Gabor Wavelet Associative Memory for Face Recognition

Haihong Zhang, Bailing Zhang, Weimin Huang, and Qi Tian

***Abstract*—This letter describes a high-performance face recognition system by combining two recently proposed neural network models, namely Gabor wavelet network (GWN) and kernel associative memory (KAM), into a unified structure called Gabor wavelet associative memory (GWAM). GWAM has superior representation capability inherited from GWN and consequently demonstrates a much better recognition performance than KAM. Extensive experiments have been conducted to evaluate a GWAM-based recognition scheme using three popular face databases, i.e., FERET database, Olivetti-Oracle Research Lab (ORL) database and AR face database. The experimental results consistently show our scheme's superiority and demonstrate its very high-performance comparing favorably to some recent face recognition methods, achieving 99.3% and 100% accuracy, respectively, on the former two databases, exhibiting very robust performance on the last database against varying illumination conditions.**

***Index Terms*—Face recognition, Gabor wavelet networks (GWNs), kernel associative memory (KAM).**

I. Introduction

Face recognition has been a very active research topic in pattern recognition and computer vision communities in recent years. It has a wide range of applications such as identity authentication, access control, surveillance and content-based indexing [1]. Despite remarkable progresses so far, the general task of face recognition remains a challenging problem due to complex patterns caused by various variations in illumination conditions, facial expressions, and poses.

The first important issue in face recognition is to select efficient representations of face images. Psychophysical studies have suggested that the visual perception tasks such as similarity judgement tend to operate on a low-dimensional representation of the sensory data [2]. Many representation approaches for face recognition have been suggested such as simple low-resolution "thumbnail" images [3] and geometric features [4]. Another methodology widely used is by holistic/local image decomposition with some special 2-D signals (so-called image kernels) such as Eigenfaces [5].

In order to accurately capture local features in face images, a spatial-frequency analysis is often desirable. Wavelet analysis is particularly useful for this purpose since it has a good characteristics of space-frequency localization. In computer vision, the multiresolution scheme in wavelet analysis has been justified by psychovisual research. In particular, among various wavelet bases Gabor functions provide a favorable tradeoff between spatial resolution and frequency resolution [6]. And there is a strong biological relevance of processing images by Gabor wavelets as they have similar shapes to the receptive fields of simple cells in the primary visual cortex (V1) [7].

Krueger has proposed for face representation a Gabor wavelet network (GWN) [8] which tries to represent an image by a set of weighted Gabor wavelets. He has also successfully applied GWNs to face recognition and other domains. However, since GWN was designed to represent a single static image, it may not effectively handle variable features that are important for recognition. Hence, a new GWN model is needed to represent a subject with variable appearances.

Manuscript received January 3, 2003; revised April 16, 2004.

H. Zhang, W. Huang, and Q. Tian are with the Institute for Infocomm Research, Singapore 119613 (e-mail: hhzhang@i2r.a-star.edu.sg).

B. Zhang is with the School of Computer Science and Mathematics, Victoria University, Melbourne VGIC3011, Australia (e-mail: bzhang@csm.edu.au).

Digital Object Identifier 10.1109/TNN.2004.841811

Another important issue in face recognition is about how to classify a query face image using the selected representation. Many discrimination techniques have been proposed, such as Fisher linear discriminant (FLD) or neural networks classifiers. Examples of applying neural networks in face recognition include: 1) the convolutional neural network (CNN) [9] and 2) the probabilistic decision-based neural network [10]. Another kind of neural network that has been actively researched in face recognition is associative memory (AM) [11]. In brief, AM-based classification learns how to do recognition by categorizing positive examples of a subject. There has been a long history of AM research and the continuous interest is due to a number of attractive features of these networks, such as content addressable memory, collective computation capabilities etc.

In a recent paper, Zhang *et al.* proposed a kernel associative memory (KAM) which introduces kernel methods to AM by nonlinearly mapping the data into some high-dimensional feature space through operating a kernel function with input space [12]. For face recognition, KAM was used with image features from wavelet transform (WT) and yielded good experimental results.

