Proceeding, Seminar of Intelligent Technology and Its Applications (SITIA' 2000)
Graha Institut Teknologi Sepuluh Nopember, Surabaya, 19 April 2000

Facial Landmark Detection using a Gabor Filter Representation and a Genetic Search Algorithm

Resmana Lim [1,2], M.J.T. Reinders [1], Thiang [2]

[1] Information and Communication Theory Group
Faculty of Information Technology and Systems
Delft University of Technology
P.O. Box 5031, 2600 GA Delft, The Netherlands
{R.Lim; M.J.T.Reinders}@its.tudelft.nl

[2] Electrical Engineering Department
Petra Christian University
Siwalankerto 121-131, Surabaya 60236, Indonesia
{resmana; thiang@petra.ac.id}

Abstract

This paper proposes a method that finds landmark points on the face image, which is one of the main tasks in a face recognition system. Salient facial landmark detection is important because it enables face normalization and leads to size and orientation invariant face recognition. The presented approach is based on an elastic graph matching technique and uses a genetic algorithm to perform the search. The feasibility of our methodology for detection tasks related to face landmark points detection has been deployed using the ORL face image database. Experiments show satisfactory results under a relatively wide conditions. The GA searching approach is essential because it reduces the search space considerably.

KEYWORDS : facial landmark detection, face recognition, elastic graph matching, genetic algorithm, Gabor filtering.

1. Introduction

Face detection and detecting facial landmarks (such as position of eyes, nose, mouth, etc.) play an important role in face recognition systems. In practical face recognition system, these subsystems determine the quality of the recognition rate because they are used to normalize the recorded image(s). Once normalized, face images can quite reliably recognized using well-known techniques like principle component analysis [1] [2]See also [3] for more information on face recognition. This paper, however focuses on the robust and accurate detection of landmark points on the face. The presented approach uses Gabor filter responses to effectively represent the landmark points. The choice of Gabor filter responses is because they biologically motivated [4]. Further, feature extraction based on Gabor filtering has been deployed successfully for texture analysis [5], character recognition [6], fingerprint recognition [7], and face recognition [8] [9]. The essence of the success of Gabor filters is that they remove most of the variability in image due to variation in lighting and contrast, at the same time being, they robust against small shift and deformation. Recent work [10] has shown that a Gabor approach for local feature extraction outperformed PCA (Principle Component Analysis), FLD (Fisher's Linear Discriminant) and LFA (Local Feature Analysis).

Our approach in finding facial landmarks is based on elastic graph matching [8] [9]. A probed image graph is match against face model. The face model is represented as a graph with its nodes encoding the landmark points of interest and edges encoding the geometry of the face. Each node represents the encoded landmark point by the local Gabor filter responses called a *jet*. The geometry (edges) needs to be flexible to account for individual variations in geometry and distortions due to rotation in depth or active expressions. This flexibility is incorporated by making the graph elastic.

The graph matching process in [9] uses a coarse-to-fine approach with the degrees of freedom of the face graph progressively: translation, scale, aspect ratio and finally local distortion. This matching process however is computationally expensive due to its high searching space. Here, we introduce an alternative matching process by using a Genetic Algorithm (GA) to optimized the matching criteria. The use of GA considerably speeds up the finding of the facial landmarks. The search we allow encodes for translation,scale and rotation of the face graph. The feasibility of our methodology has been tested using the ORL face image database. The use of GA in feature detection is not new, for example in [11] GA has been used for eye detection by fitting image distribution (mean, entropy and standard deviation) between training and probed image. Here however, we combine the GA search strategy with the Gabor representation that has a more optimal encoding of the face landmark points [10].

The remainder of this paper is organized as follows. The Gabor feature extraction and face graph representation are presented in section 2. The graph matching process based on GA is discussed in section 3. In section 4 our experiment results on the ORL face database are presented. Finally,a conclusion and directions for future work are briefly covered in the last section.

