Recognition of Amazigh characters using SURF & GIST descriptors

H. Moudni, M. Er-rouidi, M. Oujaoura
Computer Science Department
Faculty of Science and Technology
Sultan Moulay Slimane University
Béni Mellal, Morocco.

O. Bencharef
Computer Science Department
Higher School of Technology
Cadi Ayyad University
Essaouira, Morocco.

Abstract— **In this article, we describe the recognition system of Amazigh handwritten letters. The SURF descriptor, specifically the SURF-36, and the GIST descriptor are used for extracting feature vectors of each letter from our database which consists of 25740 manuscripts isolated Amazigh characters. All the feature vectors of each letter form a training set which is used to train the neural network so that it can calculate a single output on the information it receives. Finally, we made a comparative study between the SURF-36 descriptor and GIST descriptor.**

Keywords—SURF; GIST; Principal Component Analysis; Neural Network; Amazigh Characters.

I. INTRODUCTION

Today in Morocco, pattern recognition, especially the recognition of Amazigh characters has become a growing field in which several researchers work. In order to find algorithms that can solve the problems of computer pattern recognition, which are intuitively solved by humans, several research efforts have been made and some research works of Tifinagh characters are published [1, 2, 3, 4, 5]. In this context, we proposed a system that allows for the recognition of Amazigh handwritten characters using the simplest descriptors like SURF-36 and GIST Today in Morocco, pattern recognition, especially the recognition of Amazigh characters has become a growing field in which several researchers work. In order to find algorithms that can solve the problems of computer pattern recognition, which are intuitively solved by humans, several research efforts have been made. In this context, we proposed a system that allows for the recognition of Amazigh handwritten characters using the simplest descriptor, SURF-36 and GIST as descriptors, in addition to the neural network as a classifier.

For the rest of the paper, in Section 2 and Section 3, we present the SURF-36 and GIST descriptors in addition to their calculation steps. Section 4 and section 5 have been reserved respectively for the neural networks and the database used in this paper. In Section 6, we presented the principal component analysis technique and studied the possibility of its application in this case. Section 7 presents the discussion of results. Finally, in Section 8, we concluded our work.

II. SURF DESCRIPTOR

Speeded Up Robust Features (SURF) is a visual feature extraction algorithm from an image to describe it based on the detection of interest points. We worked with the reduced SURF descriptor (SURF-36) which is slightly worse compared to usual descriptor SURF. Nevertheless, it allows a very rapid adaptation, the performance remain acceptable in comparison with other descriptors in the literature.

The SURF descriptor is mainly known for its fast computation. Its algorithm consists of two main steps. The first one is to detect the interest points in the image and the second one is to describe these interest points using a vector of 36 features.

A. Detection of Interest Points

To decrease computation time, the image to be analyzed is transformed into an integral picture. The integral images allow fast calculation of convolution and rectangular areas. Let I our initial image, I (x, y) represents the pixel value of the image at coordinates x and y.

The integral image denoted IΣ (x, y) is an image of the same size as the original image, it is calculated from this image. Each pixel of the integral image contains the sum of pixels located above and left of the pixel in the original image.

The value of a pixel of the integral image IΣ (x, y) is defined on the basis of the image I by the following equation:

$$I_{\Sigma}(x,y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x,y) \qquad (1)$$

The pixels Areas in the image with high change of intensity are searched. The Hessian matrix, based on the calculation of partial derivatives of order two, is used for this. For a function of two variables f (x, y), the Hessian matrix is defined as follows:

$$H(f(x,y)) = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y)}{\partial y \partial x} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} \qquad (2)$$

If the determinant of the Hessian matrix is positive, then the eigenvalues of the matrix are both positive or both negative, which means that an extremum is present. Points of interest will therefore be located where the determinant of the Hessian matrix is maximal. Specifically, the partial derivatives of the signal are calculated by convolution with a Gaussian. To gain

speed calculation, these are approximated by a Gaussian step function called box filter.

The representation at lower levels of scale is achieved by increasing the size of the Gaussian filter. In the end, the interest points for which the determinant of the Hessian matrix is positive and which are local maximum in a neighborhood 3 * 3 * 3 (x-axis * y- axis * scale-axis) are retained.

B. Description of Interest Points

Once the interest points are extracted, the second step is to calculate the corresponding descriptor. The SURF descriptor describes the intensity of the pixels in a neighborhood around each interest point. The x and y Haar wavelets response is calculated in a neighborhood of 6s, where s is the scale at which the interest point was found. From these values, the dominant orientation of each point of interest is calculated by dragging a window orientation.

To calculate the descriptor, a square of size 20s oriented along the dominant orientation is extracted. This area is divided into 3 x 3 squares. For each of the sub-regions, Haar wavelets are calculated on 15 x 15 points.

