rayonnement (Ra)
$C_1$	23.9529	20.1267	96.7245	141.2804
$C_1$	23.0466	20.0845	99.0183	101.4764
$C_1$	22.8598	21.0972	91.9916	170.2765
$C_1$	22.4143	18.4811	92.9828	150.6747
$C_1$	22.6468	20.5018	94.9652	138.1266
$C_1$	22.2217	19.3866	98.8990	106.4932
$C_1$	22.7504	19.3385	95.0141	135.8580
$C_1$	22.6598	20.2985	92.7700	123.4251
$C_1$	22.6842	21.4558	95.3396	120.3504
$C_1$	23.6342	18.7943	95.7425	181.3780
$C_2$	42.0980	25.3587	67.4863	30.9525
$C_2$	26.4645	67.6016	36.6588	68.7579
$C_2$	41.2524	29.5941	21.5543	98.6852
$C_2$	59.6497	35.7494	46.8765	76.9934
$C_2$	31.6037	54.9822	19.7170	82.9581
$C_2$	20.1376	-15.7894	75.8149	764.8678
$C_2$	-5.9549	0.5318	59.6640	676.2691
$C_2$	-5.4897	9.7384	73.4612	802.2993
$C_2$	-11.0136	18.8692	71.4490	597.3778
$C_2$	13.7574	10.2126	42.2195	892.1061

- Estimation des paramètres (Tableau IV)

Classes	$\mu_1(x_1)$	$var_1(x_1)$	$\mu_2(x_2)$	$var_2(x_2)$
$C_1$	22.8871	0.2840	19.9565	0.9159
$C_2$	21.2505	555.2334	23.6849	620.5510
	$\mu_3(x_3)$	$var_3(x_3)$	$\mu_4(x_4)$	$var_4(x_4)$
$C_1$	95.3448	5.7662	136.9339	659.8511
$C_2$	51.4901	447.5953	409.1267	$1.3257 \times 10^5$

TAB  
LEA  
U IV  
CAL  
CUL  
DES  
MO  
YEN  
ES  
ET  
DES  
VAR  
IAN  
CES

$P(x_1 = Ta/C_1) = N(x_1 = Ta, 22.8871, 0.2840)$ ,  
 $P(x_3 = Ha/C_2) = N(x_3 = Ha, 51.4901, 447.5953)$   
 Prenons une nouvelle observation à classer du tableau V

TABLEAU V  
UNE NOUVELLE OBSERVATION A CLASSER

Etat du microclimat	Ta (°C)	Ts (°C)	Ha (%)	Ra (Lux)
Inconnu	22,3	18	95,77	154,5

$$P(C_1) = P(C_2) = \frac{10}{20} = 0.5, \quad P(x_1 = 22.3/C_1) = 0.4080, \\ P(x_2 = 18/C_1) = 0.0516 \\ P(x_3 = 95.77/C_1) = 0.1643, \quad P(x_4 = 154.5/C_1) = 0.0123, \\ P(x_1 = 22.3/C_2) = 0.0169 \\ P(x_2 = 18/C_2) = 0.0156, \quad P(x_3 = 95.77/C_2) = 0.0021, \\ P(x_4 = 154.5/C_2) = 0.0009$$

- Combinaison

$$P(C_1/Ta, Ts, Ha, Ra) \propto P(C_1) \times P(Ta/C_1) \times P(Ts/C_1) \times P(Ha/C_1) \times P(Ra/C_1) \\ P(C_1/Ta, Ts, Ha, Ra) \propto 0.5 \times 0.4080 \times 0.0516 \times 0.1643 \times 0.0123 = 2.1257 \times 10^{-5} \\ P(C_2/Ta, Ts, Ha, Ra) \propto 0.5 \times 0.0169 \times 0.0156 \times 0.0021 \times 0.0009 = 2.4053 \times 10^{-10}$$

- Décision

Cependant, pour prendre une décision sur la nature du microclimat, il suffit de choisir la classe correspondant au maximum *a posteriori* donné par l'équation comme suit :

$$C_{MAP} = \operatorname{argmax}_{C_j} P(C_j/x_1, x_2, x_3, x_4) \propto \operatorname{argmax}_{C_j} P(C_j) \times P(x_1/C_j) \times P(x_2/C_j) \times P(x_3/C_j) \times P(x_4/C_j) \quad (10)$$

Comme  $P(C_1/Ta, Ts, Ha, Ra) > P(C_2/Ta, Ts, Ha, Ra)$

Alors l'état du microclimat est, effectivement, adéquat pour l'olivier.

#### IV. EVALUATION DU MODELE DEVELOPPE

Pour évaluer le modèle établi en termes de taux de bonne classification et rapidité d'exécution, nous faisons un certain nombre de tests sur la base d'apprentissage et la base de test en appliquant deux algorithmes d'inférence différents pour pouvoir choisir le mieux adapté à la nature de nos données. Nous faisons varier aussi la taille de la base d'apprentissage. Nous calculons, à chaque fois, le taux et le temps de classification sur la base d'apprentissage et sur la base de test.

##### A. Tests sur la base d'apprentissage

Nous utilisons six échantillons de tailles respectives de 108, 540, 1080, 2700, 5400 et 16200, sur lesquels, nous avons appliqué l'algorithme d'apprentissage du maximum de vraisemblance et les deux algorithmes d'inférence l'arbre de jonction (ADJ) et l'élimination des variables (EDV). Dans le but de conclure par rapport à la validité de notre méthode, nous calculons la matrice de confusion qui permet de donner le taux de bonne classification. Dans chaque cas les tests ont

été répétés 10 fois. Pour chacune des différentes tailles, nous avons calculé le taux de reconnaissance et le temps d'exécution en effectuant les moyennes des taux de classification et des temps d'exécution obtenus pour les 10 tests. Ce taux de classification ainsi que le temps d'apprentissage sont portés dans le tableau VI.

**TABLEAU VI**  
RESULTATS OBTENUS EN APPLIQUANT LES ALGORITHMES ADJ ET EDV SUR LA BASE D'APPRENTISSAGE

Taille de la base d'apprentissage		108	540	1080	2700	5400	16200
Algorithmes ADJ et EDV	max	100	100	98.33	92.88	92.16	91
	min	100	83.33	85.55	85.55	87.96	82.55
	moy	100	91.66	91.44	88.99	89.82	90.22
Temps d'apprentissage (s) ADJ		3.84	7.45	15.04	35.43	70.32	212.1
Temps d'apprentissage (s) EDV		8.52	12.47	22.78	53.88	75.63	498.4

Les résultats présentés, ci-dessus, montrent la capacité notable d'apprentissage de cette méthode. Même lorsque le nombre de données de la base d'apprentissage augmente, le taux de bonne classification reste satisfaisant avec un taux supérieur en moyenne à 88%. Le temps d'apprentissage est plus grand lorsque la taille de la base est importante.

Les deux algorithmes donnent exactement le même taux de classification pour les mêmes exemples. La différence réside dans le temps d'apprentissage qui est plus réduit avec l'algorithme de l'arbre de jonction. En effet, l'algorithme le mieux adapté pour notre application est l'arbre de jonction.

#### B. Tests sur la base de teste

Nous avons créé une base de test avec de nouveaux exemples. Pour étudier l'effet de la taille des données d'apprentissage sur la qualité de la méthode à classer de nouveaux exemples, nous avons choisi cinq échantillons d'apprentissage de tailles différentes de 162, 270, 540, 2700 et 5400 observations, en gardant la taille de la base de test constante à chaque fois.

