Recognition of Handwritten Characters Based on Wavelet Transform and SVM Classifier

Malika Ait Aider[1], Kamal Hammouche[1], and Djamel Gaceb[2]
[1]Laboratoire Vision Artificielle et Automatique des Systèmes, Université Mouloud Mammeri, Algérie
[2]Laboratoire D'informatique en Image et Systèmes D'information, Institut National des Sciences Appliquées de Lyon, France

Abstract: *This paper is devoted to the off-line handwritten character recognition based on the two dimensional wavelet transform and a single support vector machine classifier. The wavelet transform provides a representation of the image in independent frequency bands. It performs a local analysis to characterize images of characters in time and scale space. The wavelet transform provides at each level of decomposition four sub-images: a smooth or approximation sub-image and three detail sub-images. In handwritten character recognition, the wavelet transform has received more attention and its performance is related not only to the use of the type of wavelet but also to the type of a sub-image used to provide features. Our objective here is thus to study these two previous points by conducting several tests using several wavelet families and several combinational features derived from sub-images. They show that the symlet wavelet of order 8 is the most efficient and the features derived from the approximation sub-image allow the best discrimination between the handwritten digits.*

Keywords: *Feature extraction; wavelet transform, handwritten character recognition; support vector machine; OCR.*

Received June 10, 2015; accepted May 16, 2016

1. Introduction

Optical Character Recognition (OCR) has long been an active field of research in pattern recognition and recognition of handwritten characters. Various systems and classification algorithms have been proposed in this area for many potential applications such as processing bank checks, mail sorting, automatic reading of forms and various printed and handwritten documents [5, 15]. Basically, handwritten character recognition system involves three stages: a pre-processing stage, a feature extraction stage and a classification stage. The pre-processing task such as normalization techniques [10] consists of modification of the measured data obtained from the data acquisition to be more suitable for further processing. In feature extraction stage, a set of features allowing the discrimination between the characters is extracted. The choice of these features is the most difficult task in pattern recognition because the performance of the used classifier strongly depends on their quality. A large variety of features have been presented to describe the shape of handwritten character [22]. These features include deformable templates, zoning features, features based on projection histograms, contour profiles features and Zernike moments. In addition to these features, several combinational approaches using several features have been proposed [6, 7]. Recently, new features such as invariant curvature features [23], curvelet features [12], Gabor features [19] and wavelet features have been introduced for character recognition.

Wavelet transform provides information about the image contained in time-frequency domain [13]. The performance of the wavelet transform is related to the use of the type of wavelet as well as the types of sub-images used to provide features. Different types of wavelets have been used to extract a variety of features [1, 17, 18, 20].

Several learning methods have been proposed for handwritten character recognition. They include k-nearest neighbour classifier [12] and neural networks classifiers [2]. Support Vector Machine (SVM) has been used for handwritten character recognition [16] and has achieved far better recognition results than traditional techniques. Other techniques based on combining multiple classifiers [7, 8, 14] have been proposed to improve the performance of recognition systems. However, in several works, not only the use of the type of wavelet transforms but also the types of sub-images are arbitrary and not justified. Thus, our goal in this paper is to show that the performance of the system strongly relies on: the choice of the type of wavelet, the choice of the efficiency and the relevance of each sub-image derived from the wavelet and the choice of the combinations of the above sub-images.

In this paper, we propose to invest in the feature extraction methods based on the wavelet transform and support vector machine classifier to recognize handwritten digits. Several types of wavelets are tested in order to select the most appropriate one for our application. Several experiments are also conducted by using coefficients of each sub-image and combinations of all the four sub-images as features.

The rest of this paper is organized as follows: In section 2, the process of the character recognition is described. It includes a brief introduction to the wavelet transform and its use in the feature extraction procedure followed by a short description of support vector machine classifier. Several experimental results using several types of wavelets and several strategies in the characterization of the character images based on the wavelet transform are given and discussed in section 3. In this section, a comparison study with other techniques published in the literature has also been conducted. The last section is the conclusion and some remarks.

2. Character Recognition

We consider the problem of character recognition in terms of pattern recognition in a supervised environment. The pattern recognition scheme contains three phases: feature extraction, training phase and identification phase.

