2012 International Conference on Frontiers in Handwriting Recognition

A Study of Handwritten Characters by Shape Descriptors: Doping Using the Freeman Code

Chekib Gmati
SITI Lab
ENIT
Tunis-Belvedere, Tunisia
chekibgmt2007@yahoo.fr

Sofiene Haboubi
SITI Lab
ENIT
Tunis-Belvedere, Tunisia
sofiene.haboubi@istmt.rnu.tn

Abdulqadir Alaqeeli
CRI
KACST
Riyadh, Saudi Arabia
aalaqeeli@kacst.edu.sa

Hamid Amiri
SITI Lab
ENIT
Tunis-Belvedere, Tunisia
hamidlamiri@yahoo.com

***Abstract*— In this paper, we present the role of shape descriptors in the off-line recognition process of handwritten isolated Arabic and Latin characters. We will give some statistical and structural shape descriptors and mention their performance. Then we will present an hybrid approach that uses structural shapes' descriptors from a different angle in order to improve the recognition's results from a statistical descriptor. We will therefore introduce the concept of doping aiming to raise the recognition rate.**

Keywords-shape descriptors; statistical approaches;structural approaches; freeman code; fourier descriptor; doping

I. Introduction

The basic rules of school education, for learning writing and reading, explain why the handwriting was not invaded by the scan despite its scope.

There is a very close relationship between the writer and handwriting. The study of this relationship allows us to know more about the personality of the writer. This justifies the usefulness of the study of handwriting in areas such as biomechanics, biometrics [1] and graphology [2] [3]. Also, the study of handwriting is the subject of several researches and its importance increases.

The study of handwriting has been translated by the implementation of “handwritten characters recognition systems”. The architecture of these systems comprises three main stages: the preprocessing, the feature extraction and the classification stage. The step “key” in the recognition process is the feature extraction step because it directly affects the results of the recognition system. The extraction of primitives is carried out through shape descriptors. These descriptors are the “keystone” of the handwritten character recognition system.

They are divided into two categories: statistical and structural descriptors [4].

II. Background

A. Statistical descriptors

The statistical methods perform statistical measures that lead to describe the shape in a comprehensive manner. Descriptors that adopt these statistical approaches are simple geometric descriptors, based on moments, Fourier descriptors and descriptors based on a transform of the image.

Simple geometric descriptors are fast in computation time. They have several methods such as the degree of ellipticity or degree of squareness which is the ratio of the surface with the minimum bounding box [5]. We can mention others, like the eccentricity E and C the convexity of the object:

E= Length of the major axis / Length of the minor axis (1)

C= Perimeter of the convex hull / P (2)

Descriptors based on moments are divided into two categories: the moments of orthogonal polynomials and moments non-orthogonal.

Moments of orthogonal polynomials can be continuous or discrete. Among the continuous time in this category, we include the Zernike moments that have been proposed by [6]. This descriptor is made up of complex polynomials which form a complete orthogonal set, defined on the unit disk $x^2+y^2=1$.

This descriptor is considered a descriptor robust with high descriptive power [7]. The computation time of Zernike moments is very important. Several searches were conducted in order to reduce complexity and increase the performance of Zernike moments [8] [9].

We can cite also the moments of Fourier_Mellin [10] The Pseudo-Zernike moments of generalized [11]. In the category of discrete moments of orthogonal polynomials, we quote the moments of Tchebichef [12] and Krawtchouk moments [13]. Orthogonal moments are remarkably efficient than no-orthogonal moments, a comparison between the descriptors based on moments has been reported in [14] [15].The Fourier descriptors are presented in [16] descriptors by tangent or complex representation. The complex representation is to represent the contour points of the form in the complex plane. Fourier descriptors are the coefficients of Z-transform follows:

$$Z_k = \sum_{j=1}^{N} z_j e^{(-2\pi ijk)} \quad (3)$$

Where $z_j = x_j + iy_j$ With (x_j, y_j) the coordinates of $\{P_j\}$ the contour of the object, and in the complex plane Z is

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.170

the Z -transform of Z. Ask increases, the coefficients Z_k are more numerous and therefore, the representation of the shape of the object is increasingly fine (Figure 1).