Let dx and dy be the response to the Haar wavelet, four values are calculated for each sub-regions:

$$v_{subregion} = \left[\sum dx, \sum dy, \sum |dx|, \sum |dy| \right] \quad (3)$$

Finally, each of the extracted points in the previous step is described by a vector composed of 3*3*4 values that is 36 dimensions [6].

III. GIST DESCRIPTOR

In computer vision, GIST descriptors are a representation of an image in low dimension that contains enough information to identify the scene. Actually, any global descriptor must approach the GIST to be useful.

GIST descriptor was proposed by Oliva and more precisely by Torralba. They tried to capture the GIST descriptor of the image by analyzing the spatial frequencies and orientations. The global descriptor is built by combining the amplitudes obtained in the output of K Gabor filters at different scales and orientations. For reducing the size, each image in filter output is resized to a size N*N (N between 2 and 16), which gives a vector of dimension N*N*K. This dimension is further reduced through a principal component analysis (PCA), which also gives the weights applied to different filters [7].

IV. NEURAL NETWORKS

Neural networks are composed of simple elements (or neurons) working in parallel. These were strongly inspired from biological nervous system. As in nature, the functioning of the neural network is strongly influenced by the connections between the elements. It can lead a neural network to a specific task (eg OCR) by adjusting the values of connections (or weight) between the elements (neurons).
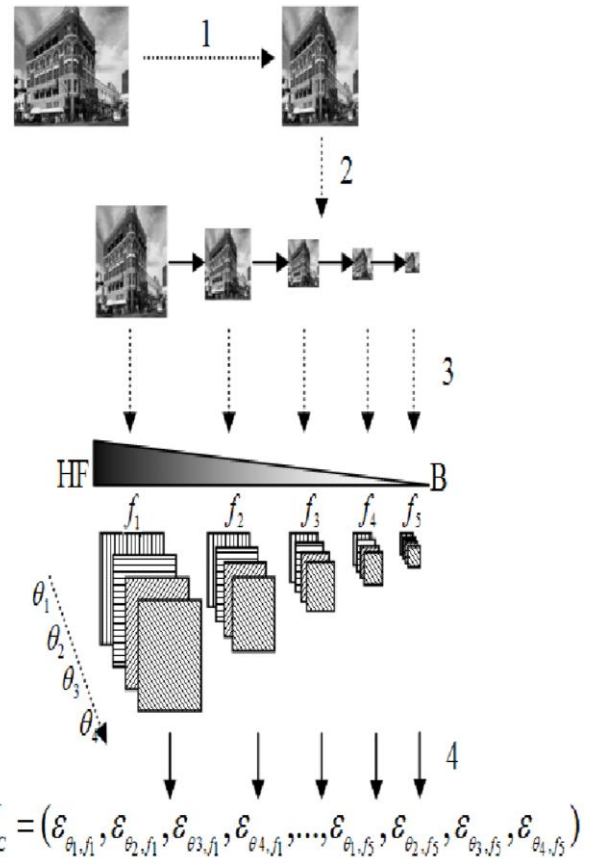

Fig. 1. GIST Descriptor

In general, the neural networks learning tasks is done and performed so that for a particular entry, the neural network give a specific target. The weight adjustment is carried out by comparison of the network response (or output) and the target, until the output corresponds at best to the target [8].
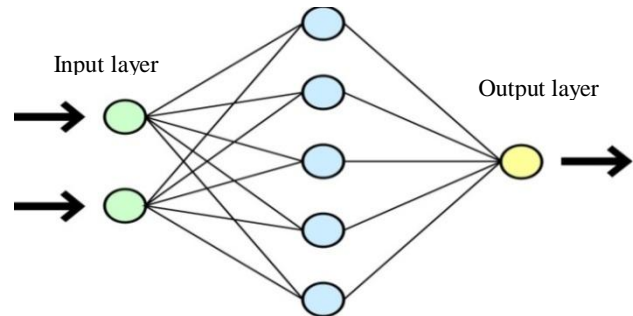