2. Gabor Feature Extraction and Face Representation

A 2-D Gabor filter is obtained by modulating a 2D sine wave (at particular frequencies and orientations) with a Gaussian envelope. We follow the notation in [7]. The 2-D Gabor filter kernel is defined by

$$f(x,y,q_k,l) = \exp\left[-\frac{1}{2}\left\{\frac{(x\cos q_k + y\sin q_k)^2}{s_x^2} + \frac{(-x\sin q_k + y\cos q_k)^2}{s_y^2}\right\}\right] .\exp\left\{\frac{2p(x\cos q_k + y\sin q_k)}{l} i\right\} \quad (1)$$

where s_x and s_y are the standard deviations of the Gaussian envelope along the x and y-dimensions, respectively. l and q_k are the wavelength and orientation, respectively. The spread of the Gaussian envelope is defined using the wavelength l. Here we set $s_x = s_y = \lambda/2$. A rotation of the $x-y$ plane by an angle q_k result in a Gabor filter at orientation q_k. q_k is defined by

$$q_k = \frac{p}{n}(k-1) \qquad k = 1,2,..,n \quad (2)$$

where n denotes the number of orientations. Here we use eight values of orientation (n=8) and four values for wavelength (l=3,5,7,10). The Gabor local feature at a point (X,Y) of an image can be viewed as the response of all different Gabor filters located at that point. A filter response is obtained by convolving the filter kernel (with specific l, q_k) with the image. For sampling point (X,Y), this response, denoted as g(.), is defined as:

$$g(X,Y,q_k,l) = \sum_{x=-X}^{N-X-1}\sum_{y=-Y}^{N-Y-1} I(X+x,Y+y)\, f(x,y,q_k,l) \quad (3)$$

where $I(x,y)$ denotes an NxN greyscale image. When we apply all Gabor filters at multiple frequencies (l) and orientations (q_k) at a specific point (X,Y) we thus get 32 (four frequencies x eight orientations) filter responses for that point. This 32 valued vector is denoted as a Gabor jet. A jet J is defined as the set $\{J_j\}$ of 32 complex coefficients obtained from one image point, and can be written as

$$J_j = a_j \exp(i f_j) \quad j=1,..,4 \quad (4)$$

where a_j is magnitude and f_j is phase of Gabor features/coefficients.

Each landmark point can thus be represented by a Gabor jet. In this paper, the face is represented by the landmark points: center of eyes, nose and mouth, see also figure 1. Using the Gabor representation, a face is thus represented by four jets (one jet, 32 Gabor responses, for each of the four landmark points). This representation can also be represented as a graph. Then the nodes (p1,..,p4) represent the landmark points and are labeled with the corresponding jet. Edges (e1,e2,e3) represent topographical information between the landmark points and labeled with the distances between the points. Because we use 4 landmark points on the face, the jet set produces 128 Gabor filter responses per face graph.

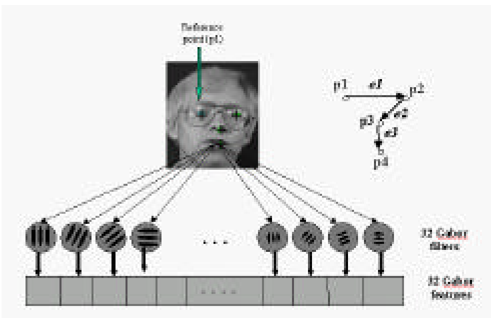

Figure 1. Extracted Gabor features.

The Face model graph is generated manually from 5 sample face images taken from the database. First, we manually mark 4 landmark points (p1,p2,p3 and p4) in the sample image (i.e. eyes, nose and mouth). Then the edges (e1,e2 and e3) are drawn between the marked landmark points and edge labels are computed as the difference between the node positions. Finally, the Gabor features are computed to provide jets of the node. This is done for all 5 sample images and the face model graph is then computed as the average of all sample graphs, i.e. for each node the average jet is calculated and for every edge the average distance is calculated.

To find the landmark points in a given unknown face image, we apply an elastic graph matching procedure on the probe image that maximizes the similarity between the graph representation of the probe image and the model graph (also reported in [9]) i.e.

$$\max_{\forall J'} S_a(J, J') \qquad (5)$$

where J is model jets and J' is jet of a probe face image. for the definition of similarity function $S_a(J, J')$ we use the definition of [9] as follows:

$$S_a(J, J') = \frac{\sum_j a_j a_j'}{\sqrt{\sum_j a_j^2 \sum_j a_j'^2}} \qquad (6)$$

The model graph has to be matched against probe image in an elastic way such that variations of graph translation, scaling and rotation are properly accounted for. In the next section, we present a matching procedure approach through evolutionary computing.