2.1. Feature Extraction based on Wavelet Transform

2.1.1. Concepts of Wavelets

Our interest in this paper is to characterize each character image of the training database by using the wavelet transform. The wavelets are set functions derived from dilatation and translation operations of the mother wavelet ψ such that:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (1)$$

The continuous real variables a and b represent the dilation and translation parameters respectively. The Continuous Wavelet Transform (CWT) of the continuous signal $x(t)$ is defined by:

$$C_{a,b}(t) = \int_{-\infty}^{+\infty} x(t)\psi_{a,b}(t)dt \quad (2)$$

In practice, a 2-D discrete wavelet transform is implemented by multistage filter bank [13]. It consists of filtering and down-sampling horizontally using 1D low-pass filter L (with an impulse response $l(i)$) and high-pass filter H (with an impulse response $h(j)$) to each row in the image $f(x, y)$, produces the coefficient matrices $f_L(x, y)$ and $f_H(x, y)$. Vertically filtering and down-sampling follow, using the low-pass and high-pass filters L and H to each column in $f_L(x,y)$ and $f_H(x, y)$ and produces four sub-images $f_{LL}(x, y)$, $f_{LH}(x, y)$, $f_{HL}(x, y)$ and $f_{HH}(x,y)$ for one level of decomposition. $f_{LL}(x, y)$ is a smooth sub-image, which represents the rough approximation of the image, $f_{LH}(x, y)$, $f_{HL}(x, y)$ and $f_{HH}(x, y)$ are detail sub-images, which represent the horizontal, vertical and diagonal directions of the image. These operations constitute the first level of the decomposition. By decomposing the approximated sub-image of each level again into four sub-images iteratively, a pyramidal image tree is obtained. Figure 1 depicts the first level in a multi-resolution pyramid decomposition of an image.

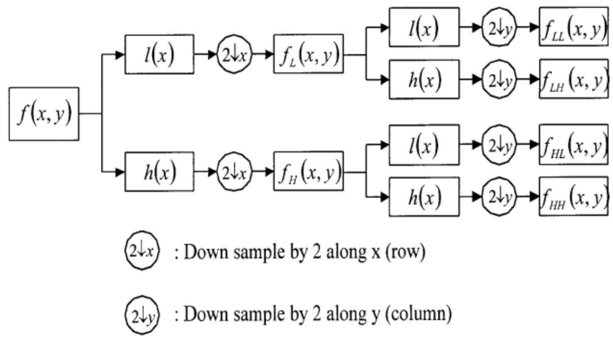

Figure 1. First level in multi-resolution image decomposition.

2.1.2. Feature Extraction

Features derived from the sub-images at each level of the decomposition must be homogeneous for same characters and different for different characters. Some methods consider all the sub-images or only the detail sub-images or only the approximation sub-image of the last decomposition level as features. In handwritten recognition, the use of the coefficients of the approximation image as features gives good results because only the low-frequency components of a character image is less sensitive to writing style variations. Other approaches consist in determining the energy of detail sub-images in each level or in the last level decomposition. Shelke and Apte [21] has proposed the use of another approach to extract features from the wavelet transform by convolving the approximation features with themselves. In our paper, several techniques (see section 3.2) are used to extract a variety of features by combining the wavelet coefficients in many ways.

2.2. Support Vector Classification

Several supervised classification techniques can be used for character recognition. We have chosen to use the SVM method as supervised classification technique for its performances in several application areas [3]. The goal of the SVM classifier is to find optimal hyper-surfaces which allow separating the data into k classes. Consider the problem of separating a set of N training vectors belonging to two family models. The training data are given as: $(X, Y)=(x_1, y_1),(x_2, y_2),\dots(x_N, y_N)$, $x_i \in \Re^M$ and $y_i \in \{+1,-1\}$. Each data x_i represented by M features $x_i = [x_{i1}, x_{i2,\dots}\ x_{iM}]$, belongs to the so called negative class 1 (y_i=-1) or positive class 2(y_i=+1).

When the data are not linearly separable, the SVM uses a non-linear function to map the training data from input space to a high dimensional feature space in which the data become linearly separable. In this case, the SVM technique seeks the hyper plane which allows separating the two classes in the new feature space.

