Local Normalization Towards Global Recognition of Arabic Handwritten Script

Samia Snoussi Maddouri[1], Hamid Amiri[1], and Abdel Belaïd[2]

[1] Laboratory of Systems and Signal Processing (LSTS), National Engineering School of Tunis (ENIT), B.P 37 Belvédère 1002, Tunisia
samia.maddouri@enit.rnu.tn
hamid.amiri@enit.rnu.tn
[2] Group READ, LORIA-CNRS Campus scientifique B.P. 239 54506 Vandoeuvre-Les-Nancy France
abdel.belaid@loria.fr

Abstract. In this paper we present a geometrical correction method of handwritten characters. This normalization method is based on Fourier coefficients of a chain-encoded contour. It is carried out in order to improve the rate and the performance of an Arabic handwritten word recognition system. The recognition is based on the global observation of apparent features associated to key-letters of the recognized word, followed by a succession of local observations of secondary features in the word. The global observation does not only try to detect key-word letters, but it also generates a set of characters giving some hints about the identity of the unknown letters present in the word image. The normalization stage leads back the features matching the letter hypothesis in order to increase the chance of recognizing the word presented to the system. Transformed parameters are the angular rotation of the first harmonic phasors and the magnitude of the semimajor axis associated to Fourier coefficients of the truncated approximation to a closed contour. Experimental tests help to evaluate distances between the normalized character and its reference.

1 Introduction

The recognition of a printed text has reached an exploitation level. However, the recognition of a handwritten text is still in a development stage. A printed text without a noise can be considered as an invariable character model. On the other hand, we can consider a handwritten text as a model of variable characters. We find character variability in position, size, rotation, slant and distortion. This variability is the most difficult aspect of a handwritten text recognition. Hence, the use of printed character recognition methods for handwriting recognition is, then, conditioned by handling this major handicap. The transformation of a handwritten character to a standard form (such as printed) is the aim of normalization. It is defined as a series of deformation of the original image in order to reduce their variability [1]. Several normalization methods are described in the literature [1], [2], [3], [4]. Some of them are largely used for Chinese handwritten script including a large variety of symbols. In the eighties, Kuhl introduced a transformation based on Fourier coefficients [5]. The idea was to describe the contour of a character by Fourier coefficients. Then, he applied geometric transformations to the Fourier coefficients in order to eliminate variability. This allows him to describe any image by invariant parameters. Recently, these transformation were used in [6] to normalize the contour of printed characters. In [7] and [8] a method for primitive extraction of long bones using a set of 2D Fourier descriptors is proposed. This set is shown to be stable, complete and endowed with geometrical invariancy properties. The invariant descriptors of the character are used for normalization and classification. It can also be used to reproduce a normalized image from the initial one. These different techniques try to normalize isolated characters (such as Asian characters, Latin characters and Arabic numbers). Normalization in this case can be independent from the context. In case of Arabic or Latin script, normalization tends to approach a new form to an already known reference, in order to reduce the handwritten variability and achieve better recognition rates.

The objective of our work is to create a relationship of similarity between a handwritten character taken from a word context and a reference set of characters proposed by a first step of a word recognition system. The normalization is specific to the handwritten character. It is considered as a step in the hole recognition system based on two complementary bottom-up and top-down process, see Fig. 1.

