3rd International Conference on Automation, Control, Engineering and Computer Science (ACECS'16)

Proceedings of Engineering & Technology (PET)
pp. 18-22

A hybrid Wavelet transform with SOM and soft kNN ensemble Classifier for Single Sample Face Recognition

Imen Mbazaâ#1, Noureddine Ellouze#2

#Department of Electrical Engineering, National Engineering School of Tunis

BP 37, Belvédère 1002 Tunis, Tunisia

[1]imene.mbazaa@enit.rnu.tn

***Abstract*— Since providing several images per subject is hardly possible in practice case, this work has been carried out under the assumption that only one image is available per person. The baseline method for the proposed automatic face recognition process consists of exploiting Discrete Wavelet Transform (DWT) for image representation and Self-Organizing Maps (SOM) neural networks combined with K-nearest neighbors (kNN) ensemble while image hybrid classification. The addressed technique leads to satisfactory recognition accuracy rates.**

***Keywords*— Face recognition, single sample face, hybrid classifier, wavelets transform, self-organizing map, k-nearest neighbor ensemble.**

I. Introduction

During the past two decades, face identification technology (FIT) has known an important evolution thanks to several factors, among others representation face model improvement. Representation models can be categorized into linear models such as fisherface and eigenface[1], and nonlinear ones.

Self-organizing maps model has proved its performance as nonlinear one, since it guaranties vector quantization in addition to topographic representation [2]. Among others, T. Huntsberger, J.Rose and S. Ramaka implemented a face recognition system called Fuzzy-Face that combines wavelet preprocessing as input to a fuzzy SOM [3].

Furthermore, the so-called "one image per person" constraining problem has been again reconsidered by researchers [4-5], since it is generally difficult to have at disposal several images per person. In fact, this restriction may affect employed techniques generalization basic property. Particularly, those built on the intra-class variations such as Linear Discriminant Analysis (LDA) [6-7], Bayesian matching methods [8] and Evolutionary Pursuit (EP) [9] that require large training sample sets to provide decent images generalization subspace. In line with this, Tan and Chen presented kernel SOM face to deal with both the nonlinear problem and the small sample problem in FI [10, 11].

In this work, we performed in a first stage discrete wavelets transform on facial images. Its multiresolutional representation provides coarse facial structure in approximation channel and fine features in detail one. Thus, thanks to discrete wavelets transform, images scale is reduced while local properties (shape and texture) are kept. Then, in a second stage, only recovered approximation coefficients were reassigned to the unsupervised and nonparametric self-organizing maps neural network to ensure images robust and compact representation. To avoid neurons wrong labeling, because of hard labeling, soft *k* nearest neighbor assignment was adopted.

The paper proceeds as follows: the proposed method is detailed in section 2. Data set and methodology are described in section 3. Experiments under FERET database are reported in section 4. Finally, conclusions and suggested future work are drawn in section 5.

II. The proposed method

Given M training images, to decide to witch one a test face x belongs to, WTSOM-kNN face recognition proposed approach processes as summarized below:

A. *Pre-processing*

All images (gallery and test) are normalized to ensure that data are impartially treated during the training step. Since we are handling grayscale images, we proceeded by histogram equalization. Thanks to this adjustment, the intensities would be better distributed on the histogram. From each normalized image, only the face is localized and cropped thanks to Viola Jone's cascade object detection algorithm.

B. *Descrete Wavelets transform*

Based on a mother wavelet, discrete wavelet transform is applied to images decomposing them into approximation and detail coefficients. Since the former ones correspond to the low-pass frequency sub-band, they enclose more than 80% of the whole image information. For this reason and in the interest of reducing features size while maintaining data, only the approximation coefficients are kept in the rest of this work. Thus, each low-pass image sub-band is partitioned into N non-overlapping and equally sized sub-blocks constituting its extracted features face.