Mathematically, the SVM problem consists in solving the following dual optimal problem:

$$\begin{cases} \max_{\lambda} \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j K(x_i, x_j) \\ \text{with the constraints} \quad 0 \leq \lambda_i \leq C, \quad i = 1, \ldots, N \text{ and} \\ \sum \lambda_i y_i = 0. \end{cases} \quad (3)$$

Where C is a penalty parameter, λ_i are the Lagrange multipliers and the kernel K(.,.)is a symmetric positive definite function which can take several forms as, for example, the Gaussian Radial Basis Function (RBF):

$$K(x_i, x_j) = exp\left(-\frac{\| x_i - x_j \|^2}{2\sigma^2} \right) \quad (4)$$

Here σ is the width of the Gaussian function. In this paper, C and σ are adjusted according to the database used. The resolution of the above dual problem yields the values of the Lagrangian multipliers λ_i^*.

2.3. Character Identification

In order to recognize the character type, the decision Function of an unknown vector $x_z = [x_{z1}, x_{z2, \ldots} X_{zM}]$ is given by the following rule:

$$Class(x_z) = sign\left(\sum_{i=1}^{N_S} \lambda_i^* y_i K(x_i, x_z) + b \right) \quad (5)$$

Where the bias term b can be obtained by:

$$b = y_i - \sum_{i=1}^{N_S} \lambda_i^* y_i K(x_i, y_j) \quad (6)$$

N_s is the number of non null support vectors.

Extensions of the binary support vector machine classifier to the multi-class classification problem are proposed in several works [4].

3. Experimental Results

In this section, we will evaluate the performance of the handwriting recognition based on the wavelet transform and SVM classifier. The MNIST database [9] is used in order to study the discriminatory power of the features extracted from the wavelet transform for character recognition. The database is split into two non-overlapping subsets. The first one is used for designing the SVM classifier and the second is used for estimating the recognition rate. This rate is evaluated as the ratio between the number of characters correctly recognized and the total number of characters. The software implementation of the Ohio State University (OSU) SVM classifier [11] is used in this paper. The parameters of SVM classifier are adjusted in an experimental manner. The penalty parameter C is fixed at 6 and the value of gamma $\gamma = \frac{1}{2\sigma^2}$ is adjusted to 0.25.

3.1. Choice of the Wavelet

The type of the wavelet plays an important role in the handwriting recognition. In this subsection, we evaluate different wavelets types as the family of symlets wavelets (sym2, sym3, sym4, sym5, sym6, sym7, sym8, sym9, and sym12), the Daubechies (db1, db2, db4 and db5), the coiflets (coif1, coif3 and coif5) wavelets and the Cohen Daubechies Feauveau wavelets (*CDF* 3/7 and *CDF* 9/7). To conduct this experiment, MNIST database is used which contains 70,000 greyscale images of handwritten digits where 60,000 images are intended for learning and 10,000 others for testing the recognition phase. The size of each image is (28x28) pixels. The decomposition of the original image using the wavelet transform is carried out at one level. This decomposition generated 4 sub-images (one approximation f_{LL} and 3 details $f_{LH}(x, y)$, $f_{HL}(x, y)$ and $f_{HH}(x, y)$. Each sub-image is characterized by a matrix of (14x14) coefficients. The coefficients of the approximation sub-image are considered as features. Hence, each character image is characterized by a feature vector composed of 196 coefficients of the approximation sub-image. Table 1 gives the recognition rates obtained by each wavelet. These results show that the features derived from the wavelet transform are able to discriminate the handwriting characters since the recognition rate is higher than 98%. However, among the tested wavelets, the symlet wavelet of order 8 provides the best recognition rate of 98.76 %.

Table 1. Recognition rates of wavelet based features.

Wavelets	Recognition rate	Wavelets	Recognition rate
Haar	98.56%	Sym4	98.70%
Db2	98.66%	Sym5	98.73%
Db4	98.68%	Sym6	98.71%
Db5	98.65%	Sym7	98.66%
Coif1	98.68%	Sym8	98.76%
Coif3	98.68%	Sym9	98.75%
Coif5	97.90%	Sym12	98.67%
Sym2	98.66%	CDF 3/7	98.52%
Sym3	98.67%	CDF 9/7	98.64%

