A simplified feature vector obtained by wavelets method for fast and accurate recognition of handwritten characters off-line

Carlos Ramírez Piña[1], Vianney Muñoz-Jiménez[1], Rosa Maria Valdovinos Rosas[1], and J. A. Hernández Servín[1] ⋆

Facultad de Ingeniería,
Universidad Autónoma del Estado de México,
Toluca, Estado de México.
{cramirezp,vmunozj,rvaldovinosr,xoseahernandez}@uaemex.mx

Abstract. This paper presents an algorithm for simplified features extraction based on a wavelet method for off-line recognition of handwritten character. The proposal is applied to a set of 3250 handwritten symbols, which include the digits and the upper and lowercase character of English alphabet. The effectiveness of our algorithm is tested by comparison against the descriptors *FKI* and *Wavelets* using the Nearest Neighbour rule as classifier. The classification is measured in percentage of overall Accuracy and the processing time obtained by each methods.

1 Introduction

The study of character recognition is divided into off-line and on-line methods mainly [1]. The difference between them lies on how handwriting is done and analyzed. For the off-line recognition, the data are taken to be a static representation of text, since it can not be establish the order on which they were produced by a machine or handwritten [2]. On the other hand, in the on-line recognition, the original data are glyphs and points, which are normally storage on regular intervals of time [3].

This paper is focused on the off-line recognition of handwritten characters. The study is based on descriptors such as FKI [4] and discrete wavelets [5]. The dataset used in this work have been generated by [6] which includes digits and characters ($0-9$, A-Z, a-z). Our proposal was compared with the descriptors FKI and the discrete wavelet, in accuracy and processing time terms using the Nearest Neighbour rule *1-NN* as classifier.

1.1 The FKI offline features

The FKI algorithm was proposed by [4] which obtain a set of geometric features that has been used in handwriting recognition. That is, given a binary image

⋆ Corresponding author

$S(x, y)$ of size $M \times N$, the method computes nine geometrical features c_i where $i \in \{1, ..., 9\}$ for each entry column x such that $1 \leq x \leq M$. This is done on each column of the image, thus the method obtain $9N$ features in total. The authors also have features such as number of black and white pixels and their transitions, centre of gravity and second order moments.

1.2 Wavelets Descriptors

The wavelets are transformations which decompose an image into multi-resolution descriptions localized in space and frequency domain providing a smaller frames of the images. The frequency domain analyse different variations that has been successfully used in many image processing applications [7].

The DWT decompose the image S into wavelet blocks, an average image of smaller size than the original for a factor of two, and three more images containing the gradients and contours of itself, according to the following definitions:

$$W_g(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} S(x, y) g^i_{jmn}(x, y) \tag{1}$$

$$W^i_h(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} S(x, y) h^i_{jmn}(x, y) \tag{2}$$

where g is $g(x) = \begin{cases} 1 & x \in [0, \frac{1}{2}] \\ -1 & x \in [\frac{1}{2}, 1] \end{cases}$ and h belongs to the *Daubechies* family of mother wavelets; where as before $i \in \{H, V, D\}$. The wavelet blocks will be denoted by $A_j = W_g(j, m, n)$, $H_j = W^H_h(j, m, n)$, $V_j = W^V_h(j, m, n)$ and $D_j = W^D_h(j, m, n)$ where j is an index that indicates level of decomposition of the image (see Figure 1 (b)).

Frequency domain analysis is the background of representation of the feature vector. Different textural and statistical values are also computed which enrich the feature vector, like mean (μ) and standard deviation (σ) [5]. The type of entropies in the reference, which we have also implemented for comparison to our proposal, are like *shannon*, *Log energy*, *threshold*, *sure* and *norm*, which are computed on approximation the A_j coefficient block, as illustrated in Figure 1 (a).

2 Our Proposal

The main objective of the proposal method is to obtain an strategy which combine feature extraction methods in handwritten characters off-line and the recognition process of these characters in an accurate way. For that, segmentation and binarization methods were used before the actual feature extraction.

