Face Recognition by Combining Kernel Associative Memory and Gabor Transforms

Bai-ling Zhang, Clement Leung
School of Computer Science and Mathematics,
Victoria University,
{bzhang, clement}@csm.vu.edu.au

Yongsheng Gao
School of Engineering,
Griffith University,
yongsheng.gao@griffith.edu.au

Abstract

Kernel associative memory (KAM) has previously been proposed as an efficient scheme for face recognition. In this paper, a hybrid method of combining KAM and Gabor wavelet transform is proposed. In this method, face images of each person are first decomposed into their spatial/frequency domains by Gabor transforms, which are then modelled by a KAM. While Gabor properties of orientation selectivity and spatial frequency selectivity provide discriminating features, KAM offers the means to capture the important intra-class variations. Experimental results obtained on two standard face databases demonstrated that the proposed method consistently improved the system performance.

1. Introduction

Face recognition technology has received much attention in recent years because of the wide range of applications. If operating under strictly controlled environments, face recognition systems can reach high accuracy. But the general face recognition problem remains one of the most challenging research topics since there are still lots of difficulties when dealing with the problems such as occlusion, illumination variations and view changes.

Two issues are most important in designing a face recognition system. The first issue is how to efficiently represent faces by appropriate features. In the literature, the appearance-based approaches of directly processing facial images as 2D holistic patterns have often been used [7]. The appearance based approaches, however, are usually sensitive to some variations from pose, illumination and expression. Another extensively used facial feature extraction method is the Gabor wavelet transform, which allows description of spatial frequency structure in the image while preserving information about spatial relations. These features are known to be robust to some variations, e.g., pose and facial expression changes.

The second issue in face recognition is how to classify the facial images based on the chosen features. Most of the face recognition research treat the problem as a multi-class classification, with many different kind of classification methods having been applied, such as the nearest neighbor classifier and Linear Discriminant Analysis. A special difficulty arisen in these multi-class classification schemes is the small sample size problem, which means only a few of sample face images available for a given subject. To solve the problem, face recognition can be studied as a one-class classification problem [6]. To facilitate this methodology, a memory-based scheme could be adopted [5], which use all the training samples for each subject to build a subject-specific model and recognition is carried out by comparing which model offers the best description of the query face. In [5], a kernel associative memory (KAM) model was proposed, which introduces kernel methods to nonlinearly map the data into some high-dimensional feature space through operating a kernel function with input space.

This paper presents a new, hybrid face recognition scheme by combining KAM and Gabor transform. In our method, face images of each person are first decomposed into their spatial/frequency domains by Gabor wavelet transforms, which are then modelled by a kernel associative memory. While Gabor properties of spatial localization, orientation selectivity and spatial frequency selectivity provide local and discriminating features, KAM offers the means to capture all intra-class variations.

2. Gabor Wavelet Transform of Face Images

Gabor wavelet has been widely applied in face recognition, for example, the Dynamic Link Architecture framework [2]. The main feature of Gabor wavelet is its description of spatial frequency structure in the image while preserving information about spatial relations, which is an ap-
proximation to the orientation and spatial frequency selective properties of simple cortical neurons. Such simultaneous spatial/frequency analysis plays a central role in feature extraction from images, and the salient visual properties captured by Gabor transform are instrumental in improving a face recognition performance, particularly the robustness to some variations.

The 2-D Gabor wavelet is a two-dimensional plane wave with wavelet vector \( \vec{k}_j \) restricted by a Gaussian envelope function with relative width \( \sigma \), which can be defined as:

\[
\psi_j(\vec{x}) = \frac{||\vec{k}_j||^2}{\sigma^2} \exp\left(\frac{||\vec{k}_j||^2}{2\sigma^2}ight) \left[\exp\left(i k_j \vec{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right]
\]

(1)

where

\[
k_j = k_m e^{i \varphi_n}, \quad k_m = \frac{0.5 \pi}{(\sqrt{2})^m}, \quad \varphi_n = \frac{n \pi}{8}
\]

(2)

In Eqn. 1, the multiplicative factor \( ||\vec{k}_j||^2 \) tunes filters to different spatial frequency bands with approximately equal energies. The first term in the square brackets determines the oscillatory part of the wavelet. The second term makes the wavelets DC-free.

