Graph Matching For Face Recognition

Using Genetic Algorithms

Ehsan Fazl Ersi

Azad University of Mashad
Faculty of Engineering
Dep. of Computer Engineering
Mashad, Iran
ehsan@fazl.ws

Kiana Hajebi

Azad University of Mashad
Faculty of Engineering
Dep. of Computer Engineering
Mashad, Iran
kiana@hajebi.ws

Abstract

Object adapted graph is a flexible and powerful representation mechanism that has been successfully applied in pattern recognition. In this mechanism objects such as human faces are described by a labelled graph, wich the facial features comprises the nodes and topographical relations form the edges. In this paper we present a system for face recognition, which is based on object adapted graphs and outcomes from making some extensions to the EBGM method in order to optimize the face graph matching process. We use packs instead of jets to describe the facial features more complete. Therefore the system doesn't need an extensive FBG to cover the possible variations in the appearance of facial features. Besides, we propose a new method for Image graph extraction, which is based on genetic algorithms and uses only the magnitude of the coefficients. The system has been tested on ORL face bank and showed satisfactory performance.

Keywords: face recognition, Gabor wavelet, Genetic algorithms, pack, ORL database.

1   Introduction

Automatic recognition of the human faces from single images is considered as one of the fundamental problems in pattern analysis and has been an area of active research for past 30 years. Face image variations in lighting conditions, size of the head, facial expression and poses, etc. are some instances that caused computerized face recognition system can't achieve the same accuracy of human vision system. The system introduced here, is based on Elastic Bunch Graph Matching system, which is presented by L Wiskott et al [8]. In the mentioned system the faces were described

by an adapted graph, which the nodes are labeled with a set of Gabor wavelet coefficients called jet, and the edges are labeled with distance information. Since there are some possible variations in the appearance of faces, to find the face graph in new images, a stack-like data structure was introduced called Face Bunch Graph (FBG), which served as a generalized representation of faces by combining graphs of individual faces. The graph matching algorithm tries to find a position for each node of the graph, which maximizes the feature similarity and minimizes the topography costs at the same time [8]. After having extracted the image graph, face recognition was accomplished by comparing an image graph to each graph in the model gallery (data base) and picking the one with the highest similarity value.

We have made some extensions to the proposed system in order to improve the recognition rate and reduce the time of matching process. A Gaussian set of jets called pack is used to describe the facial features more thoroughly and less sensitive to the exact position and states of facial features. Nodes of face graph labelled with these packs. Like most of the matching algorithms, our matching process has two phases: Global Move, which approximates the best matching position of face by preventing graph distortion, and Local Move, which allows nodes to move individually and refine the matching. We use the average FBG to find approximate face position (same as EBGM) and for local move, we propose a new method, which is based on genetic algorithms and uses only the magnitude of the coefficients.

2   Gabor Wavelet Filters

The Gabor wavelet filters, closely model the receptive field properties of cells in the primary visual cortex and remove most of the variability in images due to

variations in lighting and contrast [1]. For each filter the complex Gabor wavelet coefficients are made up from two parts: one is real part, which corresponds to even symmetry of a receptive field (RF) cell, and the other is imaginary part, which corresponds to odd symmetry of a RF cell (Fig.1).
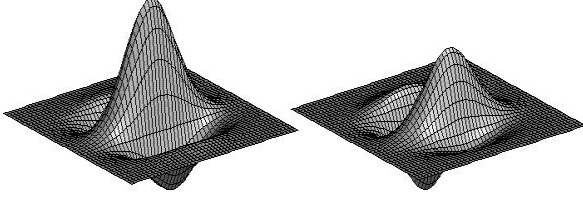

Figure 1. Real and imaginary components of a two-dimensional Gabor function

This observation can be modelled by a Cosine function for even symmetry and a Sine function for odd symmetry, as follow (1),

$\Psi_{\vec{k},+}$ corresponds to even-phase filter:

$$\Psi_{\vec{k},+}(\vec{r}) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2(\vec{r}-\vec{r}_0)^2}{2\sigma^2}\right) \times$$

$$\left[\cos(\vec{k}.(\vec{r}-\vec{r}_0)) - \exp\left(-\frac{\sigma^2}{2}\right)\right]$$

and, $\Psi_{\vec{k},-}$ corresponds to odd-phase filter:

$$\Psi_{\vec{k},+}(\vec{r}) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2(\vec{r}-\vec{r}_0)^2}{2\sigma^2}\right) \times \left[\sin(\vec{k}.(\vec{r}-\vec{r}_0))\right]$$