2.1 Binarization and segmentation

A pre-processing to the image is applied before feature extraction in order to eliminating noise of the image. In this way, firstly the images are converted into a binary type by analysing their histogram in a gray scale, in order to determine the optimal cut threshold. On a second stage, the symbol image is segmented extracting pixels corresponding to the symbol only. Finally, the symbol image are resized to a fixed size of 120×120. The size has been fixed in order to get optimal results when the wavelet transform is applied.

2.2 Feature extraction by a simplified vector feature using wavelets method

Feature extraction in the context of image processing, specifically in handwriting character recognition, is based on two types [8]; structured and statistical methods. The first one, are derived from the probability distributions of pixels, e.g. zones, first and second moments, projection and direction histograms. The second one, are based on topological and geometrical properties of the object under study.

The Wavelet transformation is used to compress an image by transforming it into the frequency domain [9]. In order to accomplish this, the image are represented using a set of basic functions produced by translation and scale up of a mother function. Let $S(x, y)$ be an input image, where x, y represent indexes, whereas $S(x, y)$ is the pixel value. In this paper, a 2D wavelet transform is used, the scaling of $S(x, y)$ is given by the functions g and h.

Coefficients wavelet analysis are obtained from three blocks; it was observed that wavelet coefficient of the third block are features of the input image, that is, it maintains representative information of the symbol. The wavelet transformation for the third state generate four images of size 15×15, A_2, H_2, V_2 and D_2 with 17 features correspondlly. The information from the approximation coefficients A_2 in third block keeps the information of the input image and the other four coefficients obtained represent 12% of the original image size and 25% of the size of the A_0 coefficient.

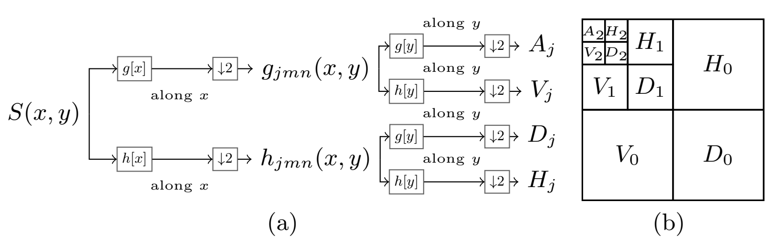

Fig. 1: (a) Block diagram for calculating the DWT , (b) Wavelet decomposition indicating the block coefficients, A_j, etc.

For each coefficient obtained, were calculated the median, entropy and standard deviation; additionally five entropy wavelets are also calculated: Shannon, Log energy, Threshold, sure and norm; with this in mind we are reducing an amount of 77% the statistical measures as compared with the original method.

The Algorithm 1 represent the feature extraction of the vector formed by 21 features proposed for this study.

Algorithm 1 Simplified vector feature using Wavelet method

Require: Gray scale input image
Ensure: Set of 21 features
1: Convert image to binary type
2: Apply the wavelet transform to obtain the coefficients of the third block A_2, H_2, V_2, D_2 thus obtainig four features.
3: Calculate the mean (μ), standard deviation (σ), entropy (E) thus giving 12 features
4: Calculate the entropies shannon, Log energy, threshold, sure, norm from A_2 thus generating five features at this stage.
5: Repeat steps 1 to 4, for each symbol image in order to form its feature vector.

3 Tools and Methods

3.1 Data set

The results here reported correspond to the experiments over the data set generated by [6], which includes digits $0-9$ with 10 classes and 527 feature vectors, the uppercase characters $A-Z$ form 26 classes and 1402 feature vectors, the lowercase characters $a-z$ with 26 classes and 1321 feature vectors.

For the data, the 10-fold cross-validation method was employed to estimate the classification error: 80% of the available patterns were for training purposes and 20% for the test set. On the other hand, we use as base classifier the 1-NN rule, expressed as [10]:

$$\delta_E(V_1, V_2) = \sqrt{\sum_{j=1}^{e}(V_1[j] - V_2[j])^2} \tag{3}$$

Where δ_E is the euclidean distance between vectors V_1 test feature and V_2 training feature .