Let \( I(x, y) \) be a gray level image. The Gabor wavelet representation of \( I \) is the convolution of the image with a family of Gabor kernels, which can be described as

\[
O_j(\vec{x}) = I(\vec{x}) * \psi_j(\vec{x})
\]

(3)

where \( \vec{x} = (x, y) \), and \( * \) denotes the convolution operator. The result \( O_j(\vec{x}) \) is also called Gaborface [4]. Figure 1 shows the Gaborfaces from a given a face image at 2 scales and 3 orientations.

![Figure 1. A sample face image represented by the magnitude part of Gabor wavelet transform.](image)

3 KAM as Computational Model of Faces

3.1 Kernel Associative Memory Revisited

For a specific class, suppose we have \( N \) prototypes \( \{x_1, x_2, \ldots, x_N\} \). A prototype is predefined as a vector in an \( I \) dimensional space. An outer-product matrix can be constructed from the prototype vectors:

\[
W = XX^T
\]

(4)

where \( X \) is an \( I \times N \) matrix in which the \( k \)th column is \( x_k \). The \( W \) can be called memory matrix, which can also be constructed by applying the pseudo-inverse to \( X \):

\[
W = XX^+\nonumber
\]

(5)

where \( X^+ \) is the pseudoinverse of \( X \). \( W \) provides a description for the corresponding class in the sense that any prototype in the class can be represented as

\[
\hat{x}_n = Wx_n, \quad n = 1, \ldots, N
\]

(6)

If a memory matrix \( W^{(k)} \) is constructed for the \( k \)th person, a query face \( y \) can be classified into one of \( C \) classes based on a distance measure of how far \( y \) is from each class. The distance can be simply the Euclidean distance

\[
d(y, \hat{x}_k) = ||y - W^{(k)}y||, \quad k = 1, \ldots, C
\]

(7)

Kernel associative memory applies kernel method to map the space of input data into a high-dimensional feature space through a mapping \( \Phi \), and then reconstruct the input via the feature space. However, we need not explicitly compute the mapped pattern \( \Phi(x) \), but the dot product between mapped patterns, which are directly available from the kernel function which generates \( \Phi(x) \). Denote by

\[
k(x, x') : = < \Phi(x), \Phi(x') >
\]

a kernel corresponding to \( \Phi \). A popular option of \( k \) is Gaussian radial basis functions:

\[
k(x, x') = \exp\left(-\frac{||x - x'||^2}{2\sigma^2}\right)
\]

(9)

The kernel associative memory can be described as

\[
\sum_{n=1}^{N} w_n k(x_n, x) \]

where \( w_n, n = 1, \ldots, N \) are weights determined by the following least square objective:

\[
J(W) = \min ||X - WK||
\]

(10)

where \( k \) is a vector in which the \( i \)th element is equal to \( k(x_i, x) \). The optimum values for \( W \) can be obtained by using LS approximation. For \( N \) kernels and \( k^{(n)} = [k_1^{(n)}, \ldots, k_N^{(n)}]^T \), \( n = 1, \ldots, N \), using matrix representation \( K = [k^{(1)}, \ldots, k^{(N)}] \), \( W \) can be calculated as
\[ W = X K^+ \]  

(11)

where \( K^+ \) is the pseudo-inverse of \( K \).

The recalling of the \( k \)-th face can be directly achieved by first calculating a kernel vector \( k^{(k)} \) from the query face image \( x_k \) (converted to vector), and then premultiplying the kernel vector \( k^{(k)} \) by the matrix \( W \):

\[ \hat{x}_k = W k^{(k)} \]  

(12)

where \( \hat{x}_k \) represents the estimation of the \( k \)-th face.

## 3.2 KAM With Gaborfaces

Suppose that we have \( C \) subjects, with each subject having \( N_j \) face images \( x^{(j)}_i \), \( j = 1, \cdots, C; i = 1, \cdots, N_j \). For each face image, Gabor wavelet transform will yield \( m \times n \) Gaborfaces as elaborated in Section 2, with \( m \) the number of scales and \( n \) the number of orientations. Accordingly, a subject specific KAM model can be constructed as follows.

