International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:2, No:9, 2008

# Face Detection using Gabor Wavelets and Neural Networks

Hossein Sahoolizadeh, Davood Sarikhanimoghadam, and Hamid Dehghani

***Abstract*—**This paper proposes new hybrid approaches for face recognition. Gabor wavelets representation of face images is an effective approach for both facial action recognition and face identification. Perform dimensionality reduction and linear discriminate analysis on the down sampled Gabor wavelet faces can increase the discriminate ability. Nearest feature space is extended to various similarity measures. In our experiments, proposed Gabor wavelet faces combined with extended neural net feature space classifier shows very good performance, which can achieve 93 % maximum correct recognition rate on ORL data set without any pre-processing step.

***Keywords*—**Face detection, Neural Networks, Multi-layer Perceptron, Gabor wavelets.

## I. Introduction

OVER last 15 years incremental activities have been seen in tackling problems such as recognizing faces in various environments [1]. Turk and Pentland [2] demonstrated that Principal Component Analysis (PCA) representation could be used to implement a very efficient and successful face recognition system. Face recognition system using Fisher's Linear Discriminate Analysis (FLD) [3] as the classifier has also been very successful. In [4], C. Liu and Wechsler introduced a Gabor-Fisher Classifier (GFC) method, which couples Gabor Wavelets, PCA and Enhanced Fisher Discriminate Model (EFM) together. The GFC method has shown significant improvement of Gabor features in face recognition. The simple neural net classifier is widely employed for face recognition task. Stan Li et al. [5] proposed a kind of Nearest Feature Line (NFL) classifier which has shown to achieve lower classification error than the NN classifier. Jen-Tzung Chien et al. [6] have proposed a Nearest Feature Space (NFS) classifier for robust decision in presence of wide facial variations, which outperformed the NFL. In our experiments, conventional NN classifier is used as base-line method in comparison with the NFS classifier. In this paper, Gabor wavelet faces are extracted by down sampled Gabor Wavelets transform on the face images. The Gabor wavelet faces after performing PCA and FLD are employed for face identification. In [6], Jen-Tzung Chien only argued L1 norm, but other similarity measure methods such as L2 norm and angle distance are also very useful for pattern classification. Our proposed method of Gabor wavelet faces using extended NFS classifier achieves 93% correct face recognition accuracy for ORL data set [7].

F. Hossein Sahoolizadeh is with the Young Researchers club of Arak, Islamic Azad University, Arak Branch, Iran (e-mail: hosein_sahooli@yahoo.com).

S. Davood Sarikhanimoghadam is with the Young Researchers club of Arak, Islamic Azad University, Arak Branch, Iran (e-mail: davood.sarikhani2008@yahoo.com).

T. Hamid Dehghani is with the Electrical Engineering Department, Maleke Ashtar University of Technology, Tehran, Iran (e-mail: hamid_deh@yahoo.com).

## II. Gabor Wavelets Ransform Features Extraction Review Stage

Gabor wavelets are widely used in image analysis and computer vision [8, 9]. The Gabor wavelets transform provides an effective way to analyze images and has been elaborated as a frame for understanding the orientation and spatial frequency selective properties of simple cortical neurons. They seem to be a good approximation to the sensitivity profiles of neurons found in visual cortex of higher vertebrates. The important advantages are infinite smoothness and exponential decay in frequency. Let $\vec{I}(z)$ be the gray level distribution of the input image, Gabor wavelets transform on $\vec{I}(z)$ can be written as a convolution of $\vec{I}(z)$ with a family of kernels $\psi_k$ :

$$\vec{O}_k(z) = I(z) * \psi_k(z) \quad (1)$$

Where * denotes the convolution operator, and $\vec{O}_k(z)$ is the convolution result at k. The Gabor wavelets (kernels) take the form of a plane wave restricted by a Gaussian envelope function [9]:

$$\psi_k(z) = \frac{\|k\|}{\sigma^2} e^{-\|k\|^2\|z\|^2/2\sigma^2}[e^{ikx} - e^{-\sigma^2/2}] \quad (2)$$

Where k determines the wavelength and orientation of the kernel $\psi_k(z)$ in image coordinates. The first term in bracket is oscillation part, and the second is dc part. The k is defined as $k(\mu, \upsilon) = k_\upsilon e^{i\varphi_\mu}$ , where μ and υ define the orientation and scale of the Gabor kernels, $k_\nu = k_{max} ax/f^\nu$ and $\varphi_\mu = \pi\mu/8$ . $k_{max}$ is the maximum frequency, and f is the spacing factor between kernels in the frequency domain.