Voici les résultats pour deux bases de test de tailles 54 et 162 observations.

**TABLEAU VII**  
RESULTATS OBTENUS SUR LA BASE DE TEST DE 54 OBSERVATIONS

La taille de la base d'apprentissage	162	270	540	2700	5400
Le taux de classification (%)	66.66	68.88	77.77	96.66	99.11
Le temps d'apprentissage (s)	6.61	10.25	14.91	50.68	107.62

**TABLEAU VII**  
RESULTATS OBTENUS SUR LA BASE DE TEST DE 162 OBSERVATIONS

La taille de la base d'apprentissage	162	270	5400	16200
Le taux de classification (%)	56.23	81,11	97.77	98.66
Le temps d'apprentissage (s)	13.421	15.566	95.383	286.266

Nous constatons d'après les résultats des deux tableaux que le taux de bonne classification est en relation directe avec la taille de la base d'apprentissage. Plus la base est riche meilleur est le taux de classification. L'enrichissement de cette base reste indispensable pour avoir des taux d'erreurs faibles. Dans le contexte où nous travaillons, nous avons besoin de contrôler l'état du microclimat avec une période bien déterminée. Alors, le test se fait, réellement, sur une mesure à la fois. Ainsi, avec une base d'apprentissage de 540 données, nous obtenons un taux de bonne classification de 100% pour les 54 états du réseau.

## V. CONCLUSION

Le travail que nous avons présenté a été consacré à l'application de la technique d'apprentissage probabiliste des réseaux bayésiens sur une fusion de données multi-capteurs dans un contexte de classification pour le contrôle et la surveillance du microclimat d'une serre agricole afin de l'adapter à la culture d'olivier. Cette technique a fait preuve de robustesse et d'efficacité sur plusieurs types de données, complètes et incomplètes et prend en considération les incertitudes et les imprécisions des données délivrées par les différents capteurs. Les réseaux bayésiens sont l'une des méthodes les plus importantes de l'intelligence artificielle et restent préférables dans plusieurs cas car ils permettent : le rassemblement de plusieurs données de natures différentes dans un seul modèle, la représentation graphique qui reste compréhensible et facile à interpréter pas uniquement par des spécialistes contrairement aux systèmes experts par exemple, sans oublier qu'avec un réseau bayésien, on peut évaluer, prévoir, diagnostiquer ou optimiser des décisions.

Les étapes que nous avons suivies pour arriver à la fin de notre étude sont : la création de la base de données suivant les paramètres adéquats pour la meilleure croissance de l'olivier, modélisation du problème en utilisant les distributions de probabilités, estimation de ces probabilités et enfin la décision sur la nature du microclimat suivant les résultats donnés par le réseau.

Malgré la structure simple du réseau bayésien naïf que nous avons utilisé, les résultats obtenus sont très convaincants avec des taux de bonne classification atteignant les 100% et des temps de simulation réduits en utilisant l'algorithme d'inférence de l'arbre de jonction qui est le mieux adapté à la nature de nos données.

Cette méthode reste généralisable par rapport au nombre de paramètres utilisés et applicable pour n'importe quelle autre

plante, à condition d'avoir les valeurs des paramètres les mieux adaptés pour sa croissance.

En perspective, il serait intéressant :

- d'étudier le positionnement et le placement de manière appropriée de ces capteurs,
- de commander les différents actionneurs pour contrôler chaque paramètre,
- de faire la réalisation pratique dans une serre réelle.

## REFERENCES

[1] David BELLOT, « *Fusion de données avec des réseaux Bayésiens pour la modélisation des systèmes dynamiques et son application en télémédecine* », 2005.

[2] M. Rombaut, « *Fusion : état de l'art et perspectives* », Laboratoire LM2S-UTT, Université de Troyes, 2001.

[3] Ayman ZUREIKI, « *Fusion de Données Multi-Capteurs pour la Construction Incrémentale du Modèle Tridimensionnel Texturé d'un Environnement Intérieur par un Robot Mobile* », Thèse de Doctorat. Université de Toulouse, 2008.

[4] David Heckerman, « *A tutorial on learning with Bayesian network* », In Michael I. Jordan, editor, *learning in Graphical Models*, Kluwer Academic Publishers, Boston, 1998.

[5] Paul Krause, « *Learning probabilistic networks* », 1998.

[6] Patrick Naim, Pierre-Henri Wuillemin, Phillippe Leray, Olivier Pourret, Anna Becker, « *Réseaux Bayésiens* », Editions EYROLLES, 2007.

[7] Isabelle Bloch, « *Fusion d'informations en traitement du signal et des images* », Hermes Science Publication, Janvier 2003.

[8] <http://www.meteociel.fr>

# Optimisation de paramètres dans une serre agricole par fusion de données multi capteurs

Boussanou Tassadit<sup>1</sup>, Mazari Chahrazad<sup>2</sup>, Diaf Moussa<sup>3</sup>

Département automatique, FGEl, UMMTO

Université Mouloud Mammeri Tizi-Ouzou

[hbtassadit@hotmail.fr](mailto:hbtassadit@hotmail.fr), [ma.charazad@hotmail](mailto:ma.charazad@hotmail), [mous.diaf@gmail.com](mailto:mous.diaf@gmail.com)

**Résumé**—Dans cet article, nous présentons une étude relative à la fusion de données. Cette technique permet d'améliorer la conduite du contrôle du climat dans une serre agricole équipée d'un système multi-capteurs. Pour ce faire, nous utilisons les fonctions de croyance pour trouver les conditions optimales du microclimat favorable à la croissance d'un olivier dans cette serre. Nous décomposons le processus de fusion en trois phases. En premier lieu, nous modélisons les informations issues des capteurs et des connaissances expertes, dans un cadre commun. En second temps, nous fusionnons les masses de croyance des capteurs par une somme conjonctive. Cette phase est suivie d'une prise de décision. En dernier lieu, nous présentons le résultat de la décision à un expert afin d'évaluer la qualité de nos résultats.

**Mots-clés**— Optimisation, serre agricole, fusion de données, multicapteurs.

## I. INTRODUCTION

La fusion multi-capteur se présente comme un moyen le plus sûr permettant le contrôle intelligent de l'environnement dans une serre agricole. La fusion d'information provenant de différents capteurs qui permettent de recueillir les paramètres d'environnement (température, humidité, rayonnement dans la serre) a fait l'objet des recherches déjà anciennes [1]. L'objectif de ces recherches est d'explorer les approches modernes pour l'amélioration de la conduite du contrôle du climat dans une serre agricole. Pour cela, différentes méthodes de fusion utilisant les données issues des systèmes multicapteurs ont été mises en œuvre. L'objectif scientifique de notre travail est l'adaptation d'un modèle des croyances transférables qui permet d'optimiser les paramètres du climat dans des intervalles bien déterminés selon le type de la plante à cultiver. La modélisation consiste, d'abord, à définir les facteurs de changement Microclimatique à l'intérieur de la serre agricole. Dans un second temps, toutes les données numériques issues des capteurs pendant la période cultivatrice d'un olivier son cumulées. Cette accumulation a permis de constituer une base de donnée que nous avons utilisée pour réaliser les travaux de modélisation, tout en nous appuyant sur des connaissances d'un expert du domaine [2].

## II. DONNEES UTILISEES

Ainsi, les données utilisées dans cette étude sont des données numériques collectées par les capteurs cités

précédemment durant la période cultivatrice de l'olivier dans la serre.