If both symmetries are combined in a complex notation with:

$$\exp(i\vec{k}\vec{r}) = \cos(\vec{k}\vec{r}) + i\sin(\vec{k}\vec{r})$$

Then Gabor biologically motivated filter can be formulated:

$$\Psi_j(\vec{r}) = \frac{k_j^2}{\sigma^2} \exp\left(-\frac{k_j^2(\vec{r}-\vec{r}_0)^2}{2\sigma^2}\right) \times$$

$$\left[\exp(\vec{k}_j.(\vec{r}-\vec{r}_0)) - \exp\left(-\frac{\sigma^2}{2}\right)\right]$$

In these equations, wave vector $\vec{k}$ determines the spatial frequency and orientation tuning:

$$\vec{k} = \begin{pmatrix} f_v \cos\varphi_\mu \\ f_v \sin\varphi_\mu \end{pmatrix}, \quad f_v = 2^{-\frac{v+2}{2}\pi}, \quad \varphi_\mu = \mu\frac{\pi}{8}$$

The result of filtering a new image *I* with a spatial Gabor function (spatial wave vector), is defined as a convolution of the image (Fig.2):

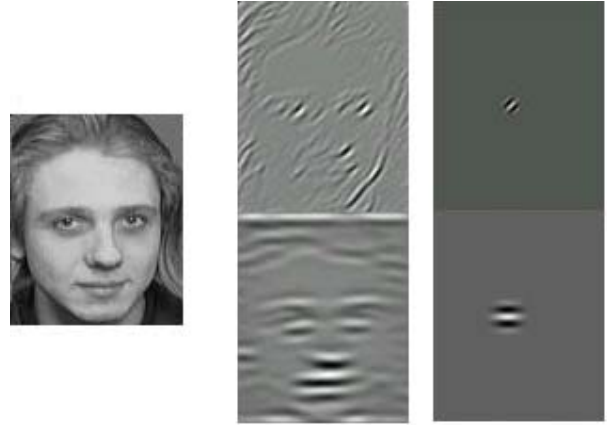$$R(\vec{r}_0) = \int \Psi(\vec{r}_0, \vec{r}) I(\vec{r}) d\vec{r}$$

Figure 2. Example of filtering a facial image with two Gabor filters used.

2.1 Jets

Typically the feature-based face recognition systems need to describe some facial point in the face image with high-level information content. One of the most powerful descriptions of these facial features is done by jets, which are based on Gabor wavelet transform.

A jet is a set of convolution coefficients for kernel of different frequencies and orientations at one image point, which describes the local neighboring pixels, surrounding that given point.

We use a set of 5 frequencies and 8 orientations; similar to the one is used in [8], in our system. Therefore for each point of the image we have 40 Gabor wavelet coefficients.

In order to compare jets, we use only the magnitudes of Gabor coefficients because they vary slowly with position, while the phases are so sensitive to location and the near image point's phases are very different, although representing almost the same local feature. The similarity between two jets is given by the normed dot product of them:

$$S_J(J_1, J_2) = \frac{\sum_i a_{1i} a_{2i}}{\sqrt{\sum_i a_{1i}^2 \sum_i a_{2i}^2}}$$

Where a_j is the jth Gabor coefficient in jet J.

2.1 Packs

Although jets are suitable for facial features description, but since they represent only a small region surrounding a given pixel in the image, they can't cover the wide range of variations in the appearance of them. For example we cannot describe the different states of someone's eyes such as opened, closed, narrow, etc. with just one jet, and this makes the recognition so sensitive.

To solve the problem we present a new structure called *pack*, which is a Gaussian set of jets in form of a $w * w$ window and is calculated as follow:

$$P = \left\{ p_{i,j} = \frac{1}{1 + \sqrt{i^2 + j^2}} J_{ij} \right\}$$

We employ $w = 7$ in our system and index $i, j = $ -3, -2 … 2, 3. The packs are comparable by jets similarity function as follow:

$$S_P(P_1, P_2) = \sum_{i,j} S_J(p_{1ij}, p_{2ij})$$

For understanding the difference between jets and packs, we explain it by an example. If you cover the face by a flat sheet, and make a small pore on it, detecting the pupil of left eye from that pore implies moving the sheet around the probable positions to get the situation and estimate the exact position of the left eye's pupil. What will happen if the left eye were closed? Of course detecting the approximate position of left eye's pupil will be so difficult and may be impossible. Now, what will happen if we make the pore wider and define the pupil, with its neighborhood, which are the other parts of eye even open, narrow or closed? Of coarse in this case we can approximate the position of left eye's pupil more accurate, although the training condition (open eye) and test condition (narrow/closed eye) were not completely similar.