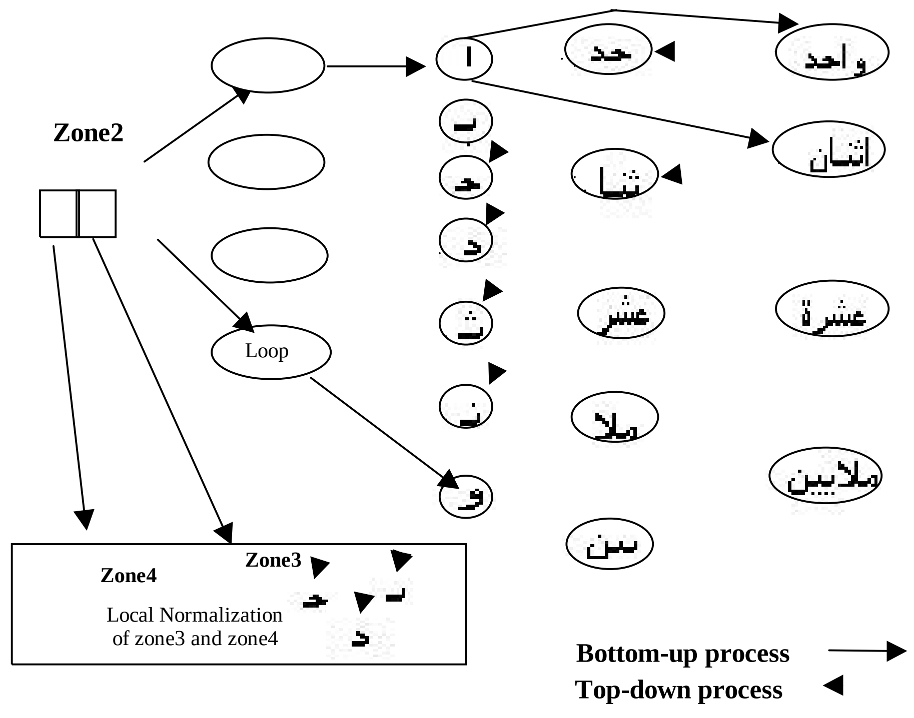


Fig. 1. Local normalization and global recognition

During the bottom-up process, the information propagates from the lower (feature) level toward the higher (word) level, and vice versa in the top-down process. At the end of the bottom-up process, some words are activated. During the top-down process, contextual information is taken into account and the activated words generate letter hypotheses which give some hints about the identity of the unknown letters present in the image [9], [10], [11]. These letters are considered as reference for the normalization step. So, the principal objective of normalization is to help the recognition system in choosing one of the characters proposed by the top-down process. This is done by approaching the local unknown form to one of the proposed characters.

This paper is organized in the following way. In section 2, we explain how the reference set is determined by the global recognition system according to the apparent morphological characteristics of the handwritten character to recognize. In section 3, we deal with the description of the character contour by Fourier coefficients. Section 4 presents the geometric transformation of these coefficients in order to eliminate the handwritten variability. In section 5, we present and discuss the experiments done on a large set of random multiscriptor characters. Finally our conclusion and perspectives are presented in section 6.

2 The Problem of Normalization

Some normalization methods try to transform different shapes into a standard one. This transformation makes the risk of moving the image away from its real shape [1], [4]... Other methods start by learning the different shapes, then they try to normalize according to these learned shapes [2], [3]... Independently from shapes other methods try to normalize descriptors [7], [8]...

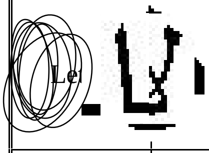
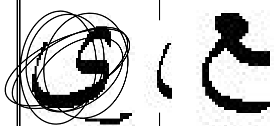
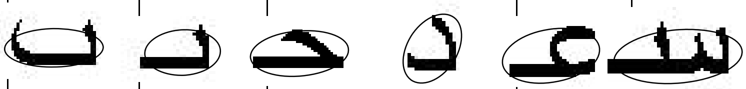
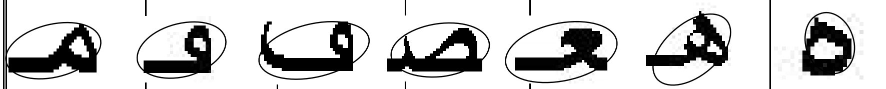
2.1 The Reference Set

In order to achieve the operation of normalization, it is interesting to choose a standard reference set. We choose the printed Arabic alphabet written in font " NESKHI " because of its simplicity and its common use in Arabic script. In order to guide the system in the choice of the nearest characters, pre-classification of printed characters is carried out on the basis of four global features : ascender, descender, loop and diacritic. The feature of printed Arabic character defined by the first harmonic phasor is then detected in order to prepare to normalization step. Table 1 present's all the printed Arabic characters with "NESKHI" font.