Quatre grandeurs de chaque paramètre ont été donc proposées par un expert pour former des combinaisons représentant les hypothèses du cadre de discernement  $\Omega$  de notre travail. Les grandeurs utilisées sont données présentées dans le tableau I

TABLEAU I  
QUATRE GRANDEURS DE CHAQUE PARAMÈTRE

Hypothèse	Température du sol en c°	Température ambiante c°	Humidité en%	Rayonnement en lux
H1	19	23	85	100
H2	20	24	95	130
H3	21	25	90	120
H4	22	26	100	170

Les données expertes qui représentent les besoins climatiques d'un olivier sous serre sont les suivantes: la température du substrat d'enracinement qui doit être entre  $[18^{\circ}, 22^{\circ}]$ , la température ambiante entre  $[22^{\circ}, 24^{\circ}]$ , l'humidité ambiante dans l'intervalle  $[90\%, 100\%]$  et la luminosité, dans  $[100\text{lux}, 200\text{lux}]$ .

## III. METHODOLOGIE

Dans le but de mener efficacement le contrôle ou la décision doit être prise de façon optimale selon les besoins de olivier cultivé dans la serre, chaque fois que la probabilité d'incertitude est la plus élevée afin d'assurer les meilleurs récolte. Ainsi, nous avons opté pour un formalisme mathématique fondé sur la théorie des fonctions de croyance transférable [3]. Le processus de la fusion d'information est décrit par quatre étapes: la modélisation, l'estimation, la combinaison et la décision. La modélisation définit le choix du formalisme qui est, dans notre cas, la théorie des fonctions de croyance (fig.1). L'estimation permet de définir les fonctions choisies dans l'étape de modélisation en fonction de l'application. La combinaison est la phase de regroupement des informations. La dernière étape consiste à prendre la décision sur le résultat de la combinaison [4].

### A. Affectation de masse

Le problème de l'affectation des masses de croyances aux différentes hypothèses est un problème majeur en fusion de données. La première étape à réaliser consiste définir le cadre de discernement qui est fortement dépendant de

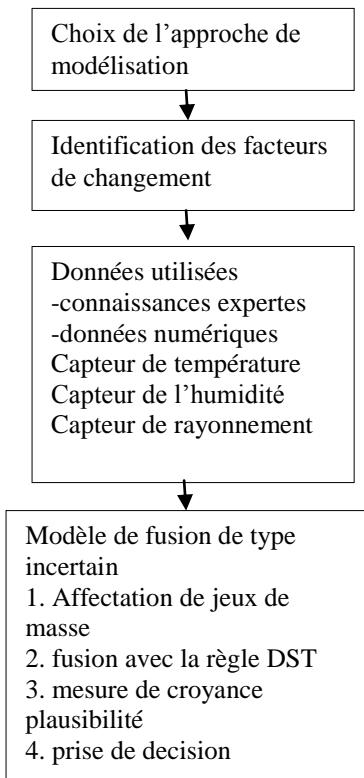


Fig.1 Méthodologie du processus de modélisation prédictive de la serre agricole

l'application. Cette première étape est généralement établie de manière supervisée par un expert. Une fois cette étape réalisée, la modélisation des fonctions de masse est un problème crucial où il n'y a pas de méthode générique. Selon le type d'application rencontrée, il existe différentes méthodes permettant d'élaborer les masses de croyances. La plus simple et la plus utilisée reste encore l'expertise humaine dans laquelle les coefficients sont établis manuellement. Pour notre travail, une affectation de masse respecte les bornes inférieures et supérieures de chaque intervalle donné précédemment. Pour initialiser ces masses de croyances, nous avons utilisé l'approche floue pour prendre en compte des imprécisions et, ainsi, améliorer la distribution de masse affectée à chaque hypothèse et ceci, même si les informations sont incomplètes. Il s'agit donc d'utiliser des degrés d'appartenance. A cet effet, pour déterminer ces masses, nous avons calculé les degrés d'appartenance des différents paramètres tel que :

$$u_{\theta_i} \in [0,1]$$

ce qui détermine le degré d'appartenance de l'élément  $x$  à un jeu flou  $\theta_i$ .

La détermination des fonctions de masse à l'aide de cette approche permet d'attribuer des valeurs de masse pour chaque hypothèse à partir des degrés d'appartenance.

Dans le cas où nous avons deux hypothèses  $\theta_1$  et  $\theta_2$ :

$u_{\theta_1}(x)$  est le degré d'appartenance de  $x$  à un jeu flou  $\theta_1$

$u_{\theta_2}(x)$ :degré d'appartenance de  $x$  à un jeu flou  $\theta_2$

$$\hat{m}(\theta_1) = u_{\theta_1}(x)$$

$$\hat{m}(\theta_2) = u_{\theta_2}(x)$$

$$\hat{m}(\theta_1 \cup \theta_2) = \hat{s}(u_{\theta_1}(x), u_{\theta_2}(x)) \quad (1)$$

où  $\hat{m}(\cdot)$  est une fonction de masse non normalisée et  $\hat{s}(\cdot)$  une estimation caractéristique de l'hypothèse composée.

### B. Application

Nous donnons dans cette section les résultats obtenus pour le calcul des degrés d'appartenance des quatre grandeurs de chaque hypothèse. Les figures suivantes donnent, respectivement, le degré d'appartenance en fonction de la température intérieure, la température ambiante, l'humidité et le rayonnement des quatre hypothèses.

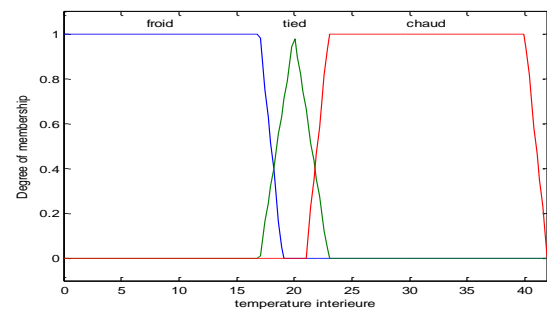


Fig.2 Degré d'appartenance en fonction de la température du sol

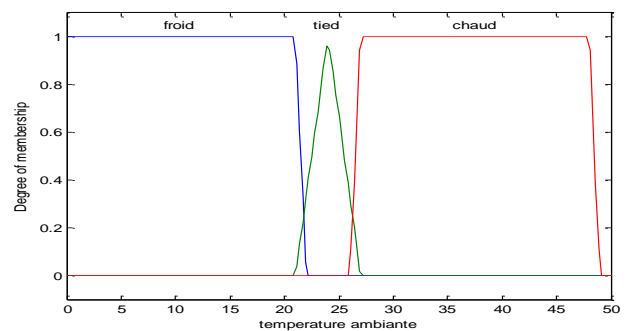


Fig.3 Degré d'appartenance en fonction de la température ambiante

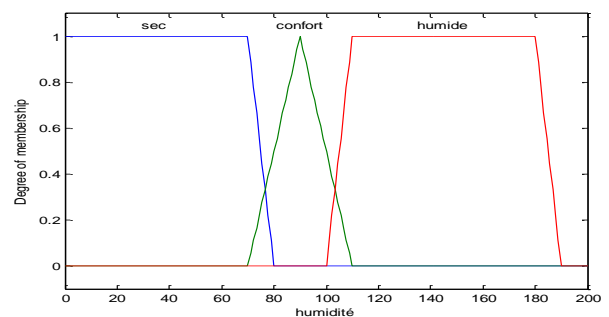


Fig. 4 Degré d'appartenance en fonction de l'humidité



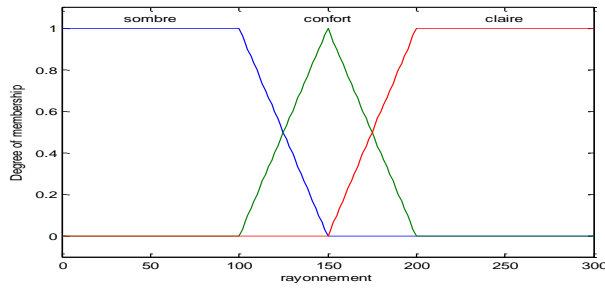


Fig.5 Degré d'appartenance en fonction de rayonnement

Les masses de ces grandeurs ont été déterminées à partir des Fig.2,3,4 et 5 ainsi que du tableau II en les équations (1).