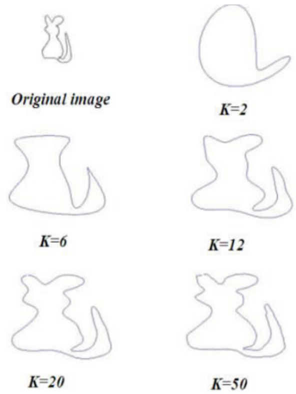

Figure 1. Pattern recognition of a synthetic object by Fourier descriptor

Our understanding of the importance of the parameter k by the following discussion:

If k=0,

$$Z_0 = \sum_{j=1}^{N} z_j \tag{4}$$

The sum of the entire complex points of a shape is the center of gravity.

If k=1,

$$Z_1 = \sum_{j=1}^{N} z_j e^{(-2\pi ij)} \tag{5}$$

It represent a circle of radius Z_1.

If k is different from 1 and 0, in which case the order k is the action of tension and pressure on the curve, respectively, for k> 0 and k <0, i.e. the phase of the complex number Z_k shows level of the action on the circle of unity. Descriptors based on a transform of the image has the Radon transform [18], the Hough transform [19]. There are other methods that adopt the time-frequency representation like Gabor filtering method.

A Gabor filter is a sinusoidal function modulated by a Gaussian envelope which is characterized by its frequency and its orientation. It preserves the temporal aspects and frequency of the signal. In space, the application of Gabor filters is performed by calculating the convolution of the image with a function adjusted to a texture. The filtering is done by a convolution product of the Gabor filter with the pixel of the original image [26].

B. *Structural Descriptors*

The structural methods are based on a description of structural type, treating the form as a set of elementary part and is presented as a structure. Among these descriptors, we quote the grammars [20], methods based on graphs [21], methods based on skeletonization ([16], [22]) vectorization [23], and methods based on edge detection. These descriptors encoding attribute to the contour of the object, of which we can cite the Freeman Code, which represented, from coordinates of a starting point and a chain coding giving the relative position of the next point of the outline.

Using 4-connectivity, the distance between the center and the four surrounding points is a pixel, the encoding of the next pixel to the right from the center is 0, the left one is 3, the top one is 1 and the bottom is 4. In the case of 8-connected, the distance between the center pixel and that of the diagonal (direction 1, 3, 5 or 7) is v2 pixel. Knowing that the directions are presented in the counterclockwise direction, the codes and the 8-connected contour reconstructed by the Freeman coding are shown in Figure 2.

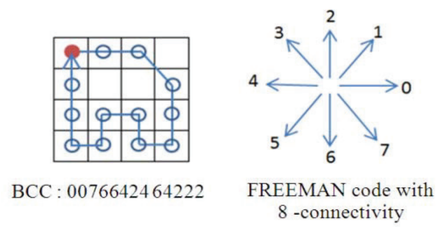

Figure 2. Representation of a form by BCC.

The properties of Freeman chains include the reduction, simplification of the road, closing a contour, calculating the length of a contour reconstructed by the Freeman chain, change of origin and the direction of travel of a contour [16]. Using the same principle, coding of LIU and ZALIC based on the variation of angle between two coding [24].

III. MOTIVATION

The feature extraction is an important step on which depend all the succeeding steps. Indeed if we consider the case where the primitive vectors are not discriminating, any classifier can compensate for this gap. The feature extraction, obtained by the shape descriptors, is strongly linked to the classification stage. Applying a statistical classification method to a large number of primitives allows achieving a learning phase longer in terms of computation time.

The obtainment of a primitive vector does not mean, necessarily, that recognition is evident.

The performance of shape descriptors depends on the usefulness of the primitive vector in the classification stage. That usefulness is defined by the size of the vector and the pertinence of its primitives so that the large size and the irrelevance of its primitives, prevents a vector from being useful for the classification stage. Thus, the main objective is to find the balance between “not having a vector of size too large” and “having relevant vector primitives”.

Nevertheless, when the primitives vector is too small, a reduction in the vector's discrimination leads us to a misclassification and consequently to a false recognition. Indeed, the goal is to reduce the computation time and increase inter-class distance. To do this, the vector must express the “singularity” of the character without taking into account the specific graphical own and unique of the object to recognize which allows to represent the characteristics of references models of the same class letter.

Our contribution consists in increasing the relevance degree of primitives the vector without increasing the size of the vector; this still depends on the invariance to the translation, rotation, and scaling.

IV. CONTRIBUTION

A. Preprocessing

The preprocessing is an essential step in the recognition process. In this step we performed binarization [25], denoising and skeletonization [16].

The skeletonization generates noise, or parasitic components, which are eliminated by the pruning. Pruning is the most important step in our method. Indeed, it allows us to determine the intersection points and the endpoints from the denoised pruned skeleton of handwritten character, in order to segment them correctly and have the sub-components which define the object. So after the preprocessing step, we obtain the sub-components of the handwritten character skeleton which are ready for the application of freeman coding.