3.2. Choice of Features

In the above experiment, the features used to classify the handwriting characters correspond only to coefficients of the approximate sub-image (f_{LL}) obtained at one level of resolution. However, other coefficients can be used as features for characterizing the characters as explained in the section 2.1.2. Hence, in this section, we will examine the relevance of the features corresponding to combinations of the four sub-images f_{LL}, f_{LH}, f_{HL} and f_{HH} generated by the wavelet transform. Next we consider the coefficients of each sub-image as features, or take as features the minimum, the maximum or the average between two sub-images only, between three sub-images only or between all sub-images. The symlet wavelet transform of order 8 (sym8) with one decomposition level is used in the second experiment. The recognition rates for different combinations are summarized in Table 2. It shows clearly that the best recognition rate is obtained when only the approximation sub-image is used as features. The use of the coefficients of details as features decreases the recognition rate. In the last line of the Table 2, all the sub-images are considered to generate the features by applying the maximum, minimum or mean operators.

Table 2. Results of recognition based on combination methods.

f_{LL}, f_{LH},f_{HL},f_{HH}	Max	Min	Mean
1000	98.76%	98.76%	98.76%
0100	83.82%	83.82%	83.82%
0010	81.19%	81.19%	81.19%
0001	70.77%	70.77%	70.77%
1100	95.84%	96.95%	97.65%
1010	95.57%	96.93%	97.40%
1001	93.58%	96.60%	97.10%
0110	83.04%	84.15%	85.75%
0101	78.66%	79.55%	77.35%
0011	77.01%	77.85%	75.03%
1110	94.55%	94.98%	93.55%
1101	93.39%	94.96%	97.82%
1011	93.73%	94.88%	91.56%
0111	82.39%	83.29%	73.56%
1111	93.71%	92.58%	94.04%

Another way to combine information from the four sub-images is to assign weights to each of them and generate new features such that: $I=C_1*f_{LL}+C_2*f_{LH}+C_3*f_{HL}+C_4*f_{HH}$.

The weights C_1, C_2, C_3 and C_4 take values between 0 and 1 and must satisfy the condition $C_1+C_2+C_3+C_4=1$. The optimal values of Ci(i=1,2,3,4) leading to the best recognition rate are given in Table 3.

Table 3. Recognition rate based weights.

Sub-image	f_{LL}	f_{LH}	f_{HL}	f_{HH}
Weights	C_1=0.7	C_2=0.1	C_3=0.1	C_1=0.1
Recognition rate	98.66%			

They are determined experimentally by exhaustive research. The value of C_1 is greater than those of others weights. It indicates that we must give more importance to the approximation information than to details. The corresponding recognition rate is also given in the Table 3. Its value of 98.66 % is higher than the rate achieved by the other combinations such as those presented in the above Table 2, but remains lower than the rate achieved by the approximation sub-image. Finally, the features extracted only from the approximation sub-image using the sym8 wavelet allow a better discrimination between the handwriting characters.

3.3. Comparison with Other Variants of Multi-Resolution Methods

To show the relevance of the features derived from the wavelet decomposition (approximation sub-image with sym8 wavelet) in the handwritten recognition, we compared them with those obtained by two other multi-resolution methods on the MNIST database.

These two multi-resolution methods allow constructing feature vectors from the wavelet decomposition with another way. The first method, called "reduced normalized image", consists in dividing each character image of size (28x28) into (14x14) non overlap blocks of (2x2) pixel size each.

The grey level mean is then evaluated in each block. Finally, the (14x14) mean values can be considered as an approximation sub-image of original image and can be used to form the feature vector. The recognition rate achieved by these features and the SVM classifier is equal to 98.55% (Table 4) and remains lower than the one achieved by the features extracted from the approximation sub-image provided by the wavelet decomposition which is 98.76%.

The application of similar multi-resolution strategy on the image gradient (Sobel operator) instead of the original image constitutes the second comparative method. This method, called here “reduced gradient image”, allows characterizing each character image by a feature vector which can also be considered as the detail sub-image. The recognition rate achieved by these other features is equal to 97.55% (Table 4). It is higher than the recognition rate provided by the features extracted from details or their combinations coming from the wavelet decomposition but remains lower than the 98.76% achieved by the features extracted from the approximation sub-image coming from the wavelet decomposition.

Table 4. Recognition rates by multi-resolution methods.

