# Facial Feature Extraction Using Genetic Algorithm

Gary G. Yen Nethrie Nithianandan
Intelligent Systems and Control Laboratory
School of Electrical and Computer Engineering
Oklahoma State University, Stillwater, OK 74074-5032, USA

**Abstract- An automatic facial feature extraction method is presented in this paper. The method is based on the edge density distribution of the image. In the preprocessing stage a face is approximated to an ellipse, and genetic algorithm is applied to search for the best ellipse region match. In the feature extraction stage, genetic algorithm is applied to extract the facial features, such as the eyes, nose and mouth, in the predefined sub regions. The simulation results validates that the proposed method is capable of automatically extracting features from various video images effectively under natural lighting environments and in the presence of certain amount of artificial noise and of multi-face oriented with angles.**

## I. INTRODUCTION

Modern multimedia technology has led to the growing demands of image and video applications in medicine, remote sensing, security, entertainment and education. Yet, effective image feature extraction, often regarded as a critical component in multimedia information processing, is not well attended. In order to efficiently retrieve an interested image in a large data base, it should be described by some keywords or automatically extracted visual features. However, in the case of a large image data base it will not be feasible to describe an image with keywords. This not only requires extensive labor, but also insufficient amounts of keywords exists to effectively characterize the contents of the image. Therefore automatic feature extraction and indexing of video images has been in a great demand. An efficient and effective image feature extraction algorithm, if successfully developed, can benefit greatly to emerging imagery applications such as biometric [1], medical diagnosis [2] and geographical information system [3], to cite a a few.

Facial feature extraction plays an important role in law enforcement forensic investigation [4], low bit video coding [5], and security access control systems [6]. State-of-the-art multimedia teleconferencing systems are based on digital video coding and transmission of television pictures. To better use the low bandwidth digital networks for this purpose, significant compression of data is mandatory. Images can be considered as having structural features such as contours and regions. These images have been exploited to encode images at low bit rates, while retaining sufficient visible structures in the reconstruction so as to maintain an acceptable level of quality [7].

Smart cards are widely used today for security access to public buildings, military installments, and computer networks. This technology is considered to be the most effective, efficient and economic means of ensuring security access. These cards store unique biometric informatics such as the fingerprint, signature, retinal pattern, voice recognition and facial features [8]. An effective facial feature extraction algorithm in this regard is considered to be an open area for further research. When fused with all matured identification approaches, a global, multi-functional smart card system can then become viable.

Due to the most recent terrorist attacks, the need of enhanced recognition technology to support automatic surveillance is becoming more important according to USA Today [9]. Surveillance cameras are fixed on the subways, city centers, parks, shopping malls, buses, and even in the historic rural villages in Britain. It is proven that this kind of security measures will not only reduce the percentage of burglary and shop lifting but also help to identify the culprits on-the-run. It is predicted that thousands of such surveillance systems will be installed across the USA at airports, train stations, stadiums, public monuments, ATM machines, and even in private businesses as a result of the September 11th attack on the World Trade Center. Yet, proven automatic facial recognition technology is not available today. Existing surveillance systems solely depend upon the human-in-the-loop process can be easily defeated.

Many facial feature extraction algorithms have been proposed in the past. They can be mainly classified into three categories, namely feature based [4], appearance based [10-11] and template based [12] approaches. Each has its advantages and limitations on its own right with respect to locating the face region, computing the cost functions and extracting the features. However, all approaches demand a fairly sophisticated model, which is often not readily available.

Developing a computational model for face recognition is quite challenging, if not at all impossible, because faces are complex and multidimensional. In the *feature based approaches*, the features are extracted based on the geometric relationships such as the positions and width of the features. The features are extracted from the vertical and horizontal integral projections of the original

0-7803-7282-4/02/$10.00 ©2002 IEEE

image. The horizontal edge map is used to extract the left and right boundaries of the face and the nose, while the vertical edge map is used to extract the eyes, mouth and nose base [10]. The peaks and valleys of the horizontal and vertical projections are analyzed with respect to a threshold to detect and extract the positions of the features. The selection of the threshold value is often sensitive to the performance achieved. Even though feature based approaches need limited memory, they are very difficult to implement, assuming an extensive knowledge beforehand.