B. Feature extraction

The method is to increase the degree of the feature vector relevance generated by the Fourier descriptor. To do this we chose to concatenate this vector by another vector. The attributes of the second vector is the results of the interpretation of sub-components by the freeman coding after the segmentation of handwritten character. We'll add to second vector two values for increase the relevance: the number of intersection points and the number of endpoints (Figure 3).

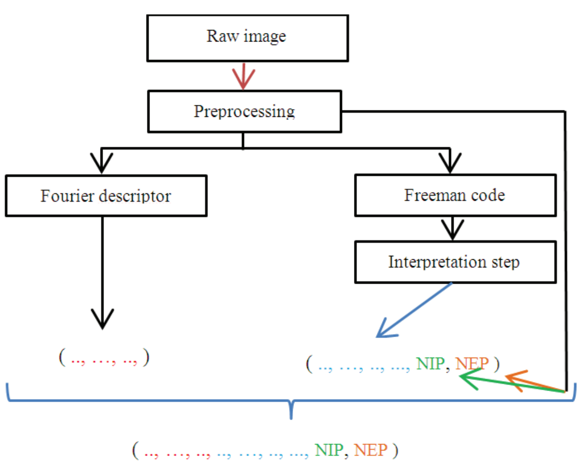

Figure 3. Doping method based on the Freeman Code

With "NIP" represents the number of intersection points and "NEP" is the number of endpoints.

Concerning the Freeman coding, it will be applied on sub-components of the object after its “segmentation”.

The segmentation of the object is made by the zeroing pixels values of intersection points and their neighbor pixels. Thereby, we will get sub-components on which we will apply, for each one, the Freeman coding (Figure 4).

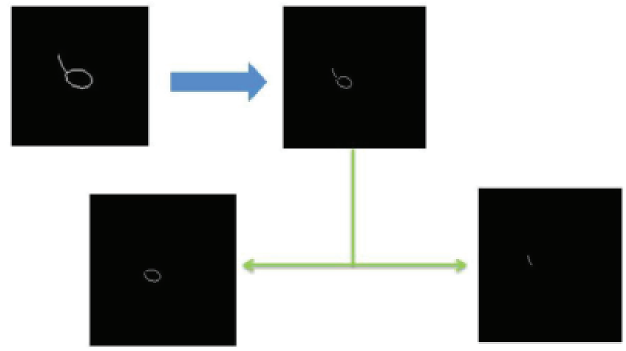

Figure 4. Segmentation of the character b by the zeroing pixels values of the intersection points

After segmentation and freeman coding, we will interpret the Freeman chain codes for each sub-component by the following:

- For each two successive elements of Freeman chain code, we can retain information, denoted i, which is the direction of angles variations carried out on the object contour path and not exceeding two values {1,-1}. Note cod1 and cod2 two successive elements of Freeman chain code:

  if cod1 < cod2 then i = 1.

  if cod1 > cod2 then i = -1.

- At this stage we must identify the dominant sense of the angles variations’ direction. To do this, we remove each two successive values of opposite signs. In Figure 5, we clearly illustrate this method:

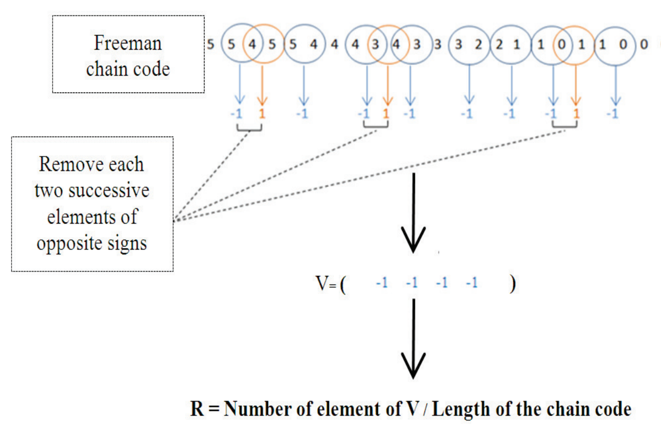

Figure 5. Method of interpretation of the Freeman chain code

The resulting vector, denoted V, will have all elements equal to 1 or -1. This vector contains elements that indicate the dominant direction of angles variations, allowing us to calculate the ratio R.