C. *SOM neural networks training*

As stated in [14], unsupervised feature clustering helps to improve the performance of the recognition system. Furthermore, it assesses a huge data bulk to few weighted items. Thus, our choice for the Self Organizing Map unsupervised neural network was belt on these arguments and the already obtained features vectors from the previous stage are submitted as input to it. After SOM training on gallery images, each sub-block of the test image is mapped to its Best

Matching Unit (BMU) which is its nearest neighbor in the SOM topological space.

The pair-wise distances between each BMU sub-block of the test image and its homologous ones of each training image, are calculated. Therefore, for a sub-block l of the test image, if the neuron n_i designate its BMU and n_j that's of the corresponding l^{th} sub-block of the j^{th} training image, the distance vector d_l :

$$d_l = [d_{i1} \quad d_{i2} \quad \ldots \quad d_{ij} \quad \ldots \quad d_{iM}]$$

represents the pair-wise calculated distances related to the l^{th} sub-block test image. For all sub-blocks, the achieved distance matrix is:

$$D = \begin{pmatrix} d_{11} & d_{12} & \ldots & d_{1j} & \ldots & d_{1M} \\ \ldots & & & & & \\ d_{i1} & d_{i2} & \ldots & d_{ij} & \ldots & d_{iM} \\ \ldots & & & & & \\ d_{N1} & d_{N2} & \ldots & d_{Nj} & \ldots & d_{NM} \end{pmatrix}$$

Consequently, the distance matrix D reports the dissimilarity/similarity between the test image x and every training image of the gallery G relatively to the SOM topological space.

D. *kNN ensemble decision*

During face recognition process, particularly when achieving the distance matrix D, to avoid neurons wrong labeling -because of hard labeling- soft k nearest neighbor assignment was performed.

According to this approach, for each sub-block and for a given number k, k nearest neighbors are identified where corresponding to the k smaller previously distances of the distance vector d_l:

$$dl^* = [d^*_{i1} \quad d^*_{i2} \quad \ldots \quad d^*_{ik} \quad \ldots \quad d^*_{iM}]$$

Where: $d^*_{i1} < d^*_{i2} < \quad \ldots \quad < d^*_{ik}$

As inspired from [11, 13], a confidence value for the k^{th} nearest neighbor is defined as:

$$c_{ik} = \frac{\log(d^*_{i1} + 1)}{\log(d^*_{ik} + 1)}$$

Finally, to characterize the test image label, all the outputs related to its N sub-blocks will be combined according to the final decision rule [13]:

$$Label(x) = arg \max_k \left(\sum_{i=1}^{N} c_{ik} \right), \qquad k = 1, \ldots, G$$

III. Data Set and Methodology

To proof the efficiency of the proposed method, namely the WTSOM-kNN algorithm, we have run several experiments on the well-known color FERET database. In fact, FERET Database contains a total of 11338 facial images of 1208 subjects, affording a wide variety across gender, ethnicity and age. It comprises 13 subsets, each of them representing images of the same pose. The carried experiments interested (*fa*) gallery images and (*fb*) probes subsets, where (*fa*) designates the regular frontal pose images subset, and (*fb*) the alternative frontal images subset, token shortly after the corresponding (*fa*) ones. In the recognition process, (*fa*) gallery set was used for training and (*fb*) probe set for testing.

Initially, training and testing images are in '.ppm' file format and of 512 by 768 pixels. During the pre-processing stage, they were converted to grayscale images ('.pgm' file format), normalized according to the histogram equalization approach and finally, belonging faces cropped to a size of 256*256 in order to perform dyadic wavelet transform.

In fact, after choosing the mother wavelet, discrete wavelet transform (DWT) is achieved on both training and testing images, with decomposition level equal to 3 to reduce their size to 32*32. Only the approximation coefficients were retained. Each corresponding low-pass sub-band was partitioned into non-overlapping sub-blocks with equal size of 4*4 and 2*2 respectively.