TABLEAU III  
LES FONCTIONS DE MASSE NON NORMALISÉES

hypothèses	Sources d'informations			
	Capteur 1 Température du sol	Capteur2 Température Ambiante	Capteur3 Humidité	Capteur4 Rayonnement
H1	0.7500	0.6666	0.6666	0.0000
H2	0.7500	1.0001	1.0001	0.6000
H3	0.9999	0.6666	0.6666	0.4000
H4	0.5001	0.3333	0.3333	0.6000

Pour normaliser ces masses d'informations, nous avons utilisé la technique de normalisation suivante.

$$m_i(.) = \frac{\hat{m}_i(.)}{\sum \hat{m}_i(.)} \quad (2)$$

ce qui nous donne le le tableau III.

TABLEAU IIIII  
LES FONCTIONS DE MASSE NORMALISÉES

Hypothèses	Plausibilité
H1	0.0000
H2	<b>0.6804</b>
H3	0.2680
H4	0.0515

Après la normalisation, nous appliquons la fusion des masses normalisées obtenues précédemment en utilisant la technique de fusion de Dempster-Shafer [5][6]. L'intérêt majeur de cette technique repose sur la possibilité de construire une fonction de masse unique par la combinaison de plusieurs fonctions de masse issues de plusieurs sources d'information distinctes comme notre cas où il s'agit de capteurs [7][8]. Ainsi, l'architecture de combinaison des quatre capteurs est représentée sur la figure (6).

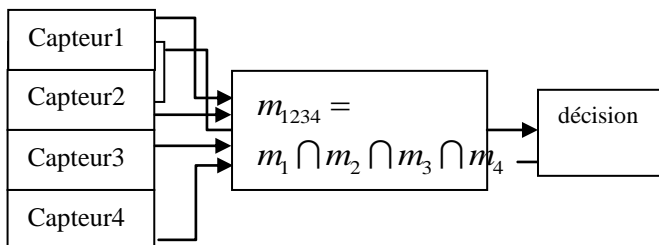


Fig.6 Architecture de combinaison des quatre capteurs

Nous obtenons le jeu de masse final  $m_{1234}$  présenté dans le tableau IV.

TABLEAU IVV  
LES FONCTIONS DE MASSE APRÈS LA COMBINAISON

Hypothèses	$m_{1234}$
H1	0.0000
H2	0.6804
H3	0.2680
H4	0.0515

### C. Décision

Une fois les sources d'informations fusionnées, il est nécessaire de sélectionner l'hypothèse la plus probable. Pour prendre une meilleure décision, nous avons utilisé la méthode de calcul de plausibilité. La plausibilité constitue le degré de confiance maximale que l'on attribue à une hypothèse. Ce critère permet de déterminer la meilleure hypothèse en donnant le maximum de chances à chacune d'elles [9][10]. La plausibilité de chaque hypothèse est donnée dans le tableau V.

TABLEAU V  
LES PLAUSIBILITÉS DES QUATRE HYPOTHÈSES

hypothèses	Sources d'informations			
	Capteur 1 Température du sol	Capteur2 Température Ambiante	Capteur3 Humidité	Capteur4 Rayonnement
H1	0.2500	0.2500	0.2500	0.0000
H2	0.2500	0.3750	0.3750	0.3750
H3	0.3333	0.2500	0.2500	0.2500
H4	0.1667	0.1250	0.1250	0.3750

## IV. CONCLUSION

Dans cet article nous avons procédé à l'étude de la fusion multi-capteurs pour l'appliquer à une serre agricole pendant la période cultivatrice d'un olivier sous serre. En effet, nous avons pu obtenir les paramètres optimaux du microclimat qui favorise le bon développement d'un olivier à partir des données incomplètes et imprécises, grâce à une approche fondée sur les fonctions de croyance. L'utilisation de ces fonctions est particulièrement intéressante lorsque l'on ne dispose pas de modèle mathématique précis de la serre et des capteurs utilisés. Pour la suite de ce travail, il serait intéressant d'effectuer des études comparatives avec plusieurs autres techniques et procéder à des tests réels sur le terrain où il s'agira particulièrement de procéder à la commander des différents systèmes de contrôle.

### REFERENCES

- [1] Berroug « Simulation numérique des besoins énergétiques d'une serre agricole sous le climat de Marrakech – Maroc » Université Cadi Ayyad, Marrakech, Maroc 2010
- [2] M Arif, "Fusion de Données : Ultime Etape de Reconnaissance de Formes, Applications à l'Identification et à l'Authentification". Thèse de doctorat. Université de Savoie, 2005.

- [3] G. Shafer. A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ,
- [4] N.-E El Faouzi, "Fusion de données", Rapport INRETS-LICIT (France), 2004.
- [5] A. Dempster, Upper and lower probabilities induced by multi-valued mapping. *Annals of Mathematical Statistics* AMS-38, pp. 325–339, 1967
- [6] M. Oussalah. Fusion de données par la théorie des possibilités, application à la localisation d'un robot mobile. Thèse, Université d'Evry Val d'Essonne (EVE), 1998
- [7] V. B. Dasarathy, « Sensor fusion potentiel exploitation-innovative architecture and illustrative applications », *Proc. of IEEE*, Vol. 85, pp. 24-39, 1997.
- [8] I Bloch, " Fusion de Données en traitement d'images : modèles d'information et décision". *Traitement du signal*, 1994, vol.11, n°6, pp 435 - 446.
- [9] E. Lefevre, Fusion adaptée d'informations conflictuelles dans le cadre de la théorie de l'évidence, application au diagnostic médical. Thèse, INSA de Rouen. 2001
- [10] D. Gruyer. Etude du traitement de données imparfaites pour le suivi multi objets : Application aux situations routières. Thèse, université de Technologie de Compiègne, 1999.

# Estimation des fréquences 3-D par la méthode ESPRIT 3-D dans un bruit Gaussien Coloré

Youness CHAWKI<sup>1</sup>, Mohammed OUANAN<sup>2</sup> et Brahim AKSASSE<sup>3</sup>

Université Moulay Ismail, Faculté des Sciences et Techniques

Equipe ASIA, Département d'Informatique

BP 509 Boutalamine, Errachidia Maroc

<sup>1</sup> youness.chawki@gmail.com

<sup>2</sup> ouanan\_mohammed@yahoo.fr

<sup>3</sup> baksasse@yahoo.com

**Résumé**— *L'estimation des paramètres d'un modèle (les fréquences 3-D), en se basant sur les méthodes d'Analyse Spectrale Haute Résolution connues par leurs performances et leurs précisions type ESPRIT 3-D, reste un problème qui s'impose dans la modélisation des signaux par une somme d'exponentielles complexes 3-D (modèle SEC 3-D) noyés dans un bruit gaussien additif. En effet, de bons résultats sont bien obtenus en présence d'un bruit blanc en utilisant les Statistiques d'Ordre Deux « SOD » (les autocorrélations), mais lorsque le bruit devient coloré, les résultats se dégradent ce qui nous oblige, pour remédier à cet inconvénient, à penser aux Statistiques d'Ordre Supérieur « SOS » (Cumulants et moments d'ordre supérieur à deux).*

**Mots clés**— Analyse Spectrale, Haute Résolution, ESPRIT 3-D, Statistique d'Ordre 2, Statistique d'Ordre Supérieur, Cumulants d'ordre 4.