3 Graphs

3.1 Face graph

Faces are represented by a labelled graph, called face graph. The facial features such as pupils, the corner of mouth, the tip of the nose, top and bottom of the ears, etc. comprise the nodes of this graph and distances between them form its edges, which describe topographical relations (a face graph from scientific point of view is given in fig.3.a). The nodes of the face graph are labelled by packs. We compare face graphs by a graph similarity function, which depends on the packs similarities and the topography of graphs in comparison with each other:

$$S_G(g_1, g_2) = \frac{1}{N} \sum_n S_p\left(P_n^{\,1}, P_n^{\,2}\right) - $$
$$\frac{2}{E} \sum_e \frac{\left(\Delta \vec{x}_e^{\,1} - \Delta \vec{x}_e^{\,2}\right)^2}{\left(\Delta \vec{x}_e^{\,2}\right)^2}$$

where $\Delta \bar{x}_e^{\,1}$ and $\Delta \bar{x}_e^{\,2}$ are the distance vectors used as labels of edges e in both graphs.

3.2 Face bunch graph

While there are considerable variations in the appearance of facial features, jets must be taken in different states of these features, to make the graph extraction possible for any face images. Therefore a more comprehensive representation is needed in order to account for all possible within-class variations. Here we use The Face Bunch Graph introduced in [8], which is a stack-like structure and constructed by combining graphs of individual sample faces (the face bunch graph from scientific point of view is given in fig.3.b). So each facial feature (each node of face graph) is represented by a set of alternatives in order to cover the variations.
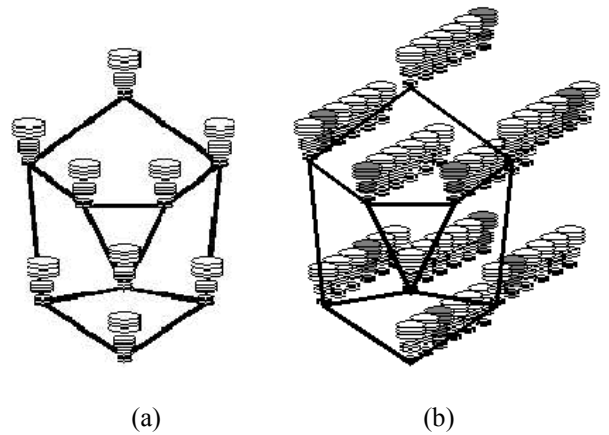

(a)　　　　　　　　(b)

Figure 3. (a) Face Graph   (b) Face Bunch Graph

4 Matching process

Like most of the matching algorithms, our matching process consists of two phases:

4.1 Global move

Our global move strategy is the same as EBGM [8]. We first make a model graph by taking the average amplitude of the packs in each bunch of the FBG, in order to condense the FBG into an average graph. Then we scan the image with a model graph and evaluate its similarity at each location, to determine the approximate face posit-ion. The position corresponding to the best fit is used as starting point for local move step.

4.1 Local move

In local move, nodes are allowed to move individually, based on an evolutionary method such as Genetic algorithm for locating in best position and refine the matching.

4.2.1 Introduction to GA

Genetic algorithm is a method for moving from one population of chromosomes to a new population by using a kind of *natural selection* together with the genetic-inspired operators of crossover and mutation. Each chromosome consists of *genes* and each gene being an instance of a particular *allele* [5].

Selection operator. The selection operator chooses those chromosomes in the population that will be allowed to reproduce. On average, the fitter chromosomes produce more offspring than the less fit ones.

Crossover operator. The Crossover operator exchanges subparts of two chromosomes, roughly mimicking biological haploids recombination organisms.

Mutation operator. The Mutation operator randomly changes the values of some locations in the chromosome.

4.1.2 Local distortion

Since all of the images we use in this system have been normalized in size and rotation, their face graph topographies are too close. Therefore the approximate position of each node, which is deter-mined in the global move stage, needs to relocate in a small surrounding region.

If we propose the face graph with a chromosome, then each gene of that chromosome containing a node position and its small surrounding region comprises the allele of that gene (fig.4).