This letter presents a new face recognition scheme called Gabor wavelet associative memory (GWAM), by incorporating advantages of both GWN and KAM. We first extend GWN to a subject dependent GWN model and then combine it with KAM in a unified framework that aims to significantly improve the representation capability of previous models.

Extensive experiments have been carried out to evaluate GWAM-based face recognition system on several publicly available benchmark face databases, including the FERET [14], the Olivetti-Oracle Research Lab (ORL), and the AR [16]face database. Compared with several latest face recognition systems, our scheme shows very high performance.

II. Subject Dependent Gabor Wavelet Networks

Wavelet networks were first introduced in [13] as a combination of feed-forward neural networks and continuous wavelet decomposition. The principle of a wavelet network consists in choosing a set of wavelets as activation functions for the second layer, adaptively according to a specific function f to be represented, such that an approximation $\hat{f}$ would be a linear combination of the wavelet set.

Krueger has extended wavelet networks to GWNs for image representation using Gabor functions [8]. A GWN is optimized for the representation of a single static image. In face recognition, we need to take into consideration the variations of a specific subject. For this we extend the GWN to a subject dependent Gabor wavelet network (SDGWN) as follows.

For a given subject and its image set $\{f_j\}, j = 1 \dots N$, a SDGWN model tries to approximate each image by a common set of Gabor wavelets $\{\psi_i\}$

$$f_j(\mathbf{x}) = \sum_{i=1}^{M} w_{ji}\psi_i. \tag{1}$$

Here ψ_i is a particular 2D Gabor function that can be defined by (following an argument in [8], we use the imaginary part of Gabor function as the wavelet basis ψ)

$$\psi_i(\mathbf{x}) = \frac{k_i^2}{\sigma_i^2} exp\left(\frac{-k_i^2\Theta_i^T(\mathbf{x}-\mathbf{x}_{i0})}{2\sigma_i^2}\right) sin\left(k_i\Theta_i^T(\mathbf{x}-\mathbf{x}_{i0})\right) \tag{2}$$

where $k_i, \Theta_i, \sigma_i, \mathbf{x}_{i0}$ define the frequency, the phase, the spatial bandwidth and the centre of a Gabor kernel, respectively. These configuration parameters can be obtained through optimizing the objective functional of total approximation error

$$E(W_1, \ldots, W_N) = \sum_{j=1}^{N} \left\| f_j - \sum_{i=1}^{M} w_{ji}\psi_i \right\|^2 \tag{3}$$

where $W_j = \{w_{j1} \ldots w_{jM}\}$ denotes a weight vector specific for an individual image f_j. As the Gabor functions $\{\psi_i\}$ are fixed, W_j can be considered as the projection of f_j into the subspace spaned by $\{\psi_i\}$. Because Gabor wavelets are nonorthogonal bases, linear projections of a new pattern on them do not produce the correct coefficients W. Instead, dual Gabor wavelets $\{\tilde{\psi}_i\}$ are adopted to obtain W for a new pattern f

$$W = \tilde{\Psi} f \ with \ \tilde{\Psi} = \Psi^+ = \left(\Psi^T\Psi\right)^{-1}\Psi^T. \tag{4}$$

Obviously, the configuration of the Gabor wavelets is crucial to representation, and learning it through the optimization of (3) plays an important role in the system. In this work, we employs nonlinear optimization toolbox provided with Matlab to attain the optimsation.

Fig. 1 shows a simple comparative example regarding image representations. Image samples are from FERET database [14]. In this test, three samples were selected for training and another image for testing. One SDGWN model was adapted to all the three images while each GWN model to an individual image. The figure successively displays the representation results to the right of the original image. Particularly, the rightmost image is the reconstruction by the SDGWN model, and it can be seen that the SDGWN model can favorably capture facial features and yield better results than the GWNs. More comparisons will be given later.

Fig. 1. Comparative representations for a new face by SDGWN and GWNs.

