A SOM-WAVELET NETWORKS FOR FACE IDENTIFICATION

Yang Zhi *Gu Ming*

School of Software, Tsinghua University, Beijing , P.R.China 100084
yangz02@mails.tsinghua.edu.cn
guming@mail.tsinghua.edu.cn

ABSTRACT

This paper describes a novel SOM-Wavelet Networks method for face recognition. We employed a SOM algorithm, which is based on the structure of a biological model, to extract shape feature of face. After the unsupervised learning, each face image will produce a shape-based vector named representative face. A Wavelet Network is applied to face identification to collect global information from a face image. Then we proposed a new approach to compute the similarity between two faces on both the global means and topological means . The experimental results are compared with other effective face identification methods and our proposed method showed a good performance.

1. INTRODUCTION

In the past two decades, Face Identification(FI) technology was advanced by many researchers. The improvement of researching on representation model of face is an important factor that promoted the performance of FI systems. By now, the most successful representation models can be mainly divided into two classes: one is the linear models such as eigenface, fisherface[1]; the other is the nonlinear models.

Self-organizing map (SOM) is one of successful nonlinear models, and the explanation of this success can be found in the fact that SOMs perform simultaneously the combination of two subtasks: vector quantization and topographic representation[2]. Recently, some effective improvement was achieved on face processing system based on SOM. T. Huntsberger, J.Rose and S. Ramaka developed a face processing system called Fuzzy-Face that combine wavelet preprocessing of input with a fuzzy SOM[3]; J. L. Alba Castro, A. Pujol and J. Villanueva applied SOM+PCA approach for FI which is based on shape and texture information[4]; Tan and Chen presented kernel SOM face to deal with both the nonlinear problem and the small sample problem in FI[5]; in biological simulation aspect, J.A. Bednar and R. Miikkulainen proposed a HLISSOM model to explain the face recognition preference of the infants[6].

0-7803-9332-5/05/$20.00 ©2005 IEEE

We are not only interested in the topological information but also concerned about global approximation of face. Wavelet Networks (WN) is a kind of universal approximator over the face and achieves faster convergence over radial basis function networks (RBFN), and it is capable of dealing with the so-called “curse of dimensionality”[7]. In WN, the radial basis functions of RBF-networks are replaced by wavelets. WN inherit the properties of wavelet decomposition and mention especially their universal approximation property, the availability of convergence rates and the explicit link between the network coefficients and the wavelet transform, and during the training phase, the network weights as well as the degrees of freedom (position, scale, orientation) of the wavelet functions are optimized [8]. Recently, wavelet network was applied in many research fields, such as Face Alignment [9], Object Representation[10], Pose Estimation [11], etc.

The work of us is to extract both shape information and global wavelet feature, and to classify them by reasonable means. To achieve this goal, we proposed a new face recognition approach based on SOM algorithm and Wavelet Networks.

The outline of this paper is as follows: section 2 will provide some background of the SOM algorithm, and an improved SOM method to get the representative face will be presented. A Wavelet Network will be applied in Face Identification in section3. In the final section, some experiment results and conclusions will be shown.

2. SHAPE PROCESSING BY SHAPE-SELF ORGANIZATION MAP

2.1 Related Background About SOM

There are two major vector quantization algorithms: the Winner Takes All(WTA) and the Winner Takes Most (WTM). WTA means that only one neuron called Best Matching Unit(BMU) is adapted when an input pattern stimulate the networks while in WTM there are more than one neurons adapted. Unlike other WTM algorithm, the SOM adapts neurons both on the pattern space and on the topological space. That means we should take the grid distance into account in grid neural networks. Translating all these ideas in formulas, the map can be defined by:

(1) a matrix $\vec{W}$, of which row $\vec{w}_r$ gives the weights or coordinates of neuron *r* in the pattern space;

(2) a function $d(q,r)$,measuring the grid distance between neurons *q* and *r* in the grid space.

Notice that the distance $d(q,r)$ may be implicitly determined either by a true position in a grid or by another more general mathematical structure like a weighted graph with neurons as nodes. This definition leaves enough freedom regarding the shape of the map or the structure of the neighborhood[8]. In this paper, we utilize the HLISSOM model as the basic topological structure(that is why we call our algorithm Shaped-LISSOM) , which has rectangular grid and each node has 8 connected neighbors in a 2D plane. Having $\vec{W}$ and $d(q,r)$, the map may begin its learning process. Assuming that:

(1) vectorial pattern $\vec{x}_i$ stimulates the map at learning time t;

(2) index $* = \arg\min_r \left\| \vec{w}_r^t - \vec{x}_i \right\|$ points to the best-matching unit,

then all neurons are adapted according to following rule: $\vec{w}_r^{t+1} = \vec{w}_r^t + \Delta\vec{w}_r^t = \vec{w}_r^t + \gamma_r^t(\vec{x}_i - \vec{w}_r^t)$ where γ_r^t is a decreasing learning rate for neuron *r* between 0 and 1 .