3. Graph Matching by Genetic Algorithm

In the previous section we showed that finding the landmark points in an unknown probe image can be achieved by solving Eq. 5 for all possible combination of jets. To cope for scale and rotation variations we even need to maximize over several scaled and rotated representations of the probe image. In this section we propose a strategy that overcomes this exhaustive searching and maintaining the invariance against scale and orientation. The strategy is based upon the evolutionary method of a genetic algorithm (GA). A GA method samples the facial landscape to detect salient locations for possible face landmark locations. It 'remember' only those salient areas whose Gabor features have a high match (i.e. similarity) with the face model graph. It then uses these high fitness areas as a starting point for sampling new areas until the algorithm converges.

GA encodes each point in the parameter (or solution) space into a binary bit string called a chromosome, and each point is associated with a 'fitness' value that is usually equal to the objective function evaluated at the point. In our case the objective function is equal to the similarity function between model graph and a given probe image graph (Eq. 6). For our problem, the solution is represented by only 4 parameters. These are 1,2) the position of the probe graph (we have taken the x and y coordinate of the left eye) reference 3) the scaling factor, and 4) rotation angle. Note that the positions of all landmark points other than the left eye can be determined from the left eye position using the face model graph in conjunction with the scale and orientation parameters.

Each parameter is coded into an 8-bit binary string (Figure 2). Our GA keeps a set of points as a population (or gene pool), which is then evolved repeatedly towards a better overall fitness value. In each generation, the GA constructs a new population using genetic operators such as crossover and mutation (Figure 3). Only population members that have high fitness values because they are more likely to survive, participate in crossover (mating) operations. After a number of generations, the population contains members with better fitness values. Note that the GA improve performance by upgrading entire populations rather than individual members. In our implementation, we choose to keep only the best 2 members when each new population is generated for making sure that the fitness value monotonically increased (elitism principle).

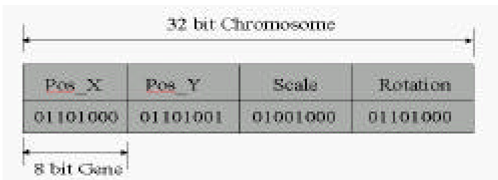

Figure 2. Chromosome format.

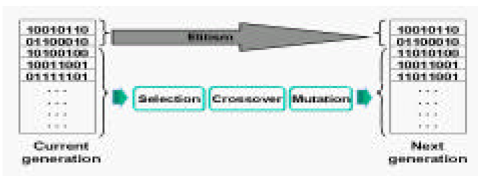

Figure 3. Producing the next generation in GAs.

Based on the above concepts, a simple GA algorithm for maximizing the similarity between the given probe image and the model graph is described as follows [12], also illustrated in Figure 3:

1. Initialization

Initialize a population with randomly generated individuals i.e. random values for the 4 parameters: position (x,y) of reference point, scaling factor and rotation angle. Evaluate the fitness value of each individual by evaluating the similarity function between the probe image graph denoted by the parameter of the individual and the stored face model graph.

2. Generate New Solution

a. Keep the best 2 individuals.
b. Select two members from the population with probabilities proportional to their fitness values.
c. Apply crossover with the probability equal to the predefined crossover rate.
d. Apply mutation with the probability equal to the predefined mutation rate.
e. Repeat (b) to (e) until enough members are generated to form the next generation.

3. Evaluate Solution

Repeat step 2 and 3 until a stopping criteria is met. The stopping criteria is met if the fitness does not improved within 4 generations or the number of generation has been exceeded a certain number (here we have used 15).

After the GA searching is converged, it provides us with values for the left-eye/reference position, scaling factor and rotation angle. These values in conjunction with the stored model graph's structure (p1,..,p4 and e1,e2,e3) results in the positions of the landmark points of the probe image that we were interested in.

4. Experiment Results

We have used the ORL database which contains a set of faces taken between April 1992 and April 1994 at the Olivetti Research Laboratory in Cambridge, UK [13]. There are 10 different images of 40 distinct subjects. For some of the subjects, the images were taken at different times. There are variations in facial expression (open/closed eyes, smiling/non-smiling), and facial details (glasses/no-glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. There is some variation in scale of up to about 10%. The images are greyscale with a resolution of 112 x 92. In addition we have used scaled version (75%) of these images to demonstrate the scaling invariance of the proposed landmark detector.

For the GA we used a constant population size of 20 individuals at each generation, the crossover rate and mutation rate are 1.0 and 0.01, respectively. The search space for each of the parameters was additionally constraint by the following boundaries:

- left-eye/reference point position X=[1 .. 55], Y=[1 ..70]
- scaling factor=[60% .. 110%]
- rotation angle=[-20° .. +20°]

Figure 4 shows same examples of the results of the detected facial landmarks. Figure 5 shows the fitness of the population for one of the experiments.