3.2 The configuration of the method

The experiments were carried out datasets with different dimension of the feature vector, depending on the method used. That is:

– The FKI method, obtain nine features by column that containing the image, therefore the feature vector will have nine features by the number of columns that containing the image.
– Wavelts method obtain 55 features. The vector dimension is computed by the matrix of A_0, which generates $(\frac{x}{2})$ $(\frac{y}{2})$ features, where, x and y are the original image size, plus 54 features which represent the statistical averages.
– The Simplified vector features using Wavelet method obtain a vector with 21 features. That is, the whole of the features is $(\frac{x}{8})$ $(\frac{y}{8})$ plus 17 features which represent the statistical averages.

4 Results and Discussion

In this paper, we study two descriptor methods: FKI and Wavelets, in comparison with our wavelets method for recognition of handwritten characters off-line, in *Accuracy* and processing time terms. The *Accuracy* is obtained as follow:

$$Accuracy = 1 - \frac{M_e}{M}\,, \tag{4}$$

where M_e is the number of misclassified samples and M is the number of training samples.

4.1 Classifier performance

Figure 2, shows the 1-NN classification resuslts for each feature selection method here studied. The y axis represents the *Accuracy*, x axis correspond to the class. Some comments about these results are: First, it is clear that the recognition obtained from each method is not uniform by each class, however, for the digit dataset our method proposed shows a uniform behaviour with an average accuracy of 93.8%. On the other hand, the upper case dataset Wavelet method shows an uniform behaviour having an average accuracy of 93.0%. Finally, with the lower case dataset, Wavelet method abtain an average accuracy of 88.0%.

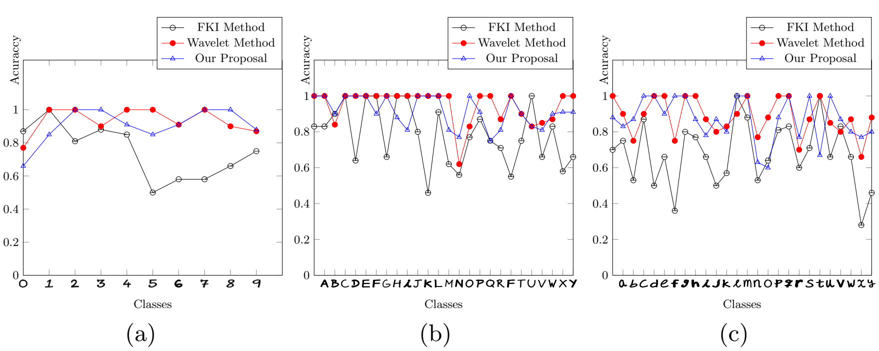

Fig. 2: Overall Accuracy (a) Digits dataset, (b) Uppercase characters dataset, (c) Lowercase characters dataset.

In order to identify the statistic significance between the methods, the Table 1, shows the average accuracy for each dataset, bold values represent the best results. For that, the rank of each method was calculated as follows [11]: For each dataset, the method with the best accuracy receives rank 1, and the worst receives rank 3. If there is a tie, the ranks are shared. Thus the overall rank of a method is the averaged rank of this method across the data set used. The results shown that the highest rank is obtained by the Wavelet method and the method with lowest rank is the FKI method.

Table 1: Overall Accuracy Performance

	FKI		Wavelet		Our proposal	
Dataset	μ	*Rank*	μ	*Rank*	μ	*Rank*
$0 \dots 9$	77	(3)	89	(2)	**93**	**(1)**
A ... Z	78	(3)	**93**	**(1)**	88	(2)
a ... z	67	(3)	**88**	**(1)**	86	(2)
Average accuracy	74		90		89	
Average Rank		3		1.3		1.6

To complete the analysis of statistical significance between the results, the Namenyi test is used [11] $DC = q_\alpha \sqrt{\frac{K(K+1)}{6N}}$, where q_α is critical value, K is the number of methods to compare and N is the number of training set used. The test obtains a critical difference (CD) to reject the assumptions on which the corresponding p value is less than the adjusted α. In this paper we compare three feature selection methods and analyse their behaviour on three different datasets; the corresponding value for q_a are: $q_{0.05}$ is 2.343 and for $q_{0.10}$ is 2.052. The critical difference for $q_{0.5}$ is 1.913 and for $q_{0.10}$ is 1.675.