The principal component analysis is a statistical approach used in *appearance based approaches* to extract facial features for recognition [10]. This approach transforms face images into a small set of characteristic feature images called “eigenfaces”, which are principal components of the initial training set of face images [13]. Eigenfaces are nothing but a set of orthonormal basis vectors. Each of these basis vectors can be displayed as a ghostly face; often referred to as an eigenface [14]. Concepts of eigenfaces can be extended to eigenfeatures, such as eigeneye, eigenmouth and eigennose. These “eigenfeatures” are used for the detection of features such as eyes, mouth and nose. In the eigenfeature representation the equivalent distance from the feature space is effectively used for detection of features. In case of a new input image, the distance from the feature space is computed at each pixel, and the minimum of the distance map is considered as the best match. However, eigenfeature approaches commonly assume that the images in low dimension cannot scale up properly.

*Template based approaches* use single or multiple templates represented by specifically designed energy functions for feature extraction. They are simple to apply, but they suffer from insufficient matching capability and large memory requirements. Additionally, it cannot deal with complicated background settings. Some feature extraction techniques use deformable templates to extract facial features [15]. These are flexible templates constructed with *a priori* knowledge of the shape and size of the different features [16]. The templates can change their size and shape so that they can match properly. Each of the templates has been evaluated by an energy function, defined in terms of peaks and valleys of the image intensity, edges and the intensity itself. The minimum value of the energy function corresponds to the best fit with the image. These methods work well in detection of the eyes and mouth, despite variations in tilt, scale, and rotation of head. However, modeling of the nose and eyebrow was always a difficult task [15-16].

The normal process of searching for the features is computationally expensive, therefore genetic algorithm (GA) is used as a search algorithm. Genetic algorithm possesses the following features that make them better suited than those traditional search algorithms [17-18].

- GA works with a coding of parameter sets and not with the parameters themselves.
- GA searches from a population of candidate solutions, not a single point.
- GA uses objective function information, not derivative or other auxiliary information
- GA uses probabilistic transition rules, not deterministic rules.
- The result is a population of solutions instead of an individual solution.

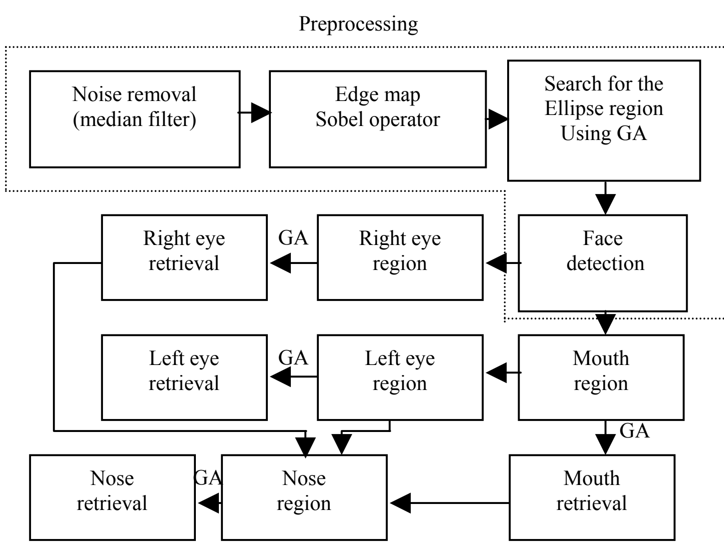

Figure 1. Block diagram of proposed feature extraction method

Figure 1 shows an overview of the process involved in extracting the facial features. The face segmentation is a critical part in the feature extraction stage. Even though there are many face segmentation algorithms based either on the

RGB color model or HSV color model, the one open problem of color information based face location is that when there are non face skin color regions, such as when hands and shoulders appeared in the background, it is difficult to locate the face by the skin color information. A method that approximates the face region to the ellipsoid is used to detect the face [19] in the preprocessing stage. The proposed feature extraction method is based on the edge density distribution using templates. In the features extraction stage, the main features such as the eye-eyebrow, nose and mouth are extracted from the face region. The features are extracted on the basis of a predefined fitness function.