III. GWAM

The original KAM can be briefly described as follows. First a correlation AM [11] $\hat{\mathbf{x}}_n = W\mathbf{x}_n$ $(n = 1, \cdots, N)$ is rewritten as

$$\hat{\mathbf{x}} = W\mathbf{x} = \sum_{n=1}^{N} \left(\mathbf{x}_n \mathbf{x}_n^T\right) \mathbf{x} = \sum_{n=1}^{N} (\mathbf{x}_n, \mathbf{x})\mathbf{x}_n \tag{5}$$

where $(\mathbf{x}_n, \mathbf{x})$ denotes the dot product between a prototype $\mathbf{x}_n$ and a probe pattern $\mathbf{x}$. Let us substitute the dot product by kernel product with a mapping function $\Phi : k(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}), \mathbf{\Phi}(\mathbf{x})' \rangle$. After generalization, the resulting network termed KAM is given by [12]

$$\hat{\mathbf{x}} = \sum_{n=1}^{N} w_n k(\mathbf{x}_n, \mathbf{x}) = W\mathbf{k}. \tag{6}$$

The weight vectors can be learned through minimizing the following objective functional:

$$J(W) = \|X - W\mathbf{k}\| \tag{7}$$

where X is the matrix consisting of row sample-vectors. A linear optimal solution to this problem is given by $W = XK^+ = X(K^TK)^{-1}K$, where $K = [k^{(1)}, \cdots, k^{(N)}]$ is the matrix of kernel products.

In this letter, we propose a GWAM, by introducing the weight vectors of SDGWN to KAM as feature patterns. As shown in Fig. 2, the model consists of five layers with a KAM model embeded as the central part. The first layer receives input images. The second and the fourth layer correspond to a same set of Gabor wavelets, which are used to encode/decode image patterns to/from the corresponding SDGWN feature space. The third layer serves as the kernel layer of KAM. The final output (i.e., the representation) is presented at the last layer. In this way, GWAM inherits the advantages of both SDGWN and KAM in object representation and in pattern learning.

For face recognition, a modular system based on GWAM is built, in which each GWAM model is associated with a particular subject. In the recognition stage, a face f will be presented to each GWAM model for reconstruction, and the similarity measurement between f and the representation $\hat{f}$ will be taken to determine which class f is from. In other words, we can pick up the network that offers the best representation for the input pattern. In particular, we adopt a widely used similarity measurement given by $\cos(\hat{f}, f) = (\hat{f}^T f)/(\|\hat{f}\| \cdot \|f\|)$.

Fig. 3 illustrates the recognition scheme. The input image on Fig. 3(a) is from the FERET database [14]. It is first encoded by various sets of Gabor wavelets to produce SDGWN representations, as shown in Fig. 3(b) where the right SDGWN model (the top one) tends to yield better results than others. The representative weight vectors $\{w_j\}$ serves as features for kernel associative memories to recall patterns $\{\hat{w}_j\}$, which are then transformed to pixel domain [see column Fig. 3(c)] by the Gabor wavelets. Fig. 3(d) displays typical images selected from the training gallery for each network. Apparently each network tries to reconstruct a pattern similar to its training patterns. It can be seen that the first network (GWAM1) yielded the best representation, thus, offered the correct recognition.

IV. Experiments

We have conducted extensive experiments to test our face recognition scheme and to compare it with other well-known methods on a few publicly available benchmark face databases including the FERET standard facial database (Release2) [14], ORL database [15] and the AR face database [16].

A. Experiments for Comparing SDGWN with GWN

For the experimental comparison between SDGWN and GWN, a randomly selected subset (containing 117 images of 10 persons) from the FERET database was used. We created a training set consisting of three images for each person, the remainder constituting the test set. When a test pattern was presented to either SDGWN or GWN model, the representation quality is evaluated as elaborated in last section. In particular, each SDGWN/GWN model used 80 Gabor wavelets, and a variant of GWN referred to as GWN-1 was tested which selects the best representation from GWNs. Table I shows the results. Clearly, SDGWN yielded generally better results for new images than GWN-1 did.