2.2 Normalization Procedure

Our normalization approach uses the boundary function of the characters. It is based on the method developed by Kuhl and Gardina [5], [6]. Figure 2 presents the different steps of a normalization system. In the first step, we start by boundary detection. In the second step, the Freeman chain code is generated.

Table 1. Global features and rotation angular printed Arabic characters (font NESKHI)

	Ascender						
Le 							
Angles	90°	100	73°	62°	85°	95°	82°
	Descender						
Letters 							
	45°	75°	67°	20°	54°		
							
35°	25°	30°	78°	80°	86°		
	Neither Ascender, nor Descender						
Letters 							
Angles	1°	3°	6°	57°	10°	5°	
							
14°	15°	13°	8°	12°	40°	18°	

A Fast Fourier Transform (FFT) algorithm calculates the Fourier coefficients of a chain-encoded contour. The forth step attempts to normalize these coefficients to cope with orientation and size variation of a handwritten character. A possible step of regeneration of a normalized character can be done by a simple application of the reverse FFT [12], [13], [14].

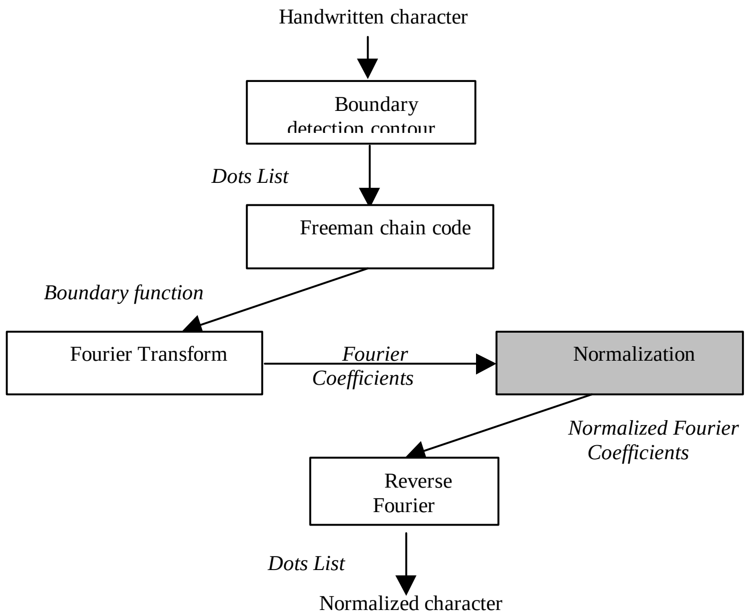


Fig. 2. character normalization system

3 Fourier Coefficients of a Script

3.1 Closed Boundary Function

A continuous contour of a character can be approximated by the Freeman chain code described by a sequence of piecewise linear fits representing standardized links. Each link is an integer between 0 and 7. It is oriented from 0 to 2π in angle steps of $\pi/4$ (as measured counter-clockwise from X axis of an X-Y coordinate system). Its length is of 1 or $\sqrt{2}$ depending on the link parity [5].

This description allow us to generate a periodic and temporal function. We can then define either continue or discrete Fourier Transform. Indeed, the time needed to go through the first k links in the chain at constant speed from an arbitrary starting point is :

$$t_k = \sum_{i=1}^{k} \Delta t_i \quad (1)$$

Where : Δt_i is the length of i^{th} link.

The basic period of the chain code is :

$$T = \sum_{i=1}^{N} \Delta t_i \qquad (2)$$

N is the total number of codes describing the character boundary.
The projection on X and Y of the first k links of the chain are respectively :

$$x_k = \sum_{i=1}^{k} \Delta x_i \quad \text{et} \quad y_k = \sum_{i=1}^{k} \Delta y_i \qquad (3)$$

Δx_i and Δy_i can be 1, 0 or –1 depending on the value and the orientation of the Freeman link. Figure 3 explains these values [5].