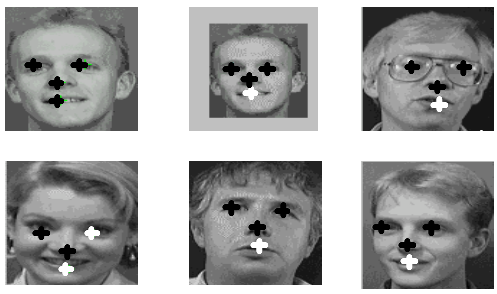

Figure 4. Facial Landmaks Finding

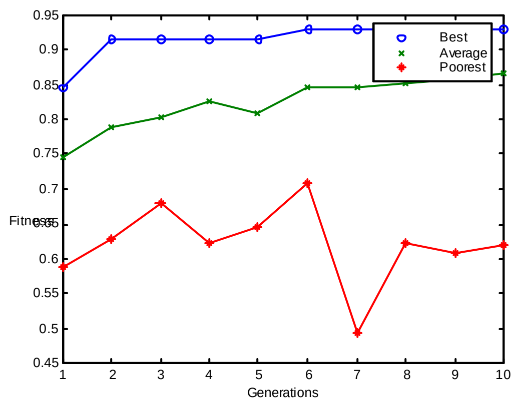

Figure 5. Fitness Function

5. Conclusion

We have proposed a detection scheme for locating facial landmarks based on graph matching using a genetic algorithm as optimization strategy. The performance of the proposed method was demonstrated on the ORL face database. The results are quite promising for frontal pose faces with small rotation and tilting. This method should be further tuned by introducing local distortions on each of the nodes of the graph for achieving even better results by giving more flexibility to the graph topograhical structure. From the results of the experiment, we conclude that the proposed method has a good prospect and should be considered in the design of face recognition systems.

References

[1]. Moghaddam, B. and Pentland, A., Face Recognition Using View-Based and Modular Eigenspaces, In *Automatic Systems for the Identification and Inspection of Humans, SPIE*, volume 2257, 1994.

[2]. Turk, M. and Pentland, A., Eigenfaces for recognition, *J. of Cognitive Neuroscience*, vol.3, pp.71–86, 1991.

[3]. Jain, L.C. et al. (eds.), Intelligent Biometric Techniques in Fingerprint and Face Recognition, CRC Press, NJ, 1999.

[4]. Jones, J. and Palemer, L., An Evaluation of the Two dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex, *Journal Neurophisiology*, vol. 58, pp.1233-1258, 1987.

[5]. Tunner, M.R., Texture Discrimintation by Gabor Functions, *Biological Cybernetics*, Vol. 55, pp.71-82, 1986.

[6]. Hammamoto, Y., Uchimura, S., Masamizu, K. and Tomita, S., Recognition of Handprinted Chinese Characters using Gabor Features, *in Proc. Of th 3rd Intl. Conf. Document Analysis and Recognition*, pp.819-823, 1995.

[7]. Hamamoto, Y., A Gabor Filter-based Method for Fingerprint Identification, *Intelligent Biometric Techniques in Fingerprint and Face Recognition, eds. L.C. Jain et al.*, CRC Press, NJ, pp.137-151, 1999.

[8]. Lades, L., J. Vorbruggen, Buhmann J., Lange, J., von der Malsburg, C., and Wurtz, R., Distortion Invariant Object Recognition in the Dynamic Link Architecture, *IEEE Trans. Computers*, Vol. 42, pp.300-311, 1993, 1993.

[9]. Wiskott, L., Fellous, J.M., Kruger, N., and von der Malsburg, C., Face Recognition by Elastic Bunch Graph Mathcing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), pp.775-779, 1997.

[10]. Donato, G., Bartlett, M.S., Hager, J.C., Ekman, P., and Sejnowski, T.J., Classifying Facial Actions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10), pp.974-989, 1999.

[11]. Huang, J., Liu, C., and Wechsler, H., Eye Detection and Face Recognition using Evolutionary Computation, *in Face Recognition: from Theory to Applications, eds. Harry Wechsler et al*. Springer-Verlag, pp.348-377, 1997.

[12]. Jang, J.S.R., Sun, C.T., and Mizutani, E., Neuro-Fuzzy and Soft Computing: A Computational approach to Learning and Machine Intelligence, Prentice-Hall, Inc. NJ, 1997.

[13]. ORL website, http://www.cam-orl.co.uk/facedatabase.html