B. Experiments with FERET Datasets

The second release of the FERET consists of 14051 8-b grayscale images of human heads with views ranging from frontal to left and right profile, and the database design takes into account variable factors such as different expressions, hairstyles and illuminations. We selected those subjects with more than five frontal or near-frontal $(\text{Pose angle} \leq$

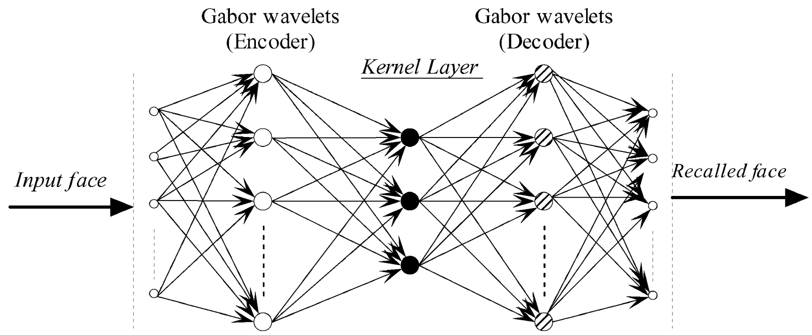

Fig. 2. GWAM.

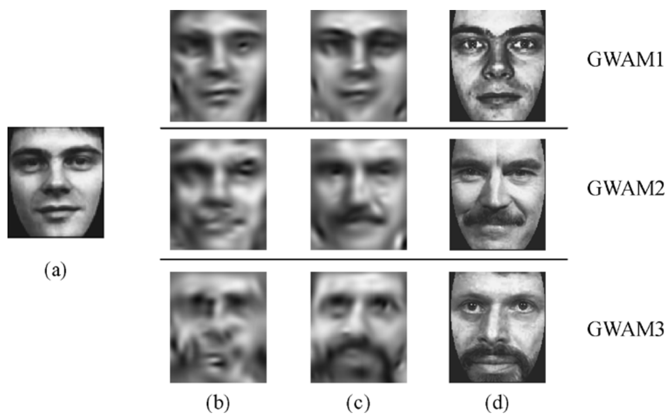

Fig. 3. Face recognition process by GWAMs. (a) Novel image to be recognized. (b) SDGWN representations as the keys for embeded KAM models in GWAMs. (c) Final reconstructions by GWAM models. (d) Typical training images of each subject. The subject model (here the top one) yields the best representation indicates the recognized person.

TABLE I
PERFORMANCE OF GWN AND SDGWN AS A FUNCTION OF APPROXIMATION ACCURACY FOR NEW IMAGES

	mean	var
GWN-1	0.843	0.028
SDGWN	0.893	0.018

15°) images, so that we can conveniently investigate all the systems over variable size of training/testing set. The resulting dataset consists of 119 persons with 927 images, which were preprocessed by a normalization program according to eye position. The final images are at the size of 130 × 150 pixels. The training set P was built by randomly selecting m ($m = 3$ or 4) samples per person from the dataset while the remainder forming the test set.

A variant of Eigenfaces technique called PCA-nearest-neighbor [3] was compared which used $n = 25$ top principal components. Another technique called ARENA [3] was also tested, which employs reduced-resolution images and we set $\delta = 10$ (see [3]) for it. All the results are given in Table II, where M stands for the number of Gabor wavelets engaged in each GWAM.

A performance evaluation method suggested by the developers of FERET [14] was implemented. In specific, a recognition will be regarded as correct if the true object is in the top n matchs. For example, if $n = 5$ and 80 recognitions out of 100 have their true identities in each top 5 matches, the cumulative score for R_5 will be $80/100 = 0.8$. Fig. 4 illustrates the cumulative scores produced by each algorithm. It can be seen that GWAM yielded clearly better results than the others.