To interpret the results it is stated that a particular method A is significantly different than B, if the overall rank $(A) + CD < \text{rank}(B)$. From results in Table 1 it is posible to identify that the behaviour of our method and the Wavelets method do not offer statistic difference, that is to say that it is competitive with the Wavelets methdo. However, comparing the resulst respect to the FKI method, the Wavelets method is significatively better (1.3 (Wavelets Rank) $+1.675(CD_{0.10}) < 3$ (FKI Rank)).

4.2 Processing time

Table 2 shows the processing time using the different methods here studied.

As Table 2 shows, with our proposal the size of feature vector has less entries, in consequence it requires less processing time compared to the others methods. If we recall the classification results from the Table 1, our algorithm proposed show better accuracy than the FKI method and clearly competes with Wavelet method, reducing execution time with the classifier 1-NN.

Table 2: Processing time

Method	Features vector	Run time(sec.)
FKI	1082	1573.63
Wavelet method	3653	5183.59
Our proposal	915	1313.74

5 Conclusions and future work

In this paper we propose a method for reducing the feature vector for handwriting recognition in comparison to the results reported by [5], in which method obtain a vector with 55 features. Our method obtain a feature vector of 21 features only, using the third moment of the wavelet transformation. This allow us to reduce processing time compared to the FKI and traditional wavelet methods. That means, our algorithm reduces the processing time from 74.65% to 16.51% and decrease in size vector from 74.87% to 15% respect to FKI and Wavelet method respectively.

The future work will be focus on the processing of the dataset generated through a simplified vector feature using Wavelet method. We are in search to improve accuracy of the classifier by using the multilayer perceptron.

Acknowledgment. This work has partially been supported by the SEP-PRODEP-3238 and 3834/2014/CIA Mexican Projects and by the Mexican Science and Technology Council (CONACYT-Mexico) through the Masters scholarship 702528.

References

1. Fotini Simistira, Vassilis Katsouros, and George Carayannis. Recognition of online handwritten mathematical formulas using probabilistic {SVMs} and stochastic context free grammars. *Pattern Recognition Letters*, 53:85 – 92, 2015.
2. Ernesto Tapia. A survey on recognition of on-line handwritten mathematical notation. In *Technical Report B-07-01*. Freie Universität Berlin, Germany, 2007.
3. Ernesto Tapia. *Understanding mathematics: A system for the recognition of on-line handwritten mathematical expressions.* PhD thesis, Freie Universität Berlin, Germany, 2005.
4. F. Álvaro, J. A. Sánchez, and J. M. Benedí. Offline features for classifying handwritten math symbols with recurrent neural networks. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 2944–2949, Aug 2014.
5. Sk Md Obaidullah, Chayan Halder, Nibaran Das, and Kaushik Roy. Numeral script identification from handwritten document images. *Procedia Computer Science*, 54:585–594, 2015.
6. Teófilo Emídio de Campos, Bodla Rakesh Babu, and Manik Varma. Character Recognition in Natural Images. *Proceedings of the International Conference on Computer Vision Theory and Applications, Lisbon, Portugal*, February 273–280, 2009.

7. K. B. Raja, S. Sindhu, T. D. Mahalakshmi, S. Akshatha, B. K. Nithin, M. Sarvajith, K. R. Venugopal, and L. M. Patnaik. Robust image adaptive steganography using integer wavelets. In *Communication Systems Software and Middleware and Workshops, 2008. COMSWARE 2008. 3rd International Conference on*, pages 614–621, Jan 2008.
8. Hedieh Sajedi. Handwriting recognition of digits, signs, and numerical strings in Persian. *Computers & Electrical Engineering*, 49:52–65, 2016.
9. R Colom, Rafael Gadea, A Sebastiá, Marcos Martínez, Vicente Herrero, and Vicente Arnau. Transformada Discreta Wavelet 2-D para procesamiento de video en tiempo real. *Actas de las XII Jornadas de Paralelismo*, 2010.
10. Cristina García Cambronero and Irene Gómez Moreno. Algoritmos de aprendizaje: knn & kmeans. *Intelgencía en Redes de Comunicación*, 2006.
11. Janez Demšar. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, 7:1–30, 2006.