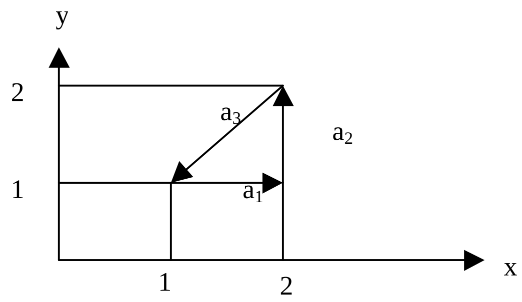


Fig. 3. The changes in the X, Y projections of the Freeman chain code. $a_{1x}=1$, $a_{1y}=0$; $a_{2x}=0$, $a_{2y}=1$; $a_{3x}=-1$, $a_{3y}=-1$.

Thus, the Fourier transform for X-Y projection of the chain of the complete contour is defined as :

$$X_N(k) = A_0 + \sum_{n=1}^{N} a_n \cos \frac{2n\pi k}{N} + b_n \sin \frac{2n\pi k}{N} \qquad (4)$$

$$Y_N(k) = C_0 + \sum_{n=1}^{N} c_n \cos \frac{2n\pi k}{N} + d_n \sin \frac{2n\pi k}{N} \qquad (5)$$

Where :
k : the k^{th} boundary point,
N : is the number of harmonics needed in the approximation of the boundary by Fourier coefficients. It correspond also to the number of boundary points.
It is shown that the number of harmonics required (such as that of the approximation of a contour is completely equal to the original contour) is N.
a_n, b_n, c_n and d_n are Fourier coefficients corresponding to the n^{ith} harmonic.
A_0 et C_0 the bias terms corresponding to frequency 0.
These coefficients are generated by the Fast Fourier Transform algorithm (FFT) [10], [11], [12]. The FFT algorithm reduces the computation time of the Fourier transform by orders of magnitude.

3.2 Fourier Approximation of Contour Image

The components of the productions ($X_n(k)$, $Y_n(k)$), n>0 are given by :

$$X_n(k) = a_n \cos \frac{2n\pi k}{N} + b_n \sin \frac{2n\pi k}{N} \tag{6}$$

$$Y_n(k) = c_n \cos \frac{2n\pi k}{N} + d_n \sin \frac{2n\pi k}{N} \tag{7}$$

The elliptic locus for the point ($X_n(k)$, $Y_n(k)$) is shown by removing the dependency on the sine and cosine terms to obtain :

$$\frac{\left(d_n^2 + c_n^2\right)X_n^2 + \left(a_n^2 + b_n^2\right)Y_n^2 - 2X_nY_n\left(a_nc_n + b_nd_n\right)}{\left(a_nd_n - b_nc_n\right)^2} = 1 \tag{8}$$

This involves that the points ($X_n(k)$, $Y_n(k)$) have elliptic loci and that the Fourier approximation to the original contour can be viewed as the addition in proper phase relationship of rotating phasors, which are defined by the projections [5], see figure 4.