Features	Training	Testing	R. Rate
Reduced normalized image	60000	10000	98.55%
Reduced gradient image	60000	10000	97.55%

3.4. Comparison with Other Works

To compare the efficiency of our character recognition technique based on the wavelet transform and the SVM classifier to those published in the literature, the tests must be made in the same conditions. But in practice, several works which have been done on the MNIST database, consider size normalization as a pre-processing technique to improve the performance of the system. However, no further size normalization is considered in our tests.

Firstly, the comparison is conducted with techniques presented in [2, 17, 18, 20] by using the MNIST database. The results are indicated in table 5. Recall that in [2] the recognition rate of 97.57% was achieved by using the features derived from the Daubechies wavelet transform and a majority voting method. However, this value is less than 98.60% found in our experiment. The result of recognition rate of 98.09% reported in [17] using the complex wavelet structural similarity based support vector machine method (CW-SSIM SVM) remains inferior to 98.76%. In [18], 98.22% of recognition rate is obtained by using the combination of the gradient features and curvature features derived from the Mexican hat continuous wavelet transform but remains lower than ours. In [20], the (WT-PCA) descriptors have been applied on the directional features extracted from the four Kirsh masks. Using individual classifiers, the best recognition rate of 98.64% was obtained. Note that we have found the same value without reduction in dimension of the features and using one type of features (Table 1). It is only by combining the different responses of five classifiers of SVM type with the Bayesian strategy that authors could improve the recognition result to reach 99.32%.

Table 5. Various recognition rates with MNIST database.

Classifiers	References	Training set	Testing set	R.Rate
SVM rbf	Our paper	60000	10000	98.76%
SVM rbf	Our paper	50000	10000	98.60%
MLP	[2]	50000	10000	97.57%
SVM	[17]	60000	10000	98.09%
MLP	[18]	60000	10000	98.22%
SVM	[20]	60000	10000	89.64%
SVMs	[20]	60000	10000	99.32%

4. Conclusion

In this paper, we have presented a technique of handwritten character recognition which combines a wavelet transform and a single support vector machine classifier. The wavelet transform allows us characterizing the character images by a set of features.

The relevance of these features depends strongly on the choice of the type of the wavelet and sub-images derived from the wavelet transform. In this paper, several tests including several wavelets and smooth and details sub-images derived from the wavelets are conducted in order to determine the best wavelet and the best sub-image in the handwritten recognition framework. Experimental results on MNIST database reveal that sym8 wavelet outperforms other types of wavelets as those used in the previous works. They show also that features extracted from the smooth sub-image allowed achieving the best recognition rate. The proposed technique is efficient in comparison with other handwritten recognition methods published in the literature. As future work, we intend on one hand to integrate a normalization operation as preprocessing procedure in order to regulate the position and shape of character images, so as to reduce shape variation between the images of same class. On other hand, to investigate other features extracted from the four sub-images.

References

[1] Berthiaume V. and Cheriet M., "Handwritten Digit Recognition by Fourier-Packet Descriptors," *Electronic Letters on Computer Vision and Image Analysis*, vol. 11, no. 1, pp. 68-76, 2012.

[2] Bhattacharya U. and Chaudhuri B., "A Majority Voting Scheme for Multiresolution Recognition of Handprinted Numerals," *in Proceedings of the 17th International Conference on Document Analysis and Recognition,* Edinburgh, pp. 16-20, 2003.

[3] Burges C. and Scholkopf B., "Improving the Accuracy and Speed of Support Vector Machines," *in Proceedings of the 9th International Conference on Neural Information Processing Systems*, Denver, pp. 375-381, 1997.

[4] Cortes C. and Vapnik V., "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.

[5] Gaceb D., Eglin V., and Lebourgeois F., "A New Mixed Binarization Method Used in a Real Time Application of Automatic Business Document and Postal Mail Sorting," *The International Arab Journal of Information Technology*, vol. 10, no. 2, pp. 179-188, 2013.

[6] Gorgevik D. and Cakmakov D., "An Efficient Three-Stage Classifier for Handwritten Digit Recognition," *in Proceedings of the 17th International Conference on Pattern Recognition*, Cambridge, pp. 507-510, 2004.

[7] Kim K., Park J., Song Y., Kim I., and Suen C., "Recognition of Handwritten Numerals Using a Combined Classifier with Hybrid Features," *in Proceedings of International Workshops on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition*, Lisbon, pp. 992-1000, 2004.