## I. INTRODUCTION

La modélisation des signaux noyés dans un bruit occupe une place très importante dans les axes de recherches ces dernières années. C'est une technique utilisée dans plusieurs domaines et applications tels que les télécommunications, traitement d'antenne et traitement de l'image.

Les méthodes d'analyse spectrale peuvent être classées en deux catégories: méthodes de balayage et méthodes Haute Résolution nommées encore méthodes analytiques ou approche de sous-espace. La première famille rassemble les méthodes qui restituent l'information spectrale par le biais d'une fonctionnelle dépendant d'un vecteur fréquentiel, il s'agit des méthodes dites du pseudo-spectre. La seconde famille regroupe les méthodes qui exploitent la structure matricielle inhérente au modèle SEC 3-D. Elle propose une phase d'estimation relative aux triplets

fréquentiels contenus dans le modèle. On peut citer la méthode MEMP [1] ou ses extensions dans le cas 3-D [2] [3], la méthode ACMP [4] et la méthode ESPRIT [5] et leur extension pour les signaux 2-D [6] [7] et dans le cas 3-D [8].

Les Statistiques d'Ordre Supérieur [3] [9] [10] [11] sont utilisées essentiellement en complément aux Statistiques d'Ordre Deux. En effet, elles donnent une description plus complète des données et de leurs propriétés et elles permettent la résolution de problèmes insolubles à l'ordre deux et cela est bien concret dans le cas où le bruit est coloré. Dans ce travail, on a appliqué l'approche des statistiques d'ordre supérieur à la nouvelle méthode ESPRIT 3-D développé dans [8].

Cet article est organisé comme suit: section 2 présente le modèle SEC 3-D. Dans section 3, les Cumulants d'ordre 4 sont développés ensuite les Cumulants de la nouvelle méthode ESPRIT 3-D sont calculés. Les résultats de la simulation et de la comparaison sont présentés dans section 4. Finalement, le travail s'achève par une conclusion et perspectives.

## II. MODELE DU SIGNAL SEC 3-D

Considérons que chaque voxel  $(m, n, t)$  de bloc d'image observée correspond à la somme de deux termes:

$$y(m, n, t) = x(m, n, t) + b(m, n, t) \quad (1)$$

Avec  $1 \leq m \leq M$ ,  $1 \leq n \leq N$  et  $1 \leq t \leq T$

Le signal utile  $x(m, n, t)$  est modélisé de la façon suivante:

$$x(m, n, t) = \sum_{i=1}^K a_i \exp(j2\pi(f_{1i}m + f_{2i}n + f_{3i}t) + j\phi_i) \quad (2)$$

- Les K composantes du signal sont définies par les

triplets fréquentiels  $\{f_{1i}, f_{2i}, f_{3i}\}$ , les amplitudes  $\{a_i\}$  et the phases  $\{\varphi_i\}$

-  $b(m, n, t)$  est le bruit additif gaussian coloré.

Notre objectif est, étant donné un bloc  $M \times N \times T$ , comment estimer les fréquences  $\{f_{1i}, f_{2i}, f_{3i}\}$  pour  $i = 1 \dots K$

### III. CUMULANTS DE LA NOUVELLE METHODE ESPRIT 3-D

#### A. Cumulant d'ordre 4 des données 3-D

Dans ce paragraphe, les indices sont notés:

$$m = (m_1, m_2, m_3), f = (f_1, f_2, f_3), h = (h_1, h_2, h_3), \alpha = (\alpha_1, \alpha_2, \alpha_3)$$

$$\text{et } m+h = (m_1+h_1, m_2+h_2, m_3+h_3)$$

Pour développer les techniques d'estimation HR à partir des Statistiques d'Ordre Supérieur, nous allons utiliser les cumulants d'Ordre 4. Avec les notations précédentes, les cumulants d'ordre 4 des observations que nous allons ensuite utiliser, sont définis par:

$$C_{4y}(h, k, l) = \text{Cum}[y(m), y^*(m+h), y(m+k), y^*(m+l)] \quad (3)$$

Où (\*) désigne l'opérateur conjugué complexe, E est l'opérateur d'espérance et le cumulant d'ordre 4 est donné par:

$$\text{Cum}(y_i, y_j, y_k, y_l) = E[y_i, y_j, y_k, y_l] - E[y_i, y_j]E[y_k, y_l] - E[y_i, y_k]E[y_j, y_l] - E[y_i, y_l]E[y_j, y_k] \quad (4)$$

En reprenant le modèle des observations bruitées donné en (1), les cumulants d'ordre 4 des observations sont données par:

$$C_{4y}(h) = C_{4x}(h) + C_{4b}(h) \quad (5)$$

Or les statistiques d'ordre supérieur à 2 de signaux gaussiens n'apportent pas d'informations: les cumulants d'ordre à 2 sont nuls. Ainsi les cumulants d'ordre 4 des observations coïncident théoriquement avec des observations non bruitées:

$$C_{4y}(h) = C_{4x}(h) \quad (6)$$

On voit bien que les statistiques d'ordre supérieur à 2 permettent d'éliminer théoriquement le bruit. Par ailleurs, les cumulants d'ordre 4 du processus harmonique 3-D considéré vérifient la relation suivante :

$$C_{4y} = -\sum_{i=1}^K a_i^4 \exp[j2\pi(-h+k+l)f_i^T] \quad (7)$$

Dans la suite, on considère uniquement la tranche diagonale appelée des cumulants d'ordre 4 et définie par:

$$C_{4y}(h) = C_{4x}(h, h, h), h = (h, h, h) \quad (8)$$

Les deux équations (6) et (7) nous permet d'écrire:

$$C_{4y} = -\sum_{i=1}^K a_i^4 \exp[j2\pi(h_1 f_{k1} + h_2 f_{k2} + h_3 f_{k3})] \quad (9)$$

Cette relation montre que la diagonale des cumulants d'ordre 4 contient toute l'information utile pour estimer les fréquences du modèle harmonique et elle est la base de l'extension des techniques HR.