The organization of the paper is structured as follows. Section II presents the proposed genetic algorithm based facial feature extraction. Simulation results are used to validate the effectiveness of the proposed approach in Section III. Section IV provides the conclusion with some pertinent observations.

#### II. FACE SEGMENTATION AND FEATURE EXTRACTION

All the images captured were head and shoulder images and in a frontal view. Smoothing filters are used for noise reduction. The median filter is used for this purpose. The face segmentation process is proceeded under the assumption that the face region can be approximated by an ellipsoid. This method works well even under the environments when the background is complex and the face contains extra features such as spectacles, beard and etc. Locating faces in images containing skin color parts such as hands and shoulders poses no serious complications. Therefore each chromosome in the population during the evolutionary search has five parameters genes, the center of the ellipse (*x* and *y*), x directional radius ($r_x$), y directional radius ($r_y$) and the angle ($\theta$). The chromosome in binary form for each parameter is coded as shown in Figure 2. The use of integer coding provides no appreciable performance improvement in simulations conducted in Section 3.

| | | | | |
|---|---|---|---|---|
| x-8bits | y-8bits | $r_x$-8bits | $r_y$-.8bits | $\theta$-7bits |

Figure 2. Chromosome for face segmentation

The fitness function is defined by the number of edge pixels in the approximated ellipse like face to the actual number of pixels in the actual ellipse. The ratio is large when both ellipses overlap perfectly. It is commonly assumed in literature that the ratio of the length to breath of the face is 1.5: 1, and therefore the same ratio is used to obtain the face area once the ellipse region is located. In the case of multiple faces in the image, faces are located until a threshold is satisfied. The threshold is based on the fitness value used to locate the faces. Once the face is detected, the fitness value in that region is assumed to be zero in the succeeding search for additional face regions, if any. The images used for feature extraction are of the size 300 by 200. The feature extraction algorithm proposed is based on horizontal edge density distribution. The horizontal edge map of the image from the segmented image is obtained in order to extract facial features. It can be observed from Figure 3 that the horizontal edge density distribution around the feature point area is very high.

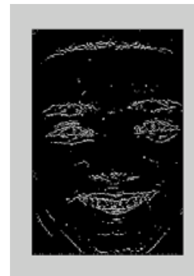

Figure 3. Edge map of the segmented image

The sizes of the templates for different features are decided according to general knowledge of the size of the features. Both the eye and eyebrow are contained in the same template. Rectangle templates of different sizes for different facial features are used. To make the search process less computationally expensive, feature extractions can be carried out in different regions of the face. Basically, the face region can be divided into three sub regions as shown in Figure 4. The sub region $R_r$ is expected to contain the right eye, the left eye is expected to be located in $R_l$ and the mouth in $R_m$. It is possible for the mouth region to contain the nose. This problem can be easily overcome by evaluating the edge density projection of the mouth region. By analyzing the peak and valley, the mouth region can be identified. The nose region ($R_n$) is obtained once the eyes and mouth are located. The search region for the nose lies between the eyes and mouth.