It is no doubt that shape information has important influence on effect of face identification system. SOM algorithm is both adequate to cluster the face space according to the shape images in order to optimize the modular eigenspace, and to keep a topological relationship within the clusters[4]. Then after training the whole image database, we will get a "shaped-based" vector set.

2.2 Learning Rules of S-LISSOM

Unlike other ANN training algorithm, SOM learning rules operate "on-line", so in cluster *j* the cluster representative B_k^j must be updated iteratively. Let $\eta_{a,b}(k)$ be the *k*-th times response of neuron(*a,b*) on B_k^j, we have:

$$\eta_{ab}(k+1) = \eta_{ab}(k) + \gamma_E^k \sum_{mn} E_{ab,mn}^k \eta_{mn}(k) - \gamma_I^k \sum_{mn} I_{ab,mn}^k(k)\eta_{mn}(k) \qquad (1)$$

where $E_{ab,mn}$ is the excitatory lateral connection weight on the connection from neuron (*m,n*) to neuron (*a,b*), $I_{ab,mn}$ is the inhibitory connection weight.

Both $E_{ab,mn}$ and $I_{ab,mn}$ can be expressed in the way of neighborhood factor or neighborhood kernel. In order to observe the self-organization property, the neighborhood factor must be a decreasing function of the grid distance $d(*,r)$ (in this paper, we utilize the Euclidean distance) between the BMU and the neuron being adapted. For example, it can be the 'Bubble' function or Gaussian kernel[2]. Here we use the Gaussian kernel v_r^k to describe the *k*-times neighborhood factor at neron r as follows:

$$v_r^k = \exp(-0.5(d(*,r)/\lambda^k)) \qquad (2)$$

where λ is a constant value between 0 and 1 in order to ensure v_r^k decrease as time goes by.

The γ_E^k and γ_I^k are the value of global decreasing learning rate. Here we use α^k to describe the *k*-times learning rate as follows:

$$\alpha^k = \exp(-k/\mu) \qquad (3)$$

where μ is the decreasing rate of above function.

2.3 Construction of S-LISSOM Representative Face

After the SOM process, we will get a vector to represent a face, and this face may be called the representative face or SOM-face[5]. Each person may have several representative faces according to different pose of his head or different face expression. So let the B_m^j be the *m*-th sample representative face of individual *j,* then sample set of this person can be defined as $B^j = \{B_m^j\}$.

4. FACE RECOGNITION

4.1 An Introduction To Wavelet Networks

A Wavelet Networks can be defined as:

Let ψ_{ni} $i = 1,....,N$ be a set of wavelets, *f* a DC-free image and w_i and n_i chosen according to the energy function(4). The two vector $W = (w_1,..., w_N)^T$ and $\Psi = (\psi_{n1},...,\psi_{nN})^T$ then define the *wavelet network* (Ψ, W) for image *f*. In our experiment, a real-valued Gabor wavelet function was used:

$$\begin{aligned}\psi_n(x,y) = \exp(-\frac{1}{2}&[s_x((x-c_x)\cos\theta-(y-c_y)\sin\theta)]^2 \\ &+[s_y((x-c_x)\sin\theta+(y-c_y)\cos\theta)]^2) \\ &\cdot\sin(s_x((x-c_x)\cos\theta-(y-c_y)\sin\theta))\end{aligned}$$

where parameter vectors $n = (c_x, c_y, \theta, s_x, s_y)$. The c_x, c_y defines the translation of the wavelet, s_x, s_y defines the dilation and θ defines the orientation of the wavelet. The parameters vector n (translation, orientation and dilation) of the wavelets maybe chosen arbitrarily at this point[8]. According to wavelet theory, any function $f \in L^2(R^2)$ can be losslessly represented by their continuous wavelet transform and thus, with arbitrary precision, by a wavelet network. We therefore interpret the image *f* to be a function of the space $f \in L^2(R^2)$ and assume further, without loss of generality that *f* is DC-free.