The ratio R is equals to the number of vector V elements divided by the length of Freeman chain code. This report provides information on the degree of convexity of each sub-component.

The segmentation generates the same sub-components in rotation, translation and scaling. Therefore, each sub-component keeps in all three cases, the same ratio R or the same degree of convexity.

V. RESULTS

The database used is composed of Latin and Arabic handwritten characters. These characters do not contain diacritical signs. On Latin characters, Have We analyze the lowercase. The database used in this study was collected from a population (with 52 people) distributed as follows:

Age :

- from 11 years to 20 years : 21,15%.
- from 21 years to 30 years : 59,61%
- from 31 years to 60 years : 19,24%

Function:

- student : 53,84%
- employed : 46,16%

Sexe:

- male :53,76%.
- female 46,24%.

Our experimental approach is to apply several classification methods on samples obtained from the step of feature extraction and this by methods validation. We seek, by analyzing the results find the added value of the doping vector.

We adopted, through WEKA, five classification algorithms and three learning methods.

Classification algorithms are:

- Bayesian probabilistic classification: « Naïves Bayes ».
- Decision tree: « J48 », « Random Forest », « Random Tree ».
- Neural Networks: « Multiplayer Perceptron ».

Validation methods are:

- «Use training set» : all data are used both to learn and test models.
- «Cross-validation» : The training set is split into 10. The algorithm will learn 10 times out of nine parties and the model will be evaluated on the remaining tenth, the 10 evaluations are then combined.
- «Percentage split» : to learn once x% and test data on (100-x) % the remaining data. In our case, we used 80% for learning.

We know that the relevance degree of the samples affect directly the classification's results. The curves below illustrate variation in the rate of correct classification for each validation method. They show the performance of the doping method compared to other descriptors.

TABLE I. RATE VARIATION IN THE CORRECT CLASSIFICATION METHOD VALIDATION: "USE TRAINING SET"

	Freeman Code	*Freeman Code + Fourier descriptor*	*Gabor filter*	*Fourier descriptor*
J 48	66.66	98.45	71.31	90.5
Random Forest	97.24	99.89	99.41	99.89
Random Tree	97.68	100	100	100
Naive Bayes	25.49	78.47	14.16	24.85
Multilayer Perceptron	39.18	97.24	52.76	80.13

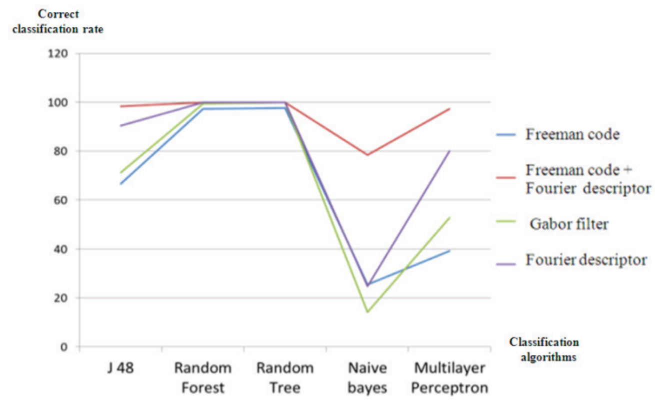

Figure 6. Variation curves of the correct classification rate in the validation method: "Use Training Set"

TABLE II. RATE VARIATION IN THE CORRECT CLASSIFICATION METHOD VALIDATION: " CROSSVALIDATION "

	Freeman Code	*Freeman Code + Fourier descriptor*	*Gabor filter*	*Fourier descriptor*
J 48	34.98	93.7	9.7	47.9
Random Forest	34.98	85.98	12.67	60.92
Random Tree	27.48	57.72	10.45	26.49
Naive Bayes	22.07	76.49	13.24	13.24
Multilayer Perceptron	32.34	86.86	18	52.42

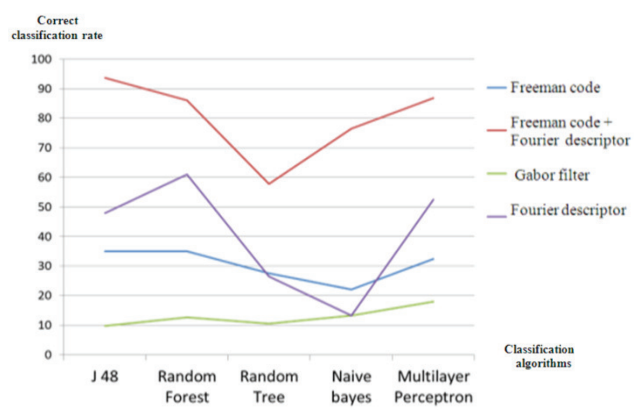

Figure 7. Variation curves of the correct classification rate in the validation method: "Cross validation"