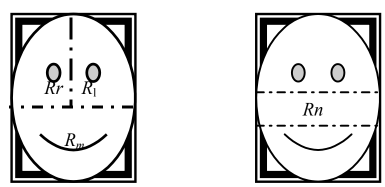


Figure 4. Sub regions of the face

In the feature extraction stage, genetic algorithm is used to search for the global maximum point when the template best matches the feature. The chromosome represents the position of the feature in the *x* and *y* direction.

| | |
|---|---|
| x-direction (7bits) | y-direction (7bits) |

Figure 5. Chromosome for face feature extraction

The fitness is evaluated in terms of the density of the template. The best template is selected when the fitness is maximized. The fitness function *F* is shown below,

0-7803-7282-4/02/$10.00 ©2002 IEEE

$$F = \frac{1}{m \times n} \sum_{x=1}^{m} \sum_{y=1}^{n} T(x, y)$$

$$\text{where } \begin{cases} T(x, y) = 1 & \text{if the pixel is white} \\ T(x, y) = 0 & \text{if the pixel is black} \end{cases},$$

and $T$ is the template, $(x, y)$ are the coordinates of the template, and $m \times n$ is the size of the template. Initially the population is chosen randomly. In each generation 20 percent of the population is considered for the reproduction. The Roulette Wheel selection scheme is applied in the selection process. In experiment, more advanced genetic selection algorithms, such as tournament selection, do not provide much improved performance. A 2-point crossover and mutation is applied by choosing the position at random without duplication. A small percentage (i.e., one percent) of individuals are inserted in every generation.

### III. SIMULATION STUDY

A set of testing images was captured for the experiment. The images were taken with a Kodak Digital Science DC260 zoom camera with a resolution of 1024 by 1536 pixels. The data consists of more than 30 different head and shoulder images with at least two images per person. Some of the testing images contained multiple faces and faces oriented to an angle. There were no auxiliary lighting used when the images were captured. To evaluate the robustness of the algorithm, images with different facial expressions, glasses and complex backgrounds were tested. Artificial noise was injected to test how the algorithm performs in the presence of noise. The images used for the feature extraction are of a size 300 by 200. The edge map obtained is using the Sobel approximation with a threshold value of 0.02.

The GA parameter setting used for both face segmentation and feature extraction in the simulation is shown in Table 1. The stopping criterion for face segmentation is 150 generations. In most cases, the face was located in less than 100 generations, accounted for approximately one second in a Pentium 500 computer running Matlab 6.0. In the process of feature extraction, the stopping criterion was set to be 100 generations. Generally, the eyes and nose were located in less than 50 generations, and the mouth was located in less than 75 generations. Figure 6 shows the located face regions, while Figure 7 displays the extracted features of the located faces. Note the face on the right is oriented to an angle, while the face in the middle shows non-nominal expression. Figures 8-12 show successful face segmentation with located features under different situations, such as complex background, various facial expressions, with glasses, with artificial noise and oriented to an angle. In all cases when the head and shoulder images are properly captured, the system developed successfully performs the face segmentation and follow-up feature extraction at 100% rate. On the other hands, existing template based approaches such as [15-16] are unable to deal with the facial images with glasses, abnormal expression and oriented to an angle, simply because there is no proven model readily available. To thoroughly validate the developed algorithm, a set of benchmark testing images should be made publicly available. The benchmark testing set should be carefully designed to exploit various challenging characteristics in real-life environments.

Table 1. GA parameters

| | Face segmentation | Feature extraction |
|---|---|---|
| Population | 100 | 50 |
| Crossover ($P_c$) | 0.8 | 0.8 |
| Mutation ($P_m$) | 0.001 | 0.001 |

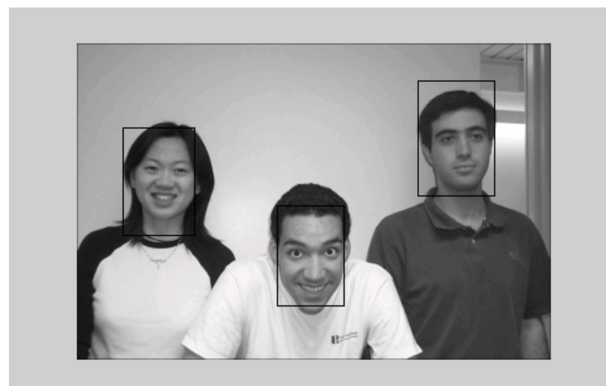

Figure 6. Image with multi-face showed located face regions