Fig. 2. Simplified diagram of a neural network

V. USED DATABASE

The database contains 33 handwritten characters Amazigh. Each character is represented by 780 ways and sizes which gives 25740 handwritten characters Amazigh. This database was developed at the Laboratory of IRF-SIC Ibn Zuhr University in Agadir, Morocco.
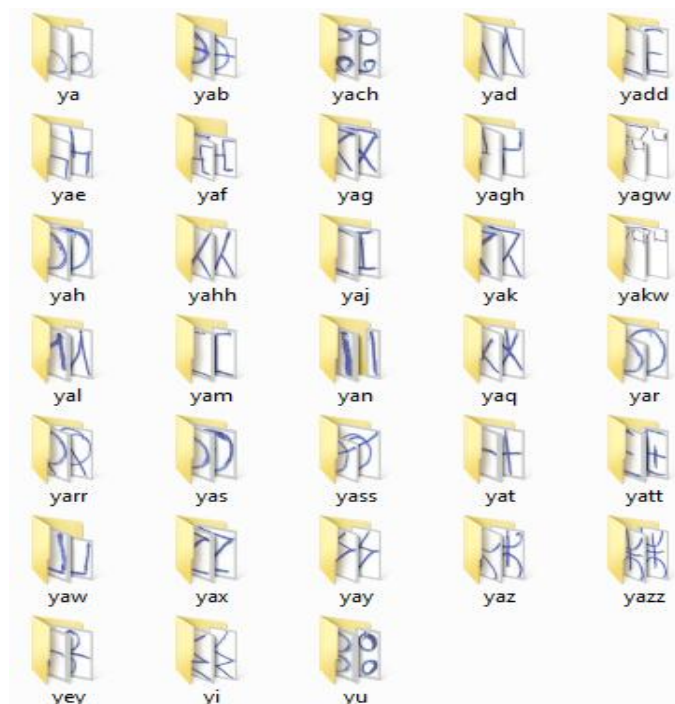
Fig. 3.   The database files.

The basic representation of Amazigh characters is given by the following figure [9]:

Fig. 4.   The basic representation of Amazigh characters.

VI.   PRINCIPAL COMPONENT ANALYSIS (PCA)

The Principal Component Analysis (PCA) is one of multivariate descriptive analyzes. The purpose of this analysis is to summarize the maximum possible information losing only the least possible information in order to:

- Facilitate the interpretation of a large number of initial data.
- Give more meaning to the reduced data.

Therefore, the PCA reduces data tables of large sizes in a small number of variables (usually 2 or 3) while keeping a maximum of information. Baseline variables are called 'metric'.

To analyze the results of PCA, it would be very helpful and useful to answer three questions:

- Are the data factorisable?
- How many factors should be retained?

- How to interpret the results?

To answer the first question, firstly, the correlation matrix should be observed. If several variables are correlated (> 0. 5), the factorization is possible. If not, the factorization has no sense or meaning and is therefore not recommended [10].

In our example we used the function corrcoef (base) from the MATLAB code to determine the correlation matrix. By examining this matrix, we find that several variables are not correlated (<0. 5).

So we cannot use the CPA for our database Amazigh characters.

VII.   RESULTS INTERPRETATION

Our learning database, sized of (25740x36), contains all the characteristic vectors of 33 Amazigh characters (each character is represented by 780 different handwriting ways).

According to Figure 5, the feature vectors are different for each character (the 780 representations of the first character are different from the 780 representations of the second character, and so on).
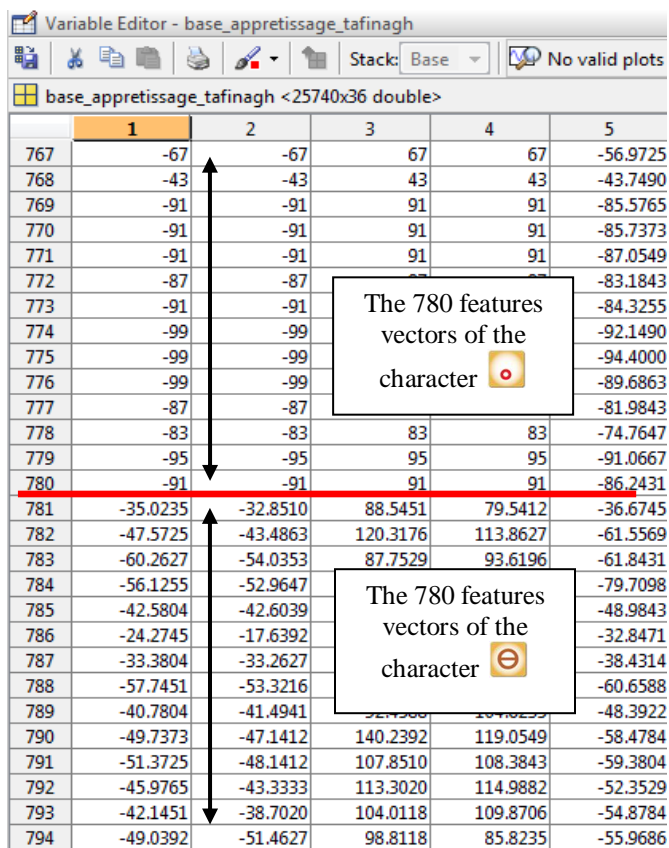

Fig. 5.   Overview of the learning database.

The produced learning database will be used as an input for the neural network. When the Neural network established its identification and final decision, we did a test on all vectors of the training database with a simple MATLAB simulation function sim(neuron, vector). To simplify this task, we created

a program that simulates the entire database and giving as a result the error rate for each used descriptor.