If we want to reconstruct image *I* by Gabor wavelet ψ_i and let w_i be the corresponding weight , we have

$$\hat{I}(\vec{x}) = \sum_{i=1}^{N} w_i \psi_i(\vec{x})$$

In order to find the WN for image f , we minimize the energy function:

$$E = \min_{n_i, w_i for \quad all \quad i} \left\| f - \sum_{i=1}^{N} w_i \psi_{ni} \right\| \tag{4}$$

If we want to compute the distance between two vectors of wavelet coefficients v and w, for example , we can compute the Euclidean distance in the (image) subspace $< \Psi >$ between the v and w:

$$\left\| \sum_{i=1}^{N} v_i \psi_{ni} - \sum_{i=1}^{N} w_i \psi_{ni} \right\| \tag{5}$$

algebraic transformations lead to

$$\|v - w\|_{\Psi} := \left[\sum_{i,j} (v_i - w_i)(v_j - w_j) \langle \psi_i, \psi_j \rangle \right]^{1/2}$$

$$= (v - w)^T (\Psi)(v - w) \tag{6}$$

Where $\|\bullet\|_{\Psi}$ computes the Euclidean distance between the two appropriate points in $< \Psi >$ and thus considers the different parameters of the wavelets. For orthogonal wavelets, the matrix $(\Psi)_{i,j} = \langle \psi_{ni}, \psi_{nj} \rangle$ is the unity matrix and no weighting is needed[8].

It is impossible to calculate w_i by a simple projection of the Gabor wavelet ψ_j onto the image[11] (as it is being done for orthogonal wavelets).Instead we have to consider the family of *dual* wavelets $\tilde{\Psi} = \{\tilde{\psi}_1, ... \tilde{\psi}_N\}$. The wavelet $\tilde{\psi}_j$ is the *dual* wavelet to the wavelet ψ_j if :

$$\langle \psi_i, \tilde{\psi}_j \rangle = \int \psi_i(\vec{x}) \tilde{\psi}_k(\vec{x}) d^2\vec{x} = \delta_{i,j} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq k \end{cases} \tag{7}$$

$\tilde{\psi}_i$ can be computed as follows:

$$\Psi_{i,j} = \langle \psi_i, \psi_j \rangle \quad \text{where} \quad \tilde{\psi}_i = \sum_j (\Psi_{i,j})^{-1} \psi_j \tag{8}$$

This function's proof can be found in [8].Then we can calculate w_i directly:

$$w_i = \langle I, \tilde{\psi}_i \rangle \tag{9}$$

4.2 Face Recognition Based On WN and S-LISSOM

In order to apply face recognition algorithm, we first extract $W_m^j = (w_1, ..., w_N)_m^j$ and $\Psi_m = (\psi_1, ..., \psi_N)_m^j$ of each training image *m* of each person ***j***. The *Wavelet Networks* $(\Psi, W)^j = ((\Psi_1, ..., \Psi_M), (W_1, ..., W_M))^j$ of each person is regarded as his or her prototype sample and stored with this person's representative face set to form a sample set $G^j = (B^j, (\Psi, W)^j)$.

When a new face image is input, we extract its coefficient vectors $V = (v_1, ..., v_N)^T$ according to eq. (9), and then optimize them based on eq.(4). After getting V_n we will compare it with W_m of each person , the difference is defined as:

$$D_m(V, W_m) = \|V - W_m\|_{\Psi_m} \tag{10}$$

The Mahalanobis distance had been proved to be effective in object matching applications[12]. Let us denote I be the input image, then compute Mahalanobis distance $M_m(I, B_m^j)$ of the corresponding the *m*-th representative face of person j as the similarity-weight u_m^j : $u_m^j = M_m(I, B_m^j)$. We have many methods to compute the difference between two faces, among them the simplest method is to calculate the sum of all D_m. As mentioned in[13], we can calculate the similarity between two faces on an average over the difference among pairs of corresponding face pose. For an input image I and a stored individual sample set G^j, the similarity between them is defined as :

$$S(I, G^J) = -\frac{1}{M} \sum_m u_m^j D_m(V, W_m) \tag{11}$$

5. EXPERIMENTS

We utilized the ORL face image database of Cambridge University to check the validity of our algorithm. 400 face images from 40 individuals in different states from the ORL have been used. The sample images vary in position, rotation, scale and expression. In this database each person has changed his face expression in each of 10 samples (open/close eye, smiling/not smiling). For some individuals, the images were taken at different times, varying facial details (glasses/no glasses). Fig1 shows some results of representative face on ORL database.

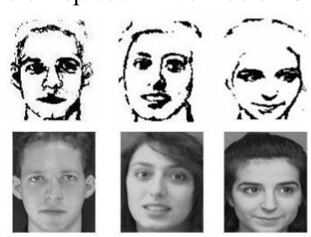

Fig1. Top row are the representative faces of images in ORL database. The bottom row are corresponding sample images of ORL database.

The first experiment was to test correct identification rate among several SOM-based face identification methods. We increased the testing gallery from 31 to 40 individuals in the experiment, and chose randomly among them to test recognition rate, each new individual with 10 images. After being added another new individual, the test was executed by randomly choosing 120 images from the current testing face database. From the Fig2 we can see our proposed method's recognition rate always higher than 90% and has an obviously better accuracy than the other two SOM based methods when the number of testing images in database increase.