Self organizing maps were trained in batch mode through the retrieved sub-blocks. The training process was run into two phases as argued in [13], specifically, a rough training phase to adjust the topological order of features vectors and a fine tune one carried on the first phase achieved map to provide an accurate quantification of the input space. In line with this, 100 updates were performed in the first phase, while 400 times in the second one. The initial weights of all neurons were set to the greatest eigenvectors of the training data, and the neighborhood widths of the neurons converged exponentially to 1 with the increase of training time.

The soft kNN ensemble classifier described before was used for final classification decision.

IV. Experiments

In this section, practical properties of the proposed WTSOM-kNN approach, such as mother wavelet, soft *kNN* parameter and gallery size, are experimented:

A. *Experiment 1:*

Considering 200 training images and 200 testing ones, we investigated the proposed method performances while varying mother wavelet and sub-block dimensions. Accomplished results are displayed in Table 1:

TABLE I
ACHIEVED ACCURACY RATES ACCORDING TO MOTHER WAVELET AND SUB-BLOCK SIZE SETTINGS

Mother wavelet	**Sub-block size**	
	4*4	**2*2**
Haar	80.5	82.5
Daubechies 2	81.5	93.5
Daubechies 4	80	93.5
Daubechies 6	82.5	95
Daubechies 8	80	94.5
Symmlet 2	81.5	94
Coiflet 1	84.5	95.5
Biorthogonale 1.1	80.5	93.5
Reverse Biorth. 1.1	82.5	93.5
Mean	**81.5**	**92.83**

We notice that achieved accuracy rates are better when sub-block is 2*2 sized, which is intuitively expected.

However, in spite of realizing higher accuracy rates with 2*2 sub-blocks, the consumed computational time has increased.

B. Experiment 2

Considering 200 training images and 200 testing ones, cumulative match scores of the proposed method was evaluated for ranks up to 10. The Daubechies 6 wavelet was set as mother one. Sub-blocks were 4*4 sized.

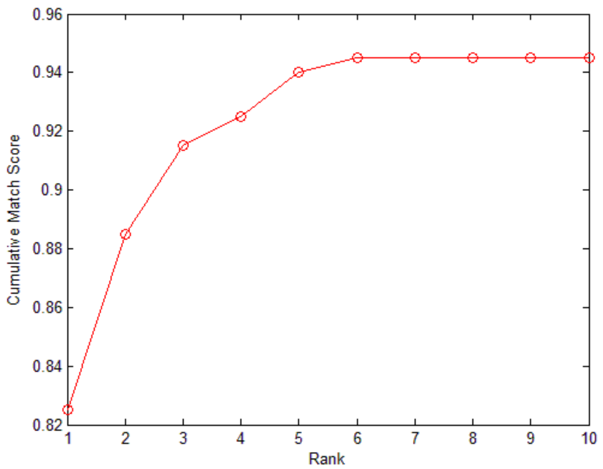

Fig. 1 Proposed method performances on FERET database

We note that the achieved accuracy recognition rate through the proposed approach is up to 82%.

C. Experiment 3

In order to explore the problem of determining the appropriate kNN classifier *k* value, we measured its top 1 recognition rate while varying *k* parameter value. Measures were performed on 850 gallery and probe images with sub-blocks 4*4 sized and Daubechies 6 mother wavelet. The extracted results are shown below:

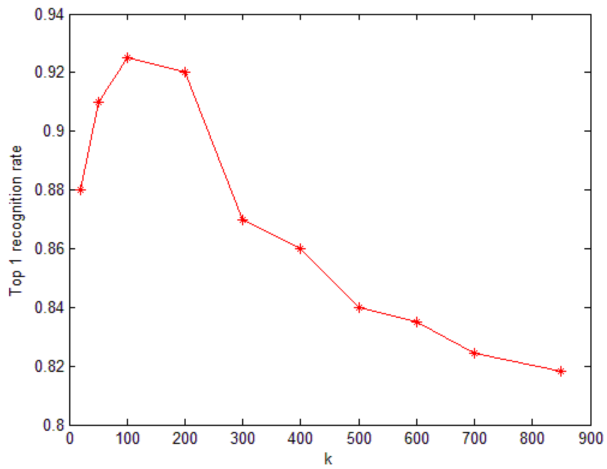

Fig. 2 Proposed method top 1 recognition rates while varying kNN classifier *k* value

