Handwritten Arabic Word Recognition based on Ridgelet Transform and Support Vector Machines

Hassiba Nemmour and Youcef Chibani

Signal Processing Laboratory, Faculty of Electronic
University of Houari Boumediene
Algiers, Algeria
hnemmour@lycos.com, ychibani@usthb.dz

Abstract

In this paper, we propose a system for handwritten Arabic word recognition based on Ridgelet transform and SVMs classifiers. Specifically, the One-against-all SVM multiclass implementation is used to recognize several Arabic words. In addition, Ridgelet features are evaluated comparatively to Radon transform as well as to the results obtained for normalized word images. Experiments are performed on samples extracted from the IFN/ENIT database. The results highlight the reliability of the Ridgelet-SVM combination for handwritten Arabic word recognition.

1. Introduction

In many document analysis applications, words are considered as the unit component in the recognition task. However, despite of the impressive progress in handwriting recognition, accuracy of word recognition systems is far from the human performance. In fact, due to the high variability in writing styles and tools as well as the unconstrained handwriting aspect, words are quite complex patterns [1]. Specifically, there are two approaches for handwritten word recognition [2,3]. The first approach is analytic; it proceeds by a segmentation into isolated characters. Then, words are recognized through character recognition. This approach is useful for problems with very large vocabulary where it becomes impossible to refer each word to a class of interest. The second approach is called holistic since it considers the whole word as a single entity corresponding to a class of interest. The holistic analysis is adequate for problems with medium and small vocabulary such as literal amount reading in bank checks [2,4]. Another advantage of this approach is that it captures all-coarticulation and variability effects into word images [1]. Recall that in the last years various research works were carried out to enhance handwritten Latin word recognition [2,5]. More recently, a progressive interest was assigned to Arabic script [4,6,7,8,9]. Nevertheless, unlike Latin, Arabic language has a set of particularities which make its recognition too challenging. For instance, Arabic is semi-cursive since a single word can include many cursive sub-words or connected-components. In addition, several diacritics are used in Arabic writing. Moreover, character shapes change according to their positions within the word [10].

In the present work, we are interested to Arabic word recognition. Note that researches in this field were inspired from methods that are already used for Latin word recognition. Thereby, Hidden Markov Models (HMM) were extensively used especially for problems with very large vocabulary [2,3,4,5,6,10,11]. Note that over the past years, HMM and then Neural networks were the classical choice to achieve the recognition task. Currently, Support Vector Machines or SVMs are the best candidate for solving all handwriting recognition tasks with medium number of classes [12]. The literature reveals a large use of SVMs especially for handwritten digit recognition [13,14]. For feature extraction, earlier works employed statistical descriptors such as relative positions, mean and variance based on profiles or histograms [15]. Due to the diacritics and spaces within Arabic words, structural features that are computed on skeletons of connected components such as the number of loops and concavities are more reliable [4]. Recently, some robust transforms such as wavelet packets have shown satisfactory results [15]. Thus, in this work we propose the use of SVM for handwritten Arabic word recognition. Besides, we investigate the applicability of Ridgelet transform for feature extraction. The Ridgelet is obtained by applying one-dimensional wavelet transform to Radon coefficients. Since, there is no reference using the Radon transform for handwritten Arabic words, the Ridgelet transform is evaluated comparatively to the results obtained with Radon features. This paper is then organized as follows. In the next section we give a brief presentation on SVMs classifiers. Section 3 presents the Ridgelet transform while the experimental results are given in section 4. The final section gives the main conclusions of the paper.

2. Support vector machines

SVMs are designed for solving binary classification. Their training consists of finding the optimal separating hyperplane between two classes [16]. Specifically, let
\((P_n,y_n) \in \mathbb{R}^M \times \{\pm 1\}\) be a set of training patterns so that \(M\): is data dimension, \(\{n = 1, \ldots, N_c\}\), \(N_c\) : is the number of samples in the class \(c\) and let \(f : \mathbb{R}^M \rightarrow \{\pm 1\}\) be a set of functions. The training process selects the function \(f\) which maximizes the margin between the two classes [17]. Then, data are classified according to the signal of the decision function so that :

\[
f(P) = \text{sign} \left( \sum_{i=1}^{SV} y_i \alpha_i k(P,P) + b \right) (1)
\]