TABLE II
RECOGNITION ACCURACY FOR FERET DATASET

n	PCA	ARENA	KAM	GWAM	
				M=80	M=16
3	54.3	55	84.7	99.3	95.8
4	55.2	55.2	91.6	99.6	99.1

C. Experiments with ORL Database

ORL database contains 40 subjects with 10 images per subject. The images at the size of 92 × 112 pixels were acquired under variable lighting condition, facial expressions and viewpoint. This database allows us to compare our system with other techniques such as SOM+CN [9] using their published results.

We randomly selected a limited number (3 or 5) of faces out of 10 to set up a GWAM model for each subject, and then count the recognition accuracy on the remaining faces. The results are given in Table III, and it can be seen that GWAM achieved perfect recognition results in the experiment.

D. Experiments with the AR Face Database

The AR face database from Purdue University contains over 3000 color images of the frontal view faces of 126 people, with roughly 26 different images per person, recorded in two different sessions separated by two weeks and each session consisting of 13 images [16]. In particular, AR face images show dramatically varying lighting conditions.

We randomly selected 42 male subjects and 44 female subjects (each subject has ten images) to set up our experimental dataset. We preprocessed all the images by a normalization process which first converted images to greyscale and then performed a geometrical normalization according to eye position such that the left/right eye in every image is at same position. 2 normal views from each person were used for training while the other 6 images for testing.

Each GWAM model employs 80 Gabor wavelets. For each kind of lighting condition, recognition performance was evaluated, respectively. The GWAM system is compared with a recently proposed technique called line edge map (LEM) [17], which uses polygonal line segments of edges to represent a face and has achieved good face recognition results on the AR database. The results are summarized in Table IV, where the performances of other techniques were duplicated from [17]). From Table IV, we can find that the GWAM significantly outperformed other techniques in general. But in "right light on" case it achieved a similar rate to that of LEM. Our system also showed robust performance under varying lighting conditions, as the maximum variation of accuracy with it is only 7%, comparing to 18.8% with LEM, 27.6% with Edge Map and 37.5% with Eigenface.

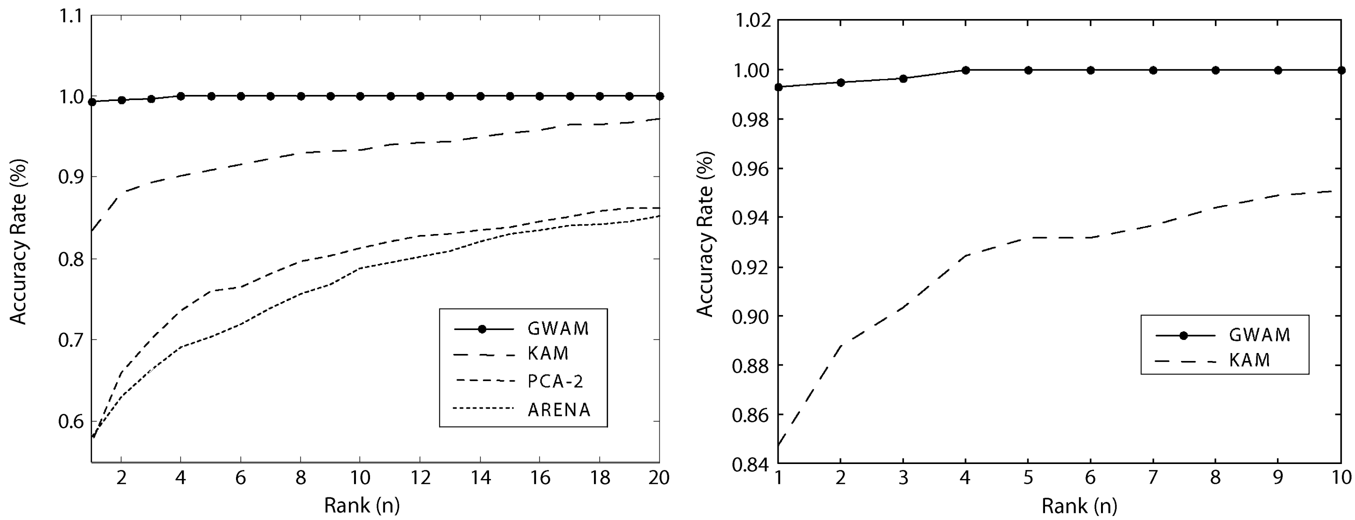

Fig. 4. Comparision of accumulated accuracy on FERET. The left one illustrates accumulated accuracies by GWAM, KAM, PCA-2, and ARENA, respectively. The right one is a zoom-in on the left-top corner of left figure to bring out the detail regarding the GWAM and KAM models. The horizontal axis denotes the rand, while the vertical axis is the percentage of correct matches.