Another experiment is to test the performance of the proposed method with reference to other template-based approaches on only one training sample. We employed

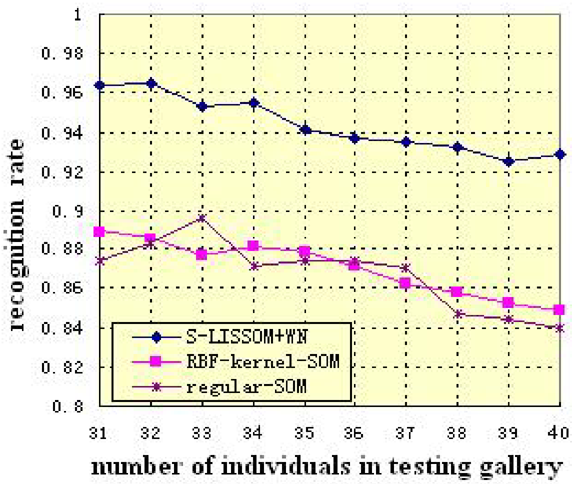

Fig2 The Recognition rate curves of 3 SOM-based Face Identification methods.

HMM, PCA and fisher-face algorithm, because they are proved to be effective in Face Identification. We tested every algorithm by randomly choosing 50 from 400 face images of ORL database. The testing results are shown in Table1. From the table, it should admitted that the HMM has the best performance. Our proposed method also gain better performance compared with rest 4 algorithms.

Table1 Comparison recognition rate of six approaches(%)

Method	Recognition rate
PCA	81.6
Fisher-Face	82.5
HMM	91.6
RBF-SOM	84.1
Regular-SOM	83.3
S-LISSOM+WN	90.8

6. CONCLUSION AND ACKNOWLEGEMENT

In this paper, we present a novel Face Identification method which not only consider the global representation of face but also take topological information into account. Based on structure of a biological model, a SOM algorithm is applied to extract the shape information of a face image. Every sample was trained under a unsupervised process and produce a corresponding vector named representative face. The Wavelet Networks was used for computing the universal feature of a face image. We proposed a new approach for computing similarity between two faces. The algorithm was tested on the ORL database and proved to be useful for face identification. Future work will be carried on locating some detail feature of a face to get a more accurate representation..

The author thank for the financial support of the National 973 Plan of China, under project number: 2004CB719400.

7. REFERENCES

[1] Belhumeur, P. Hespanha, J. Kriegman. " Eigenfaces vs. Fisherfaces: recognition using class specific linear projection". *IEEE Trans. Pattern Analysis and Machine Intelligence* 19, pp.711-720, 1997.

[2] John A. Lee, Michel Verleysen, "Self-organizing maps with recursive neighborhood adaptation", *Neural Networks,* Vol.15, pp.993–1003, 2002.

[3] Terry Huntsberger, John Rose, Shashidhar Ramaka, "Fuzzy – Face: A Hybrid Wavelet/Fuzzy Self-Organizing Feature Map System for Face Processing", *Journal of Biological Systems*, Vol. 6, No. 3, pp.281-298,1998.

[4] J. L. Alba, Albert Pujol and JJ Villanueva, "Novel SOM-PCA network for face identification" *Proceedings of SPIE Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation IV,* Vol.4479, pp.186-194,2001.

[5] 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.

[6] James. A. Bednar and Risto Miikkulainen, "Learning Innate Face Preference Report", *AI-01-291,* November 2001.

[7] S.T. Li and S.C. Chen, "Function Approximation using Robust Wavelet Neural Networks", *ICTAI02,*pp. 483 – 488, Washington DC **,**November 04 - 06, 2002 **.**

[8] V.Kruger , G.Sommer, "Wavelet networks for face processing", *JOSA-A(19)*,No. 6, pp. 1112-1119,June 2002.

[9] Changbo Hu, Rogerio Feris, Matthew Turk , " Active Wavelet Networks for Face Alignment ", *British Machine Vision Conference*, East Eaglia, UK, 2003.

[10] Volker.Kr¨uger, "Gabor wavelet networks for object representation", *Technical Report CS-TR-4245, University of Maryland,* CFAR, May 2001.

[11] Volker Krüger, Gerald Sommer, "Gabor wavelet networks for efficient head pose estimation". *Image Vision Compute. 20,*(9-10)ppt: 665-672, 2002.

[12] Josef Sivic and Andrew Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos", *ICCV03,* pp.1470-1477,2003.

[13] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, Christoph von der Malsburg , "Face Recognition by Elastic Bunch Graph Matching", *IEEE trans on pattern analysis and machine intelligence*, vol.19, NO.7, July 1997.