The optimal hyperplane corresponds to \(f(x) = 0\) while \(b\) is a bias. \(SV\) is the number of support vectors, which are training data whose Lagrange multipliers \(\alpha_i\) are not zero. The kernel \(k(\ldots)\) is any mathematical function, respecting Mercer’s conditions [17]. The Radial Basis Function (RBF) kernel provides commonly the best performances for pattern recognition applications. For two samples \(P_1\) and \(P_2\), this kernel is expressed as:

\[
k(P_1, P_2) = \exp \left( -\frac{1}{2\sigma^2} d(P_1, P_2) \right) (2)
\]

\[
d(P_1, P_2) = \| P_1 - P_2 \|^2 (3)
\]

\(\sigma\) is user-defined.

For solving multi-class problems, two multi-class implementations of SVMs are commonly used [18]. The One-Against-All (OAA), which is the earliest one, performs \(C\) binary SVMs to solve a problem with \(C\) classes. The second approach that can be used is the One-Against-One (OAO) which employs \(C(C-1)/2\) SVMs each of which is designed to separate two classes.

3. Feature extraction

Presently, we investigate the applicability of the Ridgelet transform for describing handwritten Arabic word characteristics. Recall that the use of global transform for handwriting recognition is not new. After Fourier and Hough transforms, more efficient transformations such as wavelet and wavelet packets were used for handwritten Arabic word recognition [15]. Nevertheless, the wavelet transform performs well for describing point singularities while handwritten words are composed of linear singularities. Furthermore, the Ridgelet transform has been introduced to image processing in order to improve the characterization of line singularities [19]. Ridgelet is based on Radon transform that is computed according to several angular directions. Radon coefficients correspond to projections representing the shadow of the shape at each angle [20]. Consequently, linear singularities in any direction are expressed by high magnitudes. Thenafter, in order to characterize better linear singularities, the one-dimensional wavelet transform is applied on Radon slices to yield the Ridgelet coefficients. Hence, along the Radon axis projection, the Ridgelet is constant while in the direction orthogonal to these ridges it is a wavelet [19].

For an image \(f(x_1, x_2)\) the Ridgelet transform can be computed according to the following steps:

1) Compute the Radon transform such as [19]:

\[
RA(\theta, r) = \int f(x_1, x_2) \delta(x_1 \cos \theta + x_2 \cos \theta - r) dx_1 dx_2 (4)
\]

\(\delta\): Dirac distribution.

\(\theta\): Angular variable.

\(r\): Radial variable.

2) Apply the 1D wavelet on each Radon slice in order to obtain the Ridgelet coefficients \(R(\theta, r)\).

The Ridgelet transform has been successfully used in various image processing applications such as texture classification [20]. For character recognition it was employed for printed Chinese character recognition [19]. Unfortunately, there are no references on the use of Ridgelet transform for handwriting recognition. Nevertheless, some other transforms such as the curvelet have been employed for handwritten Bangla character recognition [21] while Radon transform was used for signature verifification [22]. For our part, we propose the use of Ridgelet transform for improving handwritten Arabic word recognition in a medium vocabulary problem. We investigate also the performance of SVMs for the recognition stage.

4. Experimental results

The Ridgelet-SVM recognition system was evaluated on data extracted from the well known IFN/ENIT database [23]. This database contains 26400 images of Tunisian town names. Words are written by more than 411 scripts using different writing tools. In our experimental analysis, we collected 24 classes of interest grouping names with the largest appearance frequencies in the ENIT database. This database has a set of particularities, which make its training a challenging task. As shown in figure (1), words of classes 10 and 11 differ only in the first letter. In addition, samples of class 13 are composed of a word and digits while those of the class 22 are composed of two words. Another particularity of this database is related to the size of data, which is quite variable.
Figure 1. Examples from the selected database

The SVM implementation is based on the OAA approach using the RBF kernel. Note that several run passes were carried out in order to find the best parametric selection. Radon transform was computed without size normalization. Then, to obtain a feature vector with a unified size, the idea was to fix the number of both radial projections and angular directions specific to Radon transform. However, the results in terms of Recognition Rate (RR) showed that the derived Radon features provide weak discrimination ability. This outcome indicates the necessity to normalize the data size before computing the Ridgelet transform. Then, several sizes were tested by considering normalized images as input vectors for SVMs. The most meaningful results are listed in table (1).