### Figure 2. KAM construction for one subject.

Suppose \( x_1, x_2, \ldots, x_M \) are the training face images for a subject. As an example, we choose \( m = 3 \) and \( n = 4 \) for Gabor transforms. Then a concatenated kernel product vector \( k_{1,1}^{(1)}, k_{1,2}^{(1)}, \ldots, k_{1,12}^{(1)}, k_{2,1}^{(2)}, k_{2,2}^{(2)}, \ldots, k_{2,12}^{(2)}, \ldots, k_{M,1}^{(M)}, k_{M,2}^{(M)}, \ldots, k_{M,12}^{(M)} \) is calculated from the Gaborfaces. Using \( K \) to represent the matrix with each column the corresponding concatenated kernel product vector, and \( X \) to represent the matrix with each column the training face images (converted to vectors), the LS solution of \( W \) is given in Eqn.(10). The model construction for the a specific subject is illustrated in Figure 2.

In the testing phase, a query face image will be first transformed into \( m \times n \) Gaborfaces, which will then be compared with training Gaborfaces from different subjects to yield the corresponding concatenated kernel product vectors \( k^{(j)}_i \), \( j = 1, \cdots, C \). The recalled \( \hat{x}^{(j)}_i \) is given by Eq.(12). The closeness between the query image \( x \) and \( j \)-th KAM model can also be measured by the cosine of the angle between the vectors \( x^{(j)}_i \) and \( x \), i.e.,

\[ \cos(\hat{x}^{(j)}_i, x) = \frac{x^T \hat{x}^{(j)}_i}{\|x\| \|\hat{x}^{(j)}_i\|} \quad j = 1, \cdots, C \]  

(13)

The reconstruction for a query face image can be further illustrated in Figure 3.

### Figure 3. Illustration of the recognition stage.

## 4 Experimental Results

### 4.1 Experiments with UMIST database

The UMIST Face Database [1] consists of 564 images of 20 people. Subjects cover a range of race/sex/appearance. Examples of the UMIST database are shown in the top of Figure 4. All these images of UMIST database are cropped to the size of \( 112 \times 92 \). In this experiment, we randomly select roughly half number of the face images of each person as the training set, and the remaining images as the test set.

To evaluate the recognition performance, we use the performance measurement cumulative match scores, by which an identification is regarded as correct if the true object is in the top \( n \) matches. The bottom plot in Figure 4 illustrates the cumulative match scores of the proposed Gabor-KAM method (3 orientations and 2 scales), KAM with face images represented by the LL subband coefficient of Daubechies wavelet transform [5], and Eigenface operating on downsized images with resolution 56 \( \times \) 46. The experiment results demonstrated a better robustness of the proposed method to the pose variances.

We also compared the recognition performance with different options of the orientation/scale parameters. It seems that relatively small number of Gaborfaces are sufficient to achieve a good result. The comparative results are summarized in Table 1.
Table 1. Recognition accuracy for UMIST dataset with Gaborfaces of different orientations and scales.

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Scale</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>2.1%</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2.1%</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1.7%</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>2.7%</td>
</tr>
</tbody>
</table>

4.2 Experiments with AR Database

AR faces [3] consist of over 3200 color images of the frontal images of faces of 126 subjects. There are 26 different images for each subject. For illustration, some images from one subject are shown in top of Fig. 5.

We carried out similar experiments to evaluate the performance of the proposed Gabor KAM scheme. All the images are convoluted with a given number of Gabor kernels, yielding the Gaborfaces of same size as original image. We compared the recognition accuracies on each of the eight available data collections with three different methods: (1).the Gabor KAM method, with 3 orientations and 2 scales, (2).the KAM with Daubechies wavelet LL subband representation of faces, which is of size $52 \times 43$, and (3).the Eigenface method with downsized image $52 \times 43$ as the facial feature. The comparison of recognition performances is illustrated in the bottom of Figure 6, which showed again that the Gaborfaces is better than the Daubechies wavelet LL subband representation.

5 Conclusion

This paper further investigated the Kernel Associative Memory (KAM) in face recognition by proposing a hybrid scheme of combining KAM and Gabor wavelet transform. By decomposing each face image into their spatial/frequency domains using Gabor wavelet transforms and applying kernel methods in the construction of the descriptive associative memory for each subject, important intra-person variations can be captured. The performance of the new method has been tested using data set from the UMIST and AR databases, showing noticeably improved results comparing to the KAM with Daubechies wavelet LL subband representation of faces and the Eigenface method.

References