The $\psi_k(z)$ form a family that is self-similar under the application of the group of translations, rotations, and resizes.

The effect of the dc term becomes negligible when the parameter, which determines the ratio of the Gaussian window width to wavelength, has sufficiently large value.

Lades and et al. [9] suggest that good result can be obtained by using Gabor wavelets of five different scales, $V \in \{0, \ldots, 4\}$ and eight orientations, $\mu \in \{0, \ldots, 7\}$ ,$\sigma = 2\pi$, $k_{max} = \pi/2$ and $f = \sqrt{2}$ . We call it Gabor wavelet faces. The downsampled Gabor wavelet transform results by the factor of 64 formed a feature vector.

In this method, like many other compression methods, the input image was divided into blocks with smaller size. Then each block was entered into input layer to be compressed in the middle layer and reconstruct that block in the output layer. Implementation of this method which was done in a computer in parallel, showed its poor performance even for the trained images. To mention the works to improve this structure, we can point to hierarchal structure which two more hidden layers were used in it to consider the correlation between blocks in two output hidden layers and also between pixels in the internal layer. But this structure still was not satisfactory. Adaptive methods are other alternatives to compress the images which use several networks with different compression ratios according to their details and complexities. In [6], different methods to train such structures have been proposed. In one of these algorithms [2], blocks of the image will be separated into four group based on a criteria called “activity”. Four different networks with different compression ratios will be built and Images will be compressed based on these networks. It is shown that this structure has a good result and it is modified in this way that blocks with high activities separated in four other groups, so totally we have six different networks. This kind of separation will improve the visual quality of the reconstructed image. As another results obtained in these networks, which was introduced in [5], is using nine different networks to compress Images by classifying them with respect to block directions and the amount of edges contained in each block. In this structure image fed into the network, after each pixel value was subtracted from the included block mean. In this article, we have used entropy criterion as an estimation of information available in an image and develop a structure based on this criterion. The other method to be used is to overlap the blocks of the image which is studied hereafter and is compared with conventional JPEG compression technique.

This article consists of 3 other sections. In section II, using perceptron neural networks in image compression was studied, both in a simple direct form and adaptive one. In section III, implementation result and comparison between the proposed algorithms with conventional compression technique are derived. And finally section IV, conclusion and summary of present research work are given.

## III. Neural Network

A neural network is an information-processing system that has been developed as generalizations of mathematical models matching human cognition. They are composed of a large number of highly-interconnected processing units (neurons) that work together to perform a specific task. According to Haykin [8], a neural network is a massively parallel-distributed processor that has a natural prosperity for storing experimental knowledge. It resembles the brain in two respects:

- Knowledge is acquired by the network through a learning process;
- Inter-connected connection strengths known as synaptic weights are used to store the knowledge;
- Each neuron has an internal state called its threshold or activation function (or transfer function) used for classifying vectors. Neural classification generally comprises of four steps:
- Pre-processing, e.g., atmospheric correction, noise suppression, band rationing, Principal Component Analysis, etc;
- Training - selection of the particular features which best describe the pattern;
- Decision - choice of suitable method for comparing the image patterns with the target patterns;
- Assessing the accuracy of the classification.

## IV. Implementation Results

In this section, the implementation results of the detection algorithms based on neural algorithm are studied. Experiments are conducted by ORL dataset for our proposed Gabor waveletfaces with extended NFS classifier. The ORL dataset consists of 400 frontal faces: 10 tightly cropped images of 40 individuals with variations in pose, illumination, facial expression and accessories. The size of each image is 92 × 112 pixels, with 256 grey levels per pixel. After the extraction of waveletfaces as the features space, Multilayer Perceptron Neural networks is used to as a classifier and the above described database is used for traing.