Table 1. Overall Recognition Rates obtained for normalized data

<table>
<thead>
<tr>
<th>Normalization</th>
<th>RR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20x70 pixels</td>
<td>80,3</td>
</tr>
<tr>
<td>30x80 pixels</td>
<td>79,7</td>
</tr>
<tr>
<td>30x100 pixels</td>
<td>80,2</td>
</tr>
</tbody>
</table>

The best recognition rate is obtained when images are normalized to a size of 20x70 pixels, which allows also the smallest size for the feature vector. Therefore, the Radon transform was computed by using this normalization. Nevertheless, to reach a successful recognition with Ridgelet transform, the related Radon matrix should precisely reflect linear singularities in word images. Thereby, before computing Ridgelets, several parameter selections were tested for Radon transform. Specifically, different choices for the number of angular directions (or the number of θ) and radial projections (or the size of r in pixels) were evaluated in terms of Recognition Rate. Note that if the size of Radon matrix is too large, it will lead to a very large feature vector and the corresponding recognition rate is weak. Moreover, the best results were obtained when we used square Radon matrices. Variations of the Recognition Rate according to the size of Radon matrix are plotted in figure (2). It is easy to see that a 32×32 Radon matrix makes the best compromise between recognition accuracy and the size of the feature vector which has in this case, 1024 components.

Table 2. Performance evaluation of Ridgelet transform

<table>
<thead>
<tr>
<th>Radon matrix</th>
<th>Radon</th>
<th>Ridgelet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Size</td>
<td>RR(%)</td>
</tr>
<tr>
<td>128×128</td>
<td>16384</td>
<td>83,3</td>
</tr>
<tr>
<td>64×64</td>
<td>4096</td>
<td>83,5</td>
</tr>
<tr>
<td>32×32</td>
<td>1024</td>
<td>83,3</td>
</tr>
<tr>
<td>16×16</td>
<td>256</td>
<td>81,8</td>
</tr>
</tbody>
</table>

Roughly speaking, Ridgelet features computed on small Radon matrices (less than 64×64 Radon features) improve both the Recognition rate and the runtime. Specifically, the best performance is obtained when using a 32 × 32 Radon matrix. In this case, Ridgelet transform provides a gain of 0,7% over Radon features and more than 3% compared to normalized data. In
addition, the Ridgelet requires small runtimes because it significantly reduces the feature vector size. Note that, Ridgelet features are composed of both approximation and details of the wavelet decomposition. Then, in the last test we evaluated the feature selection of the Ridgelet transform. Since we are using the one dimensional wavelet, the number of details is controlled by the number of wavelet decompositions which depends itself on the Radon matrix size. Thereby, with a matrix of $32 \times 32$ Radon coefficients the number of decomposition levels can go to 5. The results in terms of Recognition Rate showed that after the 3rd decomposition level, the Ridgelet keeps approximately the same performance. On the contrary, when changing the wavelet coefficient selection, the Recognition Rate varies significantly. In fact, there are three possible configurations which are: approximation with all details (APP+ALL DET) which provides a feature vector of 512 components, the approximation (App) which yields a feature vector of 128 elements, and all details (ALL DET) with vectors of 384 features. As shown in figure (4), if we consider only the approximation, the recognition accuracy is altered to 70%. On the other hand, the best performance is obtained when considering all details. The most important remark is that the approximation does not improve the precision when it is added to details. Hence, the all details configuration constitutes the best selection for the Ridgelet transform since it improves the recognition rate while reducing the size of data.

![Graph showing Recognition Rate for different selections of the Ridgelet coefficients](image)

**Figure 4.** Recognition Rate for different selections of the Ridgelet coefficients

5. Conclusion

In this paper we proposed a Ridgelet-SVM combination for handwritten Arabic word recognition. The Ridgelet transform is employed in order to highlight all linear singularities within handwritten words. Using data extracted from the IFN/ENIT database, the experimental design was focused on finding the best Ridgelet configuration for Arabic word characterization. The results obtained in the different tests, highlight the efficiency of the proposed recognition system. In fact, the Ridgelet outperforms both Radon and normalized data. Nevertheless, is has been shown that an appropriate parametric selection should be carried out to reach satisfactory precision. Therefore, many other tests should be drawn such as the choice of the wavelet function as well as its parameters can improve again the recognition rates.

References


[23] www.ifnenit.com