#### B. Estimation des fréquences suivant la première dimension

Afin d'estimer les fréquences de la première dimension  $\{f_{1k}\}_{k=1}^K$  on considère la matrice des cumulants de type TBBT suivante:

$$C_{y,3} = \begin{bmatrix} C_0 & C_{-1} & \dots & C_{-P+1} \\ C_1 & C_0 & \dots & C_{-P+2} \\ \cdot & \cdot & \cdot & \cdot \\ C_{P-1} & C_{P-2} & \dots & C_0 \end{bmatrix} \quad (10)$$

Pour chaque  $p \in [-(P-1), (P-1)]$ , la matrice  $C_p$  est de taille  $Q \times Q$  ayant une structure TBT telle que:

$$C_p = \begin{bmatrix} C_p^0 & C_p^{-1} & \dots & C_p^{-Q+1} \\ C_p^1 & C_p^0 & \dots & C_p^{-Q+2} \\ \cdot & \cdot & \dots & \cdot \\ C_p^{Q-1} & C_p^{Q-2} & \dots & C_p^0 \end{bmatrix} \quad (11)$$

Et chaque bloc  $C_p^q$ , tel que  $q \in [-(Q-1), (Q-1)]$  est une matrice  $Q \times Q$  elle-même Toeplitz:

$$C_p^q = \begin{bmatrix} c_{4y}(i, j, 0) & c_{4y}(i, j, -1) & \dots & c_{4y}(i, j, -L+1) \\ c_{4y}(i, j, 1) & c_{4y}(i, j, 0) & \dots & c_{4y}(i, j, -L+2) \\ \cdot & \cdot & \dots & \cdot \\ c_{4y}(i, j, L-1) & c_{4y}(i, j, L-2) & \dots & c_{4y}(i, j, 0) \end{bmatrix} \quad (12)$$

La matrice des cumulants d'ordre 4 peut s'écrire de la manière suivante:

$$C_{y,1} = S_{[PQL, K]}^1 \Psi S_{[PQL, K]}^{1H} \quad (13)$$

Où:

$\Psi$  est la matrice diagonale de taille  $K \times K$  telle que:

$$\Psi = \text{diag}(a_i^4)_{1 \leq i \leq K} \quad (14)$$

$^H$  est l'opérateur Hermitien

$S_{[PQL,K]}^1$  est la matrice de Vandermonde 3-D de taille  $PQL \times K$  dont l'expression est:

$$S_{[PQL,K]}^1 = [s_{31,L} \otimes s_{21,Q} \otimes s_{11,P}, \dots, s_{3K,L} \otimes s_{2K,Q} \otimes s_{1K,P}] \quad (15)$$

Où:

$$s_{m,G} = [1 \quad \exp(j2\pi f_m) \quad \dots \quad \exp(j2\pi f_m(G-1))] \quad (16)$$

Pour  $m = 1, 2, 3$ ,  $n = 1, 2, \dots, K$  et  $\otimes$  est le produit de Kronecker.

La matrice de Vandermonde 3-D  $S_{[PQL,K]}^1$  peut s'écrire en fonction de la matrice de Vandermonde 2-D  $S_{[PQ,K]}^1$ :

$$S_{[PQL,K]}^1 = \begin{bmatrix} S_{[PQ,K]}^1 \\ S_{[PQ,K]}^1 \Phi_3 \\ \vdots \\ S_{[PQ,K]}^1 \Phi_3^{L-1} \end{bmatrix} \quad \text{et} \quad S_{[PQ,K]}^1 = \begin{bmatrix} S_{[P,K]}^1 \\ S_{[P,K]}^1 \Phi_2 \\ \vdots \\ S_{[P,K]}^1 \Phi_2^{Q-1} \end{bmatrix} \quad (17)$$

Avec  $S_{[P,K]}^1$  est la matrice de Vandermonde 1-D:

Où la matrice est la matrice de Vandermonde 1-D associée aux fréquences  $\{f_{m_k}\}_{k=1}^K$  définie par:

$$S_{[G,K]}^m = \begin{bmatrix} 1 & \dots & 1 \\ \exp(j2\pi f_{m_1}) & \dots & \exp(j2\pi f_{m_K}) \\ \vdots & \dots & \vdots \\ \exp(j2\pi f_{m_1}(G-1)) & \dots & \exp(j2\pi f_{m_K}(G-1)) \end{bmatrix} \quad (18)$$

Où  $m$  est la dimension spatiale  $m = 1, 2, 3$ ; et  $G$  la taille des cumulants relatifs à la dimension  $m$ , c.-à-d.  $G \in \{P, Q, L\}$

La matrice diagonale est donnée par:

$$\Phi_m = \text{diag}(\exp(j2\pi f_{mi}))_{1 \leq i \leq K} \quad (19)$$

On considère la décomposition en valeurs propres de la matrice  $C_{y,1}$ :

$$C_{y,1} = UDU^H \quad (20)$$

Avec

$$U = [u_1, \dots, u_K, u_{K+1}, \dots, u_{PQL}] \quad \text{et} \quad D = \text{diag}(\lambda_i)_{1 \leq i \leq PQL} \quad (21)$$

Où les valeurs propres obtenus sont réelles et sont classées par ordre décroissant en vérifiant la relation suivante:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{K+1} = \dots = \lambda_{PQL} = 0 \quad (22)$$

Ainsi, le sous-espace signal est engendré à la fois par les colonnes de la matrice  $S_{[PQL,K]}^1$  et par les vecteurs propres de  $C_{y,1}$  associés aux  $K$  valeurs propres non nulles  $U_{S_1}$ . Cela entraîne l'existence d'une matrice inversible  $\Theta_1$  de taille  $K \times K$  vérifiant la relation:

$$U_{S_1} = S_{[PQL,K]}^1 \Theta_1 \quad (23)$$

La matrice de Vandermonde  $S_{[PQL,K]}^1$  peut être partitionnée de deux manières différentes comme dans la relation suivante:

$$S_{[PQL,K]}^1 = \begin{bmatrix} S_{[PQ,K]}^1 \\ \vdots \\ EM1 \end{bmatrix} \updownarrow PQ = \begin{bmatrix} EM1 \Phi_3^{-1} \\ \vdots \\ S_{[PQ,K]}^1 \Phi_3^{L-1} \end{bmatrix} \updownarrow PQ \quad (24)$$

De même on considère les partitionnements pour la matrice accessibles  $U_{S_1}$

$$U_{S_1} = \begin{bmatrix} xxx \\ \vdots \\ \overline{U_{S_1}} \end{bmatrix} \updownarrow PQ = \begin{bmatrix} \underline{U_{S_1}} \\ \vdots \\ xxx \end{bmatrix} \updownarrow PQ \quad (25)$$

La matrice  $\Theta_1$  peut être utilisée de la manière suivante:

$$\begin{cases} EM1 \Theta_1 = \overline{U_{S_1}} \\ EM1 \Phi_3^{-1} \Theta_1 = \underline{U_{S_1}} \end{cases} \Rightarrow \begin{cases} EM1 = \overline{U_{S_1}} \Theta_1^{-1} \\ \overline{U_{S_1}} \Theta_1^{-1} \Phi_3^{-1} \Theta_1 = \underline{U_{S_1}} \end{cases} \quad (26)$$

$$\Rightarrow \Theta_1^{-1} \Phi_3^{-1} \Theta_1 = \overline{U_{S_1}}^\dagger \underline{U_{S_1}} = F_3$$

$^\dagger$  désigne l'opérateur pseudo inverse. Par conséquent, les fréquences  $\{f_{3_i}\}$  contenues dans la matrice  $\Phi_3$  seront estimées à partir des valeurs propres de la matrice  $F_1 = \overline{U_{S_1}}^\dagger \underline{U_{S_1}}$  par:

$$f_{1i} = \frac{1}{2\pi} \text{Im}[\log(\lambda_i [F_3])] \quad (27)$$

### C. Estimation des fréquences suivant la deuxième dimension

Pour estimer les fréquences associées à la première dimension, nous introduisons une nouvelle matrice de cumulants  $C_{y,2}$  de type TBBT construite comme précédemment. En effet, il entraîne l'existence d'une matrice inversible  $\Theta_2$  de taille  $K \times K$  vérifiant la relation:

$$Us_2 = S_{[QLP,K]}^2 \Theta_2 \quad (28)$$

La matrice de Vandermonde 3-D  $S_{[QLP,K]}^2$  est donnée par:

$$S_{[QLP,K]}^2 = \begin{bmatrix} S_{[QL,K]}^2 \\ S_{[QL,K]}^2 \Phi_1 \\ \vdots \\ S_{[QL,K]}^2 \Phi_1^{P-1} \end{bmatrix} \text{ et } S_{[QL,K]}^2 = \begin{bmatrix} S_{[Q,K]}^2 \\ S_{[Q,K]}^2 \Phi_3 \\ \vdots \\ S_{[Q,K]}^2 \Phi_3^{L-1} \end{bmatrix} \quad (29)$$

Avec  $S_{[QL,K]}^2$  et  $S_{[Q,K]}^2$  sont les matrices de Vandermonde 2-D et 1-D respectivement.