Our algorithm can detect between 77.9% and 90.3% of faces in a set of 130 test images, with an acceptable number of false detections. Depending on the application, the system can be made more or less conservative by varying the arbitration heuristics or thresholds used. The system has been tested on a wide variety of images, with many faces and unconstrained backgrounds. A fast version of the system can process a 320x240 pixel image in 2 to 4 seconds on a 2 GHz pentium4 with 512 Megabyte Ram.

There are a number of directions for future work. The main limitation of the current system is that it only detects upright faces looking at the camera. Separate versions of the system could be trained for each head orientation, and the results could be combined using arbitration methods similar to those presented here. Preliminary work in this area indicates that detecting profiles views of faces is more difficult than detecting frontal views, because they have fewer stable features and because the input window will contain more background pixels. We have also applied the same algorithm for the detection of car tires and human eyes, although more work is needed.

Even within the domain of detecting frontal views of faces, more work remains. When an image sequence is available, temporal coherence can focus attention on particular portions of the images. As a face moves about, its location in one frame is a strong predictor of its location in next frame. Standard tracking methods, as well as expectation-based methods [2], can be applied to focus the detector's attention. Other methods of improving system performance include obtaining more positive examples for training, or applying more sophisticated image preprocessing and normalization techniques. One application of this work is in the area of media technology. Every year, improved technology provides cheaper and more efficient ways of storing and retrieving visual information.

However, automatic high-level classification of the information content is very limited; this is a bottleneck that prevents media technology from reaching its full potential. Systems utilizing the detector described above allow a user to make requests of the form "Show me the people who appear in this video" or "Which images on the World Wide Web contain faces?" [6] and to have their queries answered automatically. Some sample images are shown in Fig. 1 and Fig. 2.

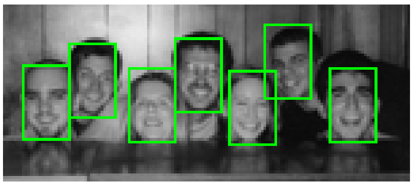

Fig. 1 Some sample images to test the proposed structure

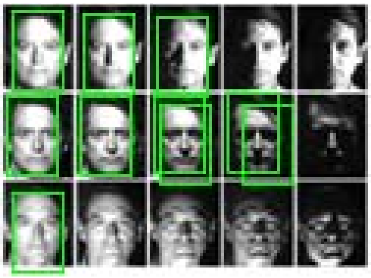

Fig. 2 Some sample images failed to work with the proposed algorithm

## V. Conclusion

In this paper, a new hybrid method of Gabor waveletfaces using extended NFS classifier has been proposed and tested with very good result in well-known standard data set.

In comparison with other methods, all Gabor waveletfaces based method is better than the base-line method. Down sampled Gabor wavelets transform of face images as features for face recognition in subspace approach is superior to pixel value. In our experiments, proposed Gabor wavelet faces combined with extended neural net feature space classifier shows very good performance, which can achieve 93 % maximum correct recognition rate on ORL data set without any pre-processing step.

## References

[1] R. W. Chellappa, C.L.; Sirohey, S., "Human and machine recognition of faces: a survey," Proceedings of the IEEE, vol. 83, pp. 705-741, 1995.

[2] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," presented at Proceedings CVPR '91., 1991. [3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," PAMI, IEEE Trans. on, vol. 19, 1997.

[3] A. M. Martinez and A. C. Kak, "PCA versus LDA," PAMI, IEEE Trans. on, vol. 23, pp. 228-233, 2001.

[4] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," Image Processing, IEEE Trans. on, vol. 11, pp. 467-476, 2002.

[5] S. Z. Li and J. Lu, "Face recognition using the nearest feature line method," Neural Networks, IEEE Trans. on, vol. 10, pp. 439-443, 1999.

[6] J.-T. Chien and C.-C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition," PAMI, IEEE Trans. on, vol. 24, pp. 1644-1649, 2002.

[7] ORL web site: www.uk.research.att.com/facedatabase.html.

[8] T. S. Lee, "Image representation using 2D Gabor wavelets," PAMI, IEEE Trans. on, vol. 18, pp. 959-971, 1996.

[9] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," Computers, IEEE Trans. on, vol. 42, pp. 300-311, 1993.