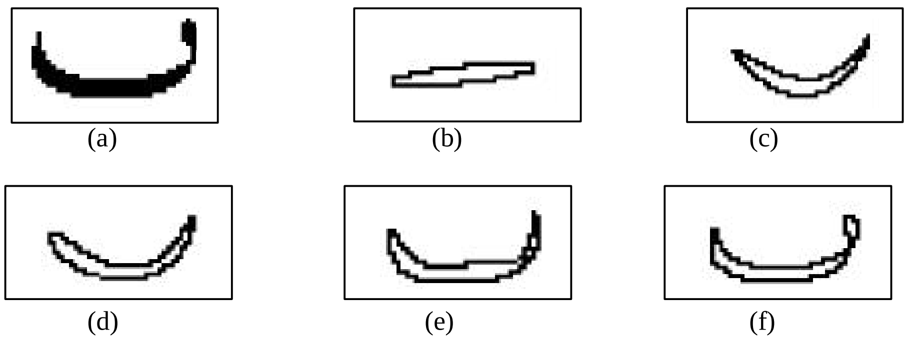


Fig. 4. Fourier approximation of the boundary of the character "ba ". Figure 4 (a) shows the 1^{st} harmonic, (b) the 1^{st} + 2^{nd} one, (c) the 1^{st} + 2^{nd} +3^{rd} one, (d) the 1^{st} + 2^{nd} +3^{rd} + 4^{th} one, (e) the 1^{st} + 2^{nd} +3^{rd} + 4^{th} + 5^{th} one, and figure 4 (f) presents the the 1^{st} + 2^{nd} +3^{rd} +.....+ 10^{th} harmonic.

4 Boundary Normalization

Since the Fourier coefficients vary according to the starting point of the Freeman chain code, the spatial rotation, the magnitude and the translation of the contour, normalization procedures are helpful. These procedures are applied on handwritten characters in order to approach them to reference printed characters.

4.1 Starting Point Transform

In order to obtain the same elliptic loci for different starting points, a step of normalization is necessary. Indeed, a difference in the starting points is displayed in the projected space as a phase shift. So, if the starting point displaced λ units in the direction of rotation (around the contour from the original starting point), we will have the following projections for $n \geq 1$:

$$\begin{aligned} X_n(k+\lambda) &= a_n \cos\frac{2\pi n}{N}(k+\lambda) + b_n \sin\frac{2\pi n}{N}(k+\lambda) \\ Y_n(k+\lambda) &= c_n \cos\frac{2\pi n}{N}(k+\lambda) + d_n \sin\frac{2\pi n}{N}(k+\lambda) \end{aligned} \qquad (9)$$

Expanding displaced X_n and Y_n and collecting terms to obtain the following Fourier coefficients transform :

$$\begin{bmatrix} a_n^* & c_n^* \\ b_n^* & d_n^* \end{bmatrix} = \begin{bmatrix} \cos\frac{2\pi n\lambda}{N} & \sin\frac{2\pi n\lambda}{N} \\ -\sin\frac{2\pi n\lambda}{N} & \cos\frac{2\pi n\lambda}{N} \end{bmatrix} \begin{bmatrix} a_n & c_n \\ b_n & d_n \end{bmatrix} \qquad (10)$$

Where coefficients a_1^*, b_1^*, c_1^*, d_1^* are corrections for the origin located at the displaced starting point. Therefore, the same elliptic loci would be obtained for different starting points. The starting point angular rotation is determined from the point (x1, y1) with elliptic loci by :

$$\begin{aligned} x_1 &= a_1 \cos\theta + b_1 \sin\theta \\ y_1 &= c_1 \cos\theta + d_1 \sin\theta \qquad \text{avec} \qquad \theta = \frac{2\pi k}{N} \end{aligned} \qquad (11)$$

$$E = \sqrt{(x_1^2 + y_1^2)} \qquad (12)$$

E is the magnitude of the first harmonic phasor
By differentiating the magnitude of the first harmonic phasor and setting the derivative equal to zero, we obtain :

$$\theta_1 = \frac{1}{2}\arctan\left[\frac{2(a_1 b_1 + c_1 d_1)}{a_1^2 + c_1^2 - b_1^2 - d_1^2}\right] \qquad (13)$$

This expression locates the first semimajor axis to occur moving away from the starting point in the direction of the rotation around the contour [5].