Or le produit de Kronecker n'est commutatif, donc les matrices  $S_{[PQL,K]}^1$  et  $S_{[QLP,K]}^2$  sont reliées par la relation suivante:

$$S_{[QLP,K]}^2 = E_1^2 S_{[PQL,K]}^1 \quad (30)$$

Avec  $E_1^2$  est la matrice de permutation donnée par:

$$E_1^2 = \sum_{i=1}^P \sum_{j=1}^Q \sum_{k=1}^L E_{i,j}^{P,Q} \otimes E_{j,k}^{Q,L} \otimes E_{k,i}^{L,P} \quad (31)$$

Et  $E_{i,j}^{P,Q}$  est la matrice de permutation élémentaire de taille  $P \times Q$  ayant la valeur 1 pour les coordonnées  $(i, j)$  et zéros ailleurs.

De même pour la matrice  $Us_2$

$$Us_2 = E_1^2 Us_1 \quad (32)$$

Les fréquences  $\{f_{2i}\}$  contenues dans la matrice  $\Phi_1$  seront estimées à partir des valeurs propres de la matrice

$F_2 = \bar{U}s_2^\dagger Us_2$  par:

$$f_{2i} = \frac{1}{2\pi} \text{Im}[\log(\lambda_i[F_2])] \quad (33)$$

### D. Estimation des fréquences suivant la troisième dimension

Comme précédemment, pour estimer les fréquences de la deuxième dimension, on va construire la matrice des cumulants  $C_{y,3}$  de type TBBT avec  $\Theta_3$  une matrice de taille  $K \times K$  vérifiant la relation suivante:

$$Us_3 = S_{[QLP,K]}^3 \Theta_3 \quad (34)$$

La matrice de Vandermonde 3-D  $S_{[LPQ,K]}^3$  est donnée par:

$$S_{[LPQ,K]}^3 = \begin{bmatrix} S_{[LP,K]}^3 \\ S_{[LP,K]}^3 \Phi_2 \\ \vdots \\ S_{[LP,K]}^3 \Phi_2^{Q-1} \end{bmatrix} \text{ et } S_{[LP,K]}^3 = \begin{bmatrix} S_{[L,K]}^3 \\ S_{[L,K]}^3 \Phi_1 \\ \vdots \\ S_{[L,K]}^3 \Phi_1^{P-1} \end{bmatrix} \quad (35)$$

Avec  $S_{[LP,K]}^3$  et  $S_{[L,K]}^3$  sont les matrices de Vandermonde 2-D et 1-D respectivement.

De même, on a les relations suivantes:

$$S_{[LPQ,K]}^3 = E_2^3 S_{[QLP,K]}^2 \quad (36)$$

$$Us_3 = E_2^3 Us_2 \quad (37)$$

Avec  $E_2^3$  est la matrice de permutation donnée par:

$$E_2^3 = \sum_{i=1}^Q \sum_{j=1}^L \sum_{k=1}^P E_{i,j}^{Q,L} \otimes E_{j,k}^{L,P} \otimes E_{k,i}^{P,Q} \quad (38)$$

Finalement, les fréquences  $\{f_{3i}\}$  contenues dans la matrice  $\Phi_2$  seront estimées à partir des valeurs propres de la matrice  $F_3 = \bar{U}s_3^\dagger Us_3$  par:

$$f_{3i} = \frac{1}{2\pi} \text{Im}[\log(\lambda_i[F_3])] \quad (39)$$

A ce niveau, les fréquences de chaque dimension sont estimées, ainsi l'étape de mise en triplet entre en jeu. En effet, on considère une matrice  $\Theta$  de taille  $K \times K$  vérifiant:

$$\Theta_1 = \Theta_2 = \Theta_3 = \Theta \quad (40)$$

Puis on construit une matrice  $F$  à partir des matrices  $F_1$ ,  $F_2$  et  $F_3$  telle que:

$$F = \alpha_1 F_1 + \alpha_2 F_2 + (1 - (\alpha_1 + \alpha_2)) F_3 = \Theta^{-1} \Delta \Theta \quad (41)$$

Où  $\alpha_1$  et  $\alpha_2$  sont des scalaires.

Dans [8] on a proposé une nouvelle méthode ESPRIT 3-D pour estimer les triplets fréquentiels s'il s'agit des multiples. En effet, on construit trois matrices de permutation  $P_1$ ,  $P_2$  et  $P_3$  telles que:

$$\begin{cases} P_1 = \Theta^{-1}\Theta_1 \\ P_2 = \Theta^{-1}\Theta_2 \\ P_3 = \Theta^{-1}\Theta_3 \end{cases} \quad (42)$$

D'où on aura les relations suivantes:

$$\begin{cases} \Phi'_1 = P_1^{-1}\Phi_1P_1 \\ \Phi'_2 = P_2^{-1}\Phi_2P_2 \\ \Phi'_3 = P_3^{-1}\Phi_3P_3 \end{cases} \quad (43)$$

Donc les nouvelles matrices  $F_1$ ,  $F_2$  et  $F_3$  deviennent:

$$\begin{cases} F'_1 = \Theta^{-1}\Phi_1^{-1}\Theta \\ F'_2 = \Theta^{-1}\Phi_2^{-1}\Theta \\ F'_3 = \Theta^{-1}\Phi_3^{-1}\Theta \end{cases} \quad (44)$$

Et finalement les fréquences  $\{f_{1i}\}$ ,  $\{f_{2i}\}$  et  $\{f_{3i}\}$  contenues dans les matrices  $\Phi'_1$ ,  $\Phi'_2$  et  $\Phi'_3$  seront estimées à partir des valeurs propres des matrices  $F'_1$ ,  $F'_2$  et  $F'_3$  par

$$\begin{cases} f'_{1i} = \frac{1}{2\pi} \text{Im}[\log(\lambda_i[F'_1])] \\ f'_{2i} = \frac{1}{2\pi} \text{Im}[\log(\lambda_i[F'_2])] \\ f'_{3i} = \frac{1}{2\pi} \text{Im}[\log(\lambda_i[F'_3])] \end{cases} \quad (45)$$

#### IV. EXEMPLE DE SIMULATION

Dans cette section, on présente quelques exemples numériques de simulation. Notre approche est testée sur un modèle SEC 3-D. Les données sont générées en fonction du modèle de l'équation (1). On considère trois modes c.à.d.  $K=3$  avec l'amplitude  $a_i=200$ , les fréquences 3-D sont données dans le tableau I. Les données et les tailles de la matrice des Cumulants sont respectivement  $(M, N, T) = (32, 32, 32)$ , et  $(P, Q, L) = (3, 3, 3)$ . Le bruit gaussien coloré est obtenu en filtrant le bruit gaussien blanc par un filtre AR 3-D [12] [13]. On va

considérer les valeurs de rapport signal sur bruit SNR=20 dB.