Belonging to Figure 2, we can deduce that the suitable kNN classifier k parameter value is approximately in the order of 10%-20% of the experimented gallery size.

Moreover, top 1 recognition rate has decreased while gallery size increased. This is due to the increasing number of classes (faces) during training process since we are dealing with single sample face problem.

V. CONCLUSION

In this paper, we set up a single sample face recognition system based on discrete wavelets transform, self-organizing maps neural network and soft k-nearest neighbor ensemble decision. The obtained results proved this system efficiency thanks to promising classification accuracy rates while using only wavelet approximation channel.

Regarding to [13], our approach allowed achieving decent results whilst reducing faces scale thanks to descrete wavelets transform downsampling property. Comparing achieved results with [13] applied to thumbnails might be an interesting issue worthy to be examined.

Moreover, as future work, we intend to investigate wavelets properties, namely: wavelets transform decomposition level and coefficients set to be kept since this issue has been rarely treated in the literature. Moreover, for the non considered details sub-images, local statistical features (such as Entropy, Energy, …) could be carried out in order to improve features vectors without increasing computational time and loosing information however small.

REFERENCES

[1] P. N. Belhumeur, J. P. Hespanha, and D. J Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19(7), 711-720, 1997.

[2] J. A. Lee, and M. Verleysen, “Self-organizing maps with recursive neighborhood adaptation”, Neural Networks, 15(8), 993-1003, 2002.

[3] T. Huntsberger, J. Rose, and S. Ramaka, “Fuzzy-Face: a hybrid wavelet/fuzzy self-organizing feature map system for face processing”, Journal of Biological Systems, 6(03), 281-298, 1998.

[4] A. M. Martinez, “Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class,” IEEE Trans. Pattern Analysis and Machine Intelligence 25(6): 748-763, 2002.

[5] D. Beymer, and T. Poggio, "Face Recognition from One Example View," Science 272(5250), 1996.

[6] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, ”Face Recognition using kernel direct discriminant analysis algorithms.” IEEE Trans. on Neural Networks, 14(1), 117-126, 2003.

[7] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection”, IEEE Trans. on Pattern Analysis and Machine Intelligence 19(7), 711-720, 1997.

[8] B. Moghaddam, T. Jebara., A. Pentland, “Probabilistic visual learning for object detection,” Proc Int Conf Computer Vision, pp. 786–793, 1995.

[9] C. Liu and H. Wechsler: "Evolutionary Pursuit and Its Application to Face Recognition", IEEE Trans. Pattern Analysis and Machine Intelligence , vol. 22, no. 6, pp. 570-582, 2000.

[10] X. Tan, S. Chen, Z.H. Zhou, F.Y. Zhang: "Robust face recognition from a single training image per person with kernel-based som-face." , ISSN (1), pp.858-863, 2004.

[11] X. Tan, S. Chen, Z.H. Zhou, F.Y. Zhang: "Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble", IEEE Transactions on Neural Networks, 16(4), 875-886, 2005.

[12] M. Fleming, and G. Cottrell, “Categorization of faces using unsupervised feature extraction,” In Proc. IEEE IJCNN International Joint Conference on Neural Networks, 1990.

[13] J. Kittler, M. Hatef, R. Duin, and J. Matas. “On Combining classifiers”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226-239, 1998.

[14] T. Kohonen, “Self-Organizing Map,“2nd edition, Berlin: Springer-Verlag,1997.