TABLE III. RATE VARIATION IN THE CORRECT CLASSIFICATION METHOD VALIDATION: " PERCENTAGE SPLIT "

	Freeman Code	*Freeman Code + Fourier descriptor*	*Gabor filter*	*Fourier descriptor*
J 48	37.56	92.26	6.97	46.4
Random Forest	34.25	83.42	9.3	56.9
Random Tree	25.41	49.17	6.97	26.51
Naive Bayes	20.99	77.34	11.04	44.19
Multilayer Perceptron	29.28	87.29	12.79	54.69

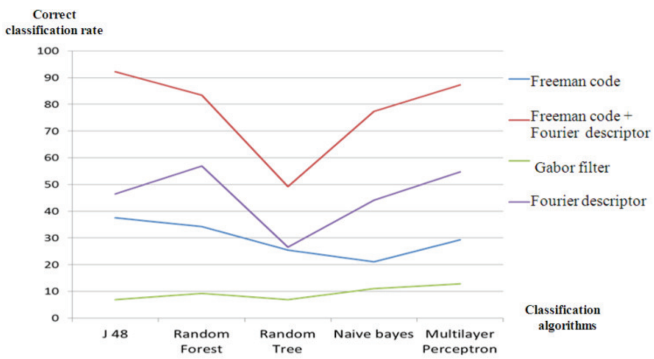

Figure 8. Variation curves of the correct classification rate in the validation method: "Percentage Split"

We notice that when we have used the method of "doping", the results were, visibly, a success. Indeed, by arranging the Freeman Code and the Fourier descriptor, we have reached a percentage of correct classification from 57.72% to 100% with the exception of a single recorded case in the use of algorithm "Random tree" as part of the validation method "percentage split" that does not exceed 49.17%.Then, we remark that the best results have been achieved through the validation method "use training set" with the algorithms "Random tree" and "Random Forest" (results between 97.24% and 100%).

For the remaining methods the highest correct results and the lowest incorrect results were achieved with the method of "doping". By contrast, the highest percentages of incorrect results and the lowest percentages of correct results were attained with the method based on Gabor filter.

Moreover, the use of the Freeman Code led to satisfactory results. But, in comparison with the method of "doping", these results are debatable and insufficient.

In addition, we find that the Freeman Code and the Fourier descriptor have generated incorrect results with higher percentages than the correct results that haven't themselves reached the level of the correct results collected through the method of "doping". In Figure 9, below, we present the curve showing the contribution of our method.

In fine, we register, through our contribution that the method of "doping", leads to the most efficiently correct results with remarkably high percentages. By contrast, the use of one method based on the Freeman Code Is imperfect and less efficient.

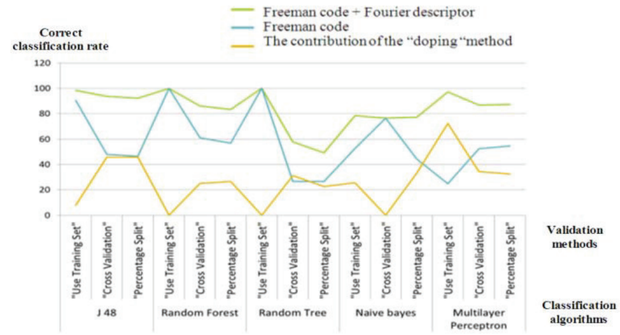

Figure 9. Contribution of the method of "doping"

VI. CONCLUSION

We find that the resulting vector of our approach has a higher relevant degree than features vectors generated by the other used shape descriptors. This is achieved by using a vector "doping" based on a structural descriptor: Freeman Code. The idea is to take advantage of both statistical and structural approaches, i.e. combine the overall and elementary representation of handwritten character. The approach is based on the interpretation of Freeman coding applied on sub-components of the object. This method can be adapted to the recognition of signatures, words, text and all other complex patterns outside of handwriting. In this case we can define the complexity of the patterns by the number of intersection points present in these patterns.

REFERENCES

[1] S. Lacour : La sécurité aujourd'hui dans la société de l'information, Harmattan edition, pp 69-96, 2007.

[2] M. Sardin : La graphologie , Eyrolles edition, 2010.