TABLE I

LES FREQUENCES 3-D UTILISEES DANS L'EXEMPLE DE SIMULATION

	$f_{1i}$	$f_{2i}$	$f_{3i}$
1 <sup>st</sup> wave	0.1100	0.1400	0.1700
2 <sup>nd</sup> wave	0.2400	0.2300	0.2100
3 <sup>rd</sup> wave	0.4500	0.4800	0.4700

TABLE II

FREQUENCE ESTIMEES AVEC SNR=20 DB

	$f_{1i}$		$f_{2i}$		$f_{3i}$	
	SOD	SOS	SOD	SOS	SOD	SOS
1 <sup>er</sup> mode	0.1102	0.1092	0.1399	0.1390	0.1701	0.1681
2 <sup>o</sup> mode	0.2401	0.2386	0.2303	0.2270	0.2098	0.2090
3 <sup>o</sup> mode	0.4500	0.4492	0.4798	0.4779	0.4702	0.4689

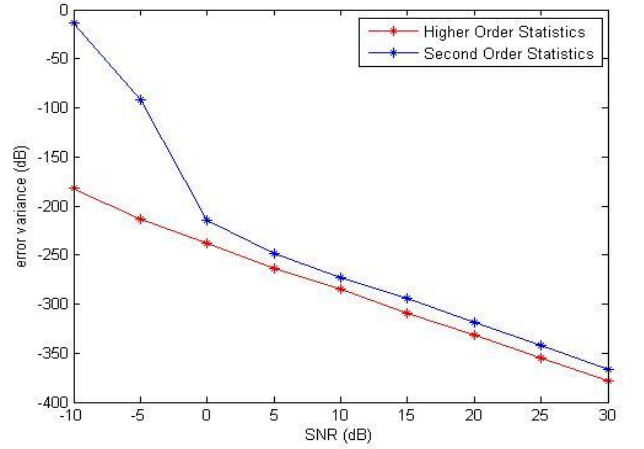


Fig. 1: Estimation-error variance versus the SNR.

#### V. CONCLUSION

Dans ce papier, nous avons montré l'intérêt des statistiques d'ordre supérieur par rapport aux statistiques d'ordre deux. En effet, l'estimation des fréquences 3-D par la nouvelle méthode ESPRIT 3-D lorsque le bruit est gaussien coloré et en utilisant les cumulants donne de bonnes performances que les autocorrélations.

## REFERENCES

- [1] Y. Hua, "Estimating Two-Dimensional Frequencies by Matrix Enhancement and Matrix Pencil", *IEEE Transactions on Signal Processing*, Vol. 40, No 9, September 1997.
- [2] B. Aksasse, M. Elansari, Y. Berthoumieu, and M. Najim, "High resolution 3-D spectral method estimation", in *Proc EUSIPCO 2002*, Vol. 2, pp. 391-394, Sept 03-06, Toulouse, France.
- [3] Y. Stitou, B. Aksasse and M. Najim, "Cumulant matrix pencil method for three-dimensional frequency estimation in the colored Gaussian noise", *proc IEEE/EURASIP ISCCSP*, Marrakech, Morocco, March 2006.
- [4] P. Vanpoucke, M. Moonen, Y. Berthoumieu, "An efficient subspace algorithm for 2-D harmonic retrieval", *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 461-464, 1994.
- [5] R. Roy, R and T. Kailath, "Esprit - Estimation Of Signal Parameters Via Rotational Invariance Techniques". *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 37, No. 7, pp. 984-995, 1989.
- [6] S. Rouquette and M. Najim, "Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods" *IEEE Trans. Signal Process.*, Vol. 49, No. 1, pp. 237-245, 2001.
- [7] Y. Berthoumieu, M. El Ansari, B. Aksasse, M Donias, M. Najim, "A 2-D Robust high resolution frequency estimation approach" *Signal Processing*, Vol. 85, No. 6, pp. 1165-1188, June 2005.
- [8] Y.Chawki, M.Ouanan and B.Aksasse "Three-Dimensional High-Resolution Spectral Analysis: New method of estimating 3-D frequencies", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 3, No. 11, pp. 1426-1432, November - 2013.
- [9] A. Swami and J.M. Mendel, "Cumulant based approach to the harmonic retrieval and related problem", *IEEE Trans. Acoust., Speech, Signal, Processing*, 5(39): pp. 1099-1109, May 1991.
- [10] H. M. Ibrahim, R. R. Ghariieb, "Estimating two-dimensional frequencies by a cumulant-based FBLP method", *IEEE Transactions on Signal Processing*, 1(47) : 262-266, January 1999.
- [11] Y. Stitou, F.Turcu, M. Najim and L. Radouane, "3-D Texture characterization based on Wold decomposition and higher order Statistics", *Dans Proc IEEE ICASSP05*, Vol. 2, pp. 165-168, March 18-23, Philadelphia, USA, 2005.
- [12] B. Aksasse, Y. Stitou and M. Najim, "Minimum eigenvalue based 3-D AR model", *IEEE, Statistical Signal Processing SSP'05 Workshop Bordeaux 2005*.
- [13] B. Aksasse, Y. Stitou, Y. Berthoumieu and M. Najim, "3-D AR Model Order Section via Rank Test Procedur", *IEEE Transactions on Signal Processing*, Vol. 54, No 7, July 2006.



<b>A</b>		<b>E</b>		<b>M</b>	
Adnane B. A.	90	El Hannachi. S. A	6	Mahi. H	51, 65
Ahmed-Foitih. Z	90			Maigret. B	80
Amara S.	107	<b>F</b>		Masmoudi. M	94
Ameur. S	111, 115	Fadhel. S	103	Mazari. C	139
Ameur. Z	111	Fizazi. H	65, 73	Medjahed .A.	107
Azizi. N	56			Merad. B. S	42
				Messabih B.	80
<b>B</b>		<b>G</b>		Mokrani. A	1
Baddeley. R	1	Gacem. A	42	Moussaoui. A	12, 85
Bekhouché. S. E	18				
Belhadj. S	90	<b>H</b>		<b>N</b>	
Belkhodja. L	103	Hadid. A	18	Nibouche. M	1
Benaïssa. M		Hamdadou. D	103		
		Hameurlaine. M	12	<b>O</b>	
Benali. A	90	Hammoudi. A	61	Ouafi. A	18
Benamrane. N	6	Houidène F.	134	Oulad Kaddour. M	99
Benatchba. K	99				
Bendani. D	122	<b>I</b>		<b>R</b>	
Bendouda. M	47	Issaad. D	107	Rabhi. S	73
Benkkadour. M. K	28			<b>S</b>	
Benlamoudi. A	18	<b>K</b>		Saadi Bachir. M	
Benouis. M	28	Kaddour El Boudadi. L	94	Samai. D	18
Berkani. D	34	Kaouadji. M	51	Senouci. M	28
		Keddi. A	103	Slimani. K	111
Berrached. N	42, 47, 122	Khelf. M	94		
Bouafia. M.	107	Khensous. G	80	<b>T</b>	
Boudjemaa F.	134	Kherchi .A. M	23	Taleb-Ahmed. A	18, 90
Bouhali. A	34				
Boukendourr. O	134	<b>L</b>			
Brahimi. M		Labed. K	65		
		Lahdir. M	115		
<b>C</b>		Lamiche. C	85		
Chalabi. Z	122	Lamy De La Chapelle M.	107		
Cherifi. M	115	Larabi. S	23		
Cheriguene. S	56	Loukil. A	94		
Cherroun. H	12				
Chouarfia. A	80				
<b>D</b>					
Dahane. A	94				
Diaf. M	139, 134				
Djaker N.	107				
Djeddou. F	61				