TABLE III
RECOGNITION ACCURACY FOR ORL DATABASE

n	PCA	SOM+CN	ARENA	KAM	GWAM	
					M=80	M=9
3	81.8	88.2	92.2	94.3	100	97.9
5	89.5	96.5	97.1	98.2	100	98.8

TABLE IV
RECOGNITION ACCURACY FOR AR DATABASE

L. Cond.	Eigenface	Edge map	LEM	GWAM
Left .	26.8%	82.1%	92.9%	96.5%
Right .	49.1%	73.2%	91.1%	90.9%
Both .	64.3%	54.5%	74.1%	89.0%

V. CONCLUSION

In this letter, we have proposed a new face recognition system based on GWAM model, which combines two recently proposed neural network models, namely GWN and KAM. The GWAM model aims to provide an efficient approach to the representation and recognition of a particular subject. Extensive experiments have been conducted, and the results demonstrates that our system can offer excellent performance for face recognition in comparison with other techniques.

REFERENCES

[1] R. Chellappa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," *Proc. IEEE*, vol. 83, no. 5, pp. 705–740, May 1995.
[2] S. Edelman, *Representation and Recognition in Vision*. Cambridge, MA: MIT Press, 1999.
[3] T. Sim, R. Sukthankar, M. Mullin, and S. Baluja, "High-performance memory-based face recognition for visitor identification,", Tech. Rep. JPRC-TR-1999-001-01, 1999.
[4] T. Kanada, "Picture processing by computer complex and recognition of human faces," Dept. Inform. Sci., Kyoto Univ., 1973.
[5] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neurosci.*, vol. 3, pp. 71–86, 1991.
[6] D. Gabor, "Theory of communications," *J. Inst. Elect. Eng.*, vol. 93, pp. 429–557, 1946.
[7] J. Daugman, "Complete discrete 2D Gabor transform by neural networks for image analysis and com-pression," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 7, pp. 1169–1179, Jul. 1988.
[8] V. Krueger, "Gabor wavelet networks for object representation," Ph.D. dissertation, Christian-Albrecht Univ., Germany, 2001.
[9] S. Lawrence *et al.*, "Face recognition: a convolutional neural network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.
[10] S. Lin, S. Kung, and L. Lin, "Face recognition/detection by probabilistic decision-based neural network," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 114–132, Jan. 1997.
[11] T. Kohonen, *Associative Memory: A System Theoretic Approach*. Berlin, Germany: Springer-Verlag, 1977.
[12] B. Zhang, H. Zhang, and S. S. Ge, "Face recognition by applying wavelet subband representation and kernel associative memory," *IEEE Trans. Neural Netw.*, vol. 15, no. 1, pp. 166–177, Jan. 2004.
[13] Q. Zhang and A. Benveniste, "Wavelets," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, pp. 889–898, Nov. 1992.
[14] P. Phillips *et al.*, "The FERET evaluation methodology for face-recognition algorithms,", Tech. Rep. NISTIR 6264, 1998.
[15] F. S. Samaria and A. Harter, "Parametrization of a stochastic model for human face identification," in *Proc. IEEE Workshop Applications Computer Vision*, 1994, pp. 138–142.
[16] A. Martinez and R. Benavente, "The AR face database,", CVC Tech. Rep. #24, 1998.
[17] Y. Gao and K. H. Leung, "Face recognition using line edge map," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 6, pp. 764–778, Jun. 2002.