We presented the summary of results in Table 1. The error rate is 25% for the SURF-36 descriptor and 17% for the GIST descriptor.

TABLE I.   ERROR RATES AND EXECUTION TIME OF SURF-36 & GIST DESCRIPTORS

Descriptor	SURF-36	GIST
Error rate	25%	17%
Execution Time in Min	8	180

From Figures 6 and 7, we observe that the GIST descriptor is more powerful than the SURF-36 descriptor. But SURF-36 is faster than GIST. Indeed, the performance of SURF-36 took 8 minutes to process our database of Amazigh characters while GIST took over 3 hours.

VIII.   CONCLUSION

We note that the SURF-36 descriptor is a little less efficient than the GIST descriptor, but it allows very rapid adaptation in recognition of Amazigh handwriting characters. The combination with the neural network, both for the SURF-36 descriptor or GIST descriptor gives insufficient results. In future work, other descriptors and other classifiers will be used to improve the recognition rate of the Amazigh characters recognition system.
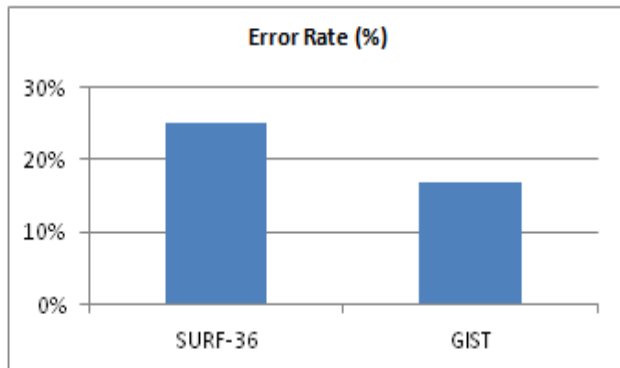
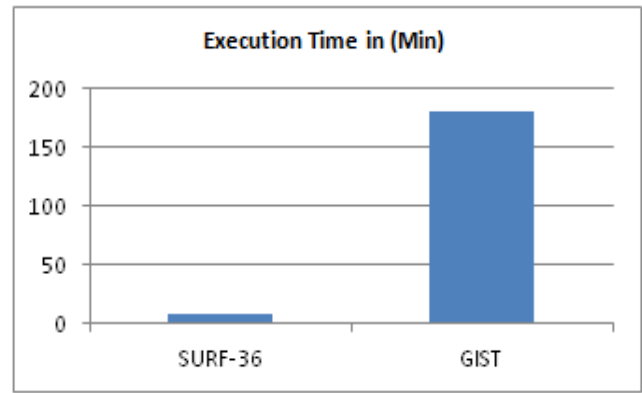Fig. 6.   SURF-36 and GIST descriptors error rate comparison.

Fig. 7.   Execution time Comparison of SURF-36 and GIST descriptors.

REFERENCES

[1] O. Bencharef, M. Fakir, B. Minaoui et B. Bouikhalene. (2011). Tifinagh Character Recognition Using Geodesic Distances, Decision Trees & Neural Networks. (IJACSA) International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence, 1-5.

[2] Y. Es Saady, A. Rachidi, M. El Yassa, D. Mammass, « Printed Amazigh Character Recognition by a Syntactic Approach using Finite Automata », ICGST-GVIP Journal, Volume 10, Issue 2, June 2010.

[3] O. Bencharef, M. Fakir, N. Idrissi, B. Bouikhalen et B. Minaoui, « Application de la géométrie riemannienne à la reconnaissance des caractères Tifinaghe », Agadir-Maroc, 06-07 Mai 2011. pp : 179 :188.

[4] R. El Ayachi, K. Moro, M. Fakir et B. Bouikhalene, « On The Recognition Of Tifinaghe Scripts ». JATIT, vol. 20, No. 2, pp: 61-66, 2010.

[5] R. El Ayachi, M. Fakir et B. Bouikhalen. (2012). Transformation de Fourier et moments invariants appliqués à la reconnaissance des caractères Tifinaghe, Revue eTI, Numéro 6. 2012.

[6] Olivier Augereau, Nicholas Journet, Jean-Philippe Domenger, "Reconnaissance et Extraction de Pièces d'identité : Une application industrielle à la détection de cartes d'identité et de passeports," 1re soumission à CIFED 2012, le 8 décembre 2011.

[7] TOUSCH A.M. « Hiérarchies sémantiques pour l'annotation multifacette d'images ». Thèse de doctorat (Mathématiques, Informatique) soutenue le 15 février 2010 à l'Ecole des Ponts - Paris Tech.

[8] http://www.obs.univbpclermont.fr/atmos/enseignement/cours-Master-2A/cours_RN_2006.pdf

[9] http://upload.wikimedia.org/wikipedia/commons/a/a5/Alphtifingh.jpg

[10] http://www.lemoal.org/download/spss/ACP.