[3] D. De Castilla : Idée reçue : La graphologie , Le Cavalier Bleu, coll. Idées Reçues, 2001.

[4] L.Heutte, T.Paquet, J.V.Moreau, Y.Lecourtier, C.Olivier: A structural/statistical feature based vector for handwritten character recognition, Pattern recognition letters, pp.629-641, 1998.

[5] P. L. Rosin : Measuring shape: ellipticity, rectangularity, and triangularity , Mach Vis App, 14, 3, 172-184, 2003.

[6] M.R. Teague: Image analysis via the general theory of moments , J. Optical Soc. Am., pp.920–930, 1980.

[7] J. Revaud : Une nouvelle mesure de distance entre descripteurs de moments de Zernike pour une similarité optimale et un angle de rotation entre les images, CORESA – Toulouse, University of Lyon, 2009.

[8] S.O. Belkasim, M. Shridar, M. Ahmadi: Pattern recognition with moment invariants: a comparative study and new results, Patt Recog, 24, pp. 1117–1138, 1991.

[9] L. Kotoulas, I. Andreadis: An efficient Technique for the Computation of ART. IEEE Transactions on circuits and systems for video technology, vol. 18, no. 5, pp. 682–686, 2008.

[10] H. Zhang and al: Construction of a complete set of orthogonal Fourier-Mellin moment invariants for pattern recognition applications, Image and Vision Computing Volume 28, Issue 1, 2010.

[11] T. Xia, H. Zhu, H. Shu, P. Haigron, L. Luo: Image description with generalized pseudo-Zernike moments, Optics, image science, and vision, 24, pp. 50–59, 2007.

[12] R. Mukundan, S.H. Ong, P.A. Lee: Image analysis by Tchebichef moments, IEEE Transactions on Image Processing, vol. 10, no. 9, pp. 1357–1364, 2001.

[13] P.-T. Yap, R. Paramesran, S.-H. Ong: Image Analysis by Krawtchouk Moments, IEEE Transactions on Image Processing, vol. 12, no. 11, pp. 1367–1377, 2003.

[14] H.Zhua et al: Image analysis by discrete orthogonal dual Hahn moments, Pattern Recognition Letters, pp.1688-1704, 2007.

[15] C.-H. The, R.T: Chin. On image analysis by the methods of moments, IEEE Transactions on Pattern Analy Mach Intell, vol. 4, no. 4, pp. 496 – 513, 1988.

[16] H. Maitre : Le traitement des images, Hermes edition science publications – lavoisier, 2003.

[17] Ch. Chettaoui et al : Etude de formes des globules drépanocytaires par traitement numérique des images », Sciences of Electronic, Technologies of Information and Telecommunications SETIT 2005, 2005.

[18] N. Nacereddine, S. Tabbone, D. Ziou, L. Hamami : Un descripteur efficace pour la reconnaissance des symboles graphiques basé sur la transformée de Radon, Symp on Writing and Document pp. 201-216, 2010.

[19] E. Tisserand, J. Pautex, P. Schweitzer : Analyse et traitement des signaux Méthodes et applications au son et à l'image, second edition Algeria-Educ.com, DUNOD edition, Paris, pp287- 288, 2008.

[20] Y. Es Saady, M. Amrouch, A. Rachidi, M. El Yassa, D. Mammass: Reconnaissance de Caractères Amazighes Imprimés par le Formalisme des Automates à états finis, published in SITACAM'09, Agadir-Maroc, 2009.

[21] D. Arrivault : Apport des Graphes dans la Reconnaissance Non-Contrainte de Caractères Manuscrits Anciens, Thesis, Poitiers University, 2006.

[22] Y.Y. Zhang, P.S.P. Wang: A new parallel thinning methodology, Patt Recog Arti Intell, pp.999–1011, 1994.

[23] L.Wenyin, W.Zhang, L.Yan : An interactive example-driven approach to graphics recognition in engineering drawings », IJDAR, vol. 9, n° 1, pp. 13-29, 2007.

[24] Y. K. Liu, B. Zalik : An efficient chain code with Huffman coding, Patt Recog, 38, pp553-557, 2005.

[25] E.Tisserand, J.F. Pautex, Patrick Schweitzer : Analyse et traitement des signaux : Méthodes et applications au son et à l'image, DUNOD Edition, paris, pp 273-277, 2008.

[26] S.Haboubi, S. Maddouri, H. Amiri: Primitive Invariants for Handwritten Arabic Script, International Journal of Computer and Information Science and Engineering, 3(1): p. 40 – 45, 2009.