4.2 Harmonic Phasor Normalization

The rotating phasor provides the basis of a most convenient mode of normalization of the elliptic harmonic phasor. Indeed, the X, Y coordinate axis in which the contour was originally oriented are rotated into new U, V coordinate axis defined by the major and minor axis of the ellipse. The spatial rotation is determined from the Fourier coefficients a_1* and c_1* that are correct for starting point displaced θ_1. So, the point (x1*, y1*) with elliptic loci is :

$$
\begin{aligned}
x_1^*(k^*) &= a_1^* \cos \frac{2\pi}{N} k^* + b_1^* \sin \frac{2\pi}{N} k^* \\
y_1^*(k^*) &= c_1^* \cos \frac{2\pi}{N} k^* + d_1^* \sin \frac{2\pi}{N} k^*
\end{aligned}
\tag{14}
$$

Since k*=0 when the first harmonic phasor is aligned with the semimajor axis, ψ_1 is obtained as :

$$
\psi_1 = \arctan\left[\frac{y_1^*(0)}{x_1^*(0)}\right] = \arctan \frac{c_1^*}{a_1^*} \tag{15}
$$

This step of normalization transforms the first harmonic phasor to 0 as proposed by [5]. This is efficient to eliminate rotation variability between characters belonging to the same family, but there is a risk to reduce also the variability between the different characters. In order to avoid this risk, we introduce another idea in this paper : The change of the first harmonic phasor of handwritten character to the first harmonic phasor of the reference character ψ_f . This phasor transformation is described as follows:

$$
\begin{bmatrix} a_1^* & c_1^* \\ b_1^* & d_1^* \end{bmatrix} = \begin{bmatrix} \cos(y_1 - y_f) & \sin(y_1 - y_f) \\ -\sin(y_1 - y_f) & \cos(y_1 - y_f) \end{bmatrix} \begin{bmatrix} a_1 & c_1 \\ b_1 & d_1 \end{bmatrix} \tag{16}
$$

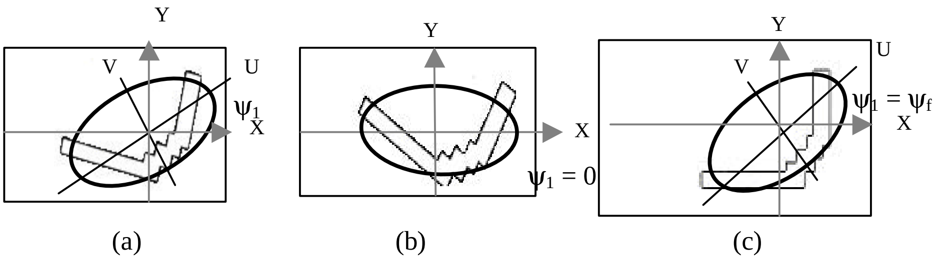


Fig. 5. Normalization of the first harmonic phasor of the character " Ra ". Figure 5(b) illustrates the normalization of the first harmonic phasor of character " Ra " (figure 5(a)) according to the proposition of [5]. His idea is to coincide U, V coordinate axis with X, Y coordinate axis. Our idea try to coincide U, V coordinate axis with those of the reference character. Figure 5(c) illustrates this.

4.3 Size Normalization

The normalization of the size can be made by dividing each of the coefficients by the magnitude of the semimajor axis defined as :

$$
E^*(0) = \sqrt{(x_1^*(0)^2 + y_1^*(0)^2} = \sqrt{a_1^{*2} + c_1^{*2}} \tag{17}
$$

This transformation proposed in [5] tries to reduce the first harmonic size of all characters to 1. This normalizes the variation of size between characters belonging to the same family but it reduces the variation between different characters. This can create ambiguity between characters having similar shapes, since their main difference occurs in their size. It is the case of " Ra" and " Del" characters (see table1 (5,2) resp (10,5) cell). In order to avoid this kind of problems, we propose to devise each of the coefficients of the handwritten character by the magnitude of the semimajor axis of the reference character :

$$\sqrt{a_f^{*2} + c_f^{*2}} \tag{18}$$

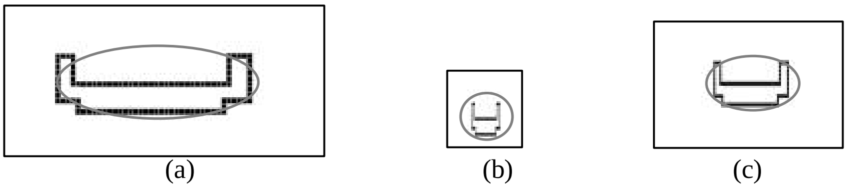


Fig. 6 Size transformation

4.4 Position Normalization

To make Fourier coefficients of our handwritten character independent of translation, we can ignore the bias terms A_0 and C_0. Figure 7 illustrates the impact of position normalization on the coordinate axis change [5].