After encoding the problem and making a random initial population, we allocate a fitness value to each chromosome based on its similarity to FBG. And perform the selection using *roulette wheel* to choose the individuals in the population in order to do crossover. Then we use two-point crossover, in which two loci are chosen at random and the segment between them are exchanged to create offspring for the next generation. Offspring are subject to point mutation in which a random locus is selected and its rate changes with a random value from its allele. Creating the new generation will continue until the average fitness value of the chromosomes in the population is close to optimum.
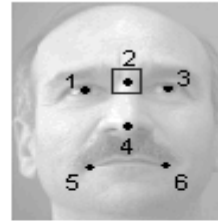
$$\text{Chromosome} = (\,\underbrace{\vec{r_1}, \vec{r_2}, \vec{r_3}, \vec{r_4}, \vec{r_5}, \vec{r_6}}_{\text{genes}}\,)$$

$$\vec{r_2}(x,y) = \begin{cases} x \in [\vec{r}_{2,x} - n] \\ y \in [\vec{r}_{2,y} - n] \end{cases} \rightarrow \begin{array}{c} \text{Allele} \\ \text{of} \\ \vec{r_2} \end{array}$$

n : Surrounding region length

Figure 4. Example of chromosome, gene, and allele in our model.

5 Recognition

After locating the face graph in the fittest position, the recognition is possible by comparing the extracted image graph to all model graphs in the gallery and choosing the one with the highest similarity value. The comparison is done by graph similarity function without the grid distortion effect:

$$S_G(g_1, g_2) = \frac{1}{N} \sum_n S_p\left(P_n^1, P_n^2\right)$$

With this little computational effort, the time of recognition among a medium database is so promising.

6   Experiments

6.1 Databases
We have tested our system on two different databases. The first one is ORL face bank (fig.5), which contains 400 images from 40 individuals persons, in different poses and the other one is a local bank (fig.6), we made from our university in the same size as ORL.

Figure 5. Example of ORL database.

Figure 6. Example of our local database.

6.2 Results

To test our system we first normalize the images in size and rotation, and for each image make the inter-eye line horizontal with distance of 25 pixels and cut it to standard size 60 x 60.
The system has been tested on ORL database and showed a 97.7 correct recognition rate.
To test the system on our local database we first omit the tenth picture, which shows the individuals wearing sunglasses, and the correct recognition rate was as high as 99.1 percent. Whereas sunglasses cover the eye features, adding the tenth image of individuals to the database has decreased our recognition rate to 93.7 percent.

References

[1] Daugman, J. G., 1988, "Complete discrete 2-D Gabor transform by neural networks for image analysis and compression", IEEE Transactions on Acoustics, Speech and Signal Processing, 36, 1169-1179.

[2] Kalocsai P, Zhao W, Elagin E, 1998, "Face similarity space as perceived by humans and artificial systems", Proceedings of IEEE Confe-rence on Automatic Face and Gesture Recogni-tion, 3, 177-180.

[3] Lades, M., Vorbruggen, J. C., Buhmunn, J., Lange, J., Malsburg C.v, C., R. P., Konen, W., 1993, "Distortion invariant object recognition in the dynamic link architecture", IEEE Transactions on Computers, 42, 300-311.

[4] Lyons M, Plante A, Jehan S, Inoue S, Akamatsu S., 1998, "Avatar Creation using Automatic Face Recognition", ACM Multimedia, 98, 427-434.

[5] Mitchell M, 1999, "An Introduction to Genetic Al-gorithms", MIT Press, Cambridge, Massachusetts.

[6] Penev P, Atick J, 1996, "Local feature analysis: a general statistical theory for object representation", Network: Computation in neural systems, 7, 477-500.

[7] Pentland A, Moghaddam B, Starner T, 1994, "View-based and modular eigenspace for face recognition", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 94, 84-91.

[8] Wiskott L, Fellous J.M., Kruger , Malsburg C.v, 1997," Face recognition by elastic bunch graph matching", Pattern Analysis and Machine Intelligence, 19, 775-779.

[9] Penev and Sirovich, 2000, "The global dimensionality of face space," In proc. 4th Int'l Conf. On Automatic Face and Gesture Recognition, 264-270.

[10] Moghaddam and Pentland, 1998, "Probabilistic visual learning for object representation," IEEE Trans. On pattern analysis and machine Intelligence, 19, 669-710.

[11] C. Liu and H. Wechsler, 2002, "Gabor Feature

Based Classification Using the Enhanced Fisher Linear Discriminant Model For Face Recognition," IEEE Trans. Image Processing, 11,467-476