[8] Lauer F., Suen C., and Bloch G., "A Trainable Feature Extractor for Handwritten Digit Recognition," *Pattern Recognition*, vol. 40, no. 6, pp. 1816-1824, 2007.

[9] Lecun Y., "The MNIST Database of Handwritten Digits," *http://yann.lecun.com/exdb/mnist*, Last Visited, 2010.

[10] Liu C., Nakashima K., Sako H., and Fujisawa H., "Handwritten Digit Recognition: Investigation of Normalization and Feature Extraction Techniques," *Pattern Recognition*, vol. 37, no. 2, pp. 265-279, 2004.

[11] Ma J., Zhao Y., and Ahalt S., "OSU SVM Classifier Matlab Toolbox," Ohio State University, USA.

[12] Majumdar A., "Bangla Basic Character Recognition Using Digital Curvelet Transform," *Journal of Pattern Recognition Research1*, vol. 2, no. 1, pp. 17-26, 2007.

[13] Mallat S., "Multifrequency Channel Decompositions of Images and Wavelet Models," *IEEE Transactions on Acoustics, Speech and Signals Processing*, vol. 37, no. 12, pp. 2091-2110, 1989.

[14] Niu X. and Suen C., "A Novel Hybrid CNN-SVM Classifier for Recognizing Handwritten Digits," *Pattern Recognition*, vol. 45, no. 4, pp. 1318-1325, 2012.

[15] Pfister M., Behnke S., and Rojas R., "Recognition of Handwritten ZIP Codes in a Real-World Non-Standard-Letter Sorting System," *Applied Intelligence*, vol. 12, no. 1-2, pp. 95-114, 2000.

[16] Rajput G. and Mali S., "Fourier Descriptor Based Isolated Marathi Handwritten Numeral Recognition," *International Journal of Computer Applications*, vol. 3, no. 4, pp. 9-13, 2010.

[17] Rehman A., Gao Y., Wang J., and Wang Z., "Image Classification Based on Complex Wavelet Structural Similarity," *in Proceedings of International Conference on Image Processing*, Brussels, pp. 1-9, 2011.

[18] Romero D., Ruedin A., and Seijas L., "Wavelet-Based Feature Extraction for Handwritten Numerals," *in Proceedings of International Conference on Image Analysis and Processing*, Vietri sul Mare, pp. 374-383, 2009.

[19] Ryszard S., "Feature Extraction of Gray-Scale Handwritten Characters Using Gabor Filters and Zernike Moments," *in Proceedings of Computer Recognition Systems 2*, Berlin, pp. 340-347, 2007.

[20] Seijas L. and Segura E., "A Wavelet-Based Descriptor for Handwritten Numeral Classification," *in Proceedings of International Conference on Frontiers in Handwriting Recognition*, Bari, pp. 653-658, 2012.

[21] Shelke S. and Apte S., "A Multistage Handwritten Marathi Compound Character Recognition Scheme Using Neural Networks and Wavelet Features," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 4, no. 1, pp. 81-93, 2011.

[22] Trier O., Jain A., and Taxt T., "Feature Extraction Methods for Character Recognition-A Survey," *Pattern Recognition*, vol. 29, no. 4, pp. 641-662, 1996.

[23] Yang L., Suen C., Bui T., and Zhang P., "Discrimination of Similar Handwritten Numerals Based on Invariant Curvature Features," *Pattern Recognition Society*, vol. 38, no. 7, pp. 947- 963, 2005.

Malika Ait Aider She is a permanent research scientist in the Automatic Control Department of Electrical Engeneering and Computer Science Faculty at the University Mouloud Mammeri of Tizi Ouzou, Algeria. Her research activity is mainly focused on feature extraction and classification for character recognition.

Kamal Hammouche He is currently Professor in the Automatic Control Department of the Electrical Engineering and Computer Science Faculty at the University Mouloud Mammeri of Tizi Ouzou, Algeria. His research interests are in the area of image processing and pattern recognition.

Djamel Gaceb He has a PhD in computer science from INSA of Lyon. Currently he is working on the topic of business document image recognition, retrieval and analysis, industrial vision image processing on mobile devices and real time applications.