Fig. 7 Position normalization

5 Experimentation and Results

This method of normalization taking into account the reference characters is experimented and analysed. A metric distance D(ma, Im), between images of handwritten characters and their printed references is defined as follow :

$$D(ma, Im) = \frac{\sum_{i=1}^{N} (a_{ma}^2 - a_{Im}^2) + (b_{ma}^2 - b_{Im}^2) + (c_{ma}^2 - c_{Im}^2) + (d_{ma}^2 - d_{Im}^2)}{\sum_{i=1}^{N} (a_{ma}^2 + a_{Im}^2 + b_{ma}^2 + b_{Im}^2 + c_{ma}^2 + c_{Im}^2 + d_{ma}^2 + d_{Im}^2)} \tag{19}$$

Where :
a_{ma}, b_{ma}, c_{ma}, d_{ma} : Fourier coefficients associated to the boundary of handwritten character.
a_{Im}, b_{Im}, c_{Im}, d_{Im} : Fourier coefficients associated to the boundary of printed character considered as the reference character for normalization.

Table 2 present the impact of normalization on the reconciliation of the handwritten character and the printed one. It shows the effect of the same normalization method in the separation between different characters. Indeed, the purpose of normalization is not only to decrease the variability between characters belonging to the same family but also to increase the variability between different characters. The two principal transformations that we present here correspond to the harmonic phasor and the size. Results of the four combinations of normalization are presented in table 2. We use the following abbreviations:

HC : Handwritten Character,
PC : Printed Character,
DbN : Distance before Normalization,
ADaN1 : Average of Distance after Normalization of first harmonic phasor to first harmonic phasor of reference character,
ADaN2 : Average of Distance after normalization of all harmonic phasors to first harmonic phasor of reference character,
DaN3 : Normalization N1 and size normalization by dividing each Fourier coefficients by the magnitude of the semimajor axis of the first harmonic,
DaN4 : Normalization N2 and size normalization by dividing each Fourier coefficients by the magnitude of the semimajor axis of the first harmonic,

For each character, 30 samples of multiscriptor handwritten characters are analysed and a computation of average distances is elaborated. We normalize each character in comparison to the angular rotation of its reference. A problem exists when a proposed character is different from the handwritten one. In this case, the distance between the two characters can be reduced by normalization but it remains greater than the distance between the normalized character and its right reference.

In this table each character is normalized using N1, N2, N3 and N4 and taking into account three different references. For instance, the handwritten character “ Alif ” is normalized by referring to printed characters “ Alif ”, “ ba ” and “ Ra ”. We notice that distances, in comparison to the right reference, remain the lowest for each kind of normalization. We can also remark that the normalization of angular rotation (regardless the size) improves the distances between the normalized character and its reference. However, the distance between the normalized “ Ra ” and its reference increases due to the great distortion of the handwritten character. This needs other kinds of normalization that we are working on. The normalization of the size reduces the distance between the characters belonging to the same family. It increases the distance between different characters.

6 Conclusion and Perspectives

In this paper, we present a normalization method for Arabic handwritten characters. Our purpose is to approximate handwritten forms (usually perturbed) to printed characters (more revealing extracted characteristics) in order to confirm the presence of some characters and the orthography completeness of the recognized word. Indeed, some handwritten characters cannot be easily described because of their large variability. The proposed normalization method tends to transform these characters to a well-known form (chosen as reference) in order to associate to them characteristics that would be comprehensible by the recognition system. This method is based on the Fourier Transform of Arabic character contours. It describes the boundary of handwritten Arabic characters by Fourier coefficients . Then, it applies geometric transformation to reduce handwritten variability within these coefficients. Hence, slant, size and position variability are mathematically calculated and adjusted.

So, this method allows us to produce an invariant description of the Arabic characters independently of their scriptor. This will perform the character recognition since it increases the distance between the distinct characters and reduces the distance between the identical characters.

We are also trying to transform not only the character slant, size and position, but also some of its details by using more than one harmonic. Indeed, the other harmonics can better describe the characteristics of the characters and would give other chances to distinguish between the different characters having close shapes.

Table 2. Distances between different characters (before and after normalization)

HC	**PC**	DbN	**DaN1**	**DaN2**	**DaN3**	**DaN4**
		0.0029	0.0016	0.0016	**0.0006**	**0.0006**
		0.0372	0.0212	0.0217	**0.0104**	0.0110
		0.0365	0.0944	0.0961	0.1012	**0.0995**
		0.0439	0.0456	0.0475	**0.0425**	0.0437
		0.0054	0.0048	0.0036	0.0053	**0.0038**
		0.1027	**0.0777**	0.0861	0.0799	0.0903
		0.0117	0.4694	0.3432	0.4234	**0.3430**
		0.1354	**0.4231**	0.4334	0.7917	1
		0.0008	**0.0535**	0.0823	**0.0535**	0.0823

7 References

1. Lee, S.W.,Park, J.S., Tang,Y.Y., Performance Evaluation of Nonlinear Shape Normalization Methods for the Recognition of Large-Sat Handwritten Characters ", ICDAR, IEEE, (1993), 402-407
2. Wakahara, T., Odaka, K., Adaptative Normalization of Handwritten Characters Using Global/Local Affine Transformation ", ICDAR-IEEE, (1997) 28-33
3. Hamanaka, M., Yamada, K., Tsulumo, J., On-Line Japanese Character Recognition Experiments by an Off-line Methid Based on Normalization-Cooperated Feature Extraction , ICDAR, IEEE, (1993), 204-207
4. Kim, S.Y., Lee, S.W., Nonlinear Shape Normalization Methods for Gray-Scale Handwritten Character Recognition, ICDAR-IEEE, (1997) 479-482
5. Kuhl,F.P., Elliptic Fourier Features of a Closed Contour , Computer Graphics and Image Processing, vol. 18, (1982), 236-258
6. Szmulo, M., Boundary Normalization for Recognition of Non Touching Non-Degraded Characters, ICDAR, IEEE, (1997), 463-466
7. F.Ghorbel, Stabilty of Invariant Fourier Descriptors and its inference in the Shape Classification, International Conference on Pattern Recognition, ICPR, IEEE, (1992), 130-134
8. Burdin, V., Ghorbel, F., de Bougrenet de la Tocnaye, J.L., Roux, C., A Tree-Dimentional Primitive Extraction of Long Bones Obtained From bi-Dimensional Fourier Descriptors, Pattern Recognition Letters 13, (1992), 213-217
9. Côté, M., Lecolinet, E., Cheriet, M., Suen, C.Y., Automatic Reading of Cursive Scripts Using a Reading Model and Perceptuel Concepts, IJDAR, Vol1, (1998), 3-17.
10. Snoussi Maddouri, S., Amiri, H., Une Méthode de Reconnaissance de Mots Manuscrits Arabes par Réseaux de Neurones Transparent, Conférence Internationale en Informatique (CII'99), Annaba, Algérie, (1999)
11. Snoussi Maddouri, S., Reconnaissance de l'Ecriture Arabe Manuscrite par Réseau de Neurones Transparent et Transformée de Fourier, Journée des jeunes chercheurs sur l'Ecrit et le Document, Colloque International Francophone sur l'Ecrit et le Document (JED-CIFED), Lyon, France, (2000).
12. Kunt, M., Traitement numérique des signaux, Edition Georgi, (1981)
13. Oppenhein, A.V., and Schafer, R.W., Digital Signal Processing , USA, (1995)
14. Biquard, F & M., Signaux Systèmes Linéaires et Bruit en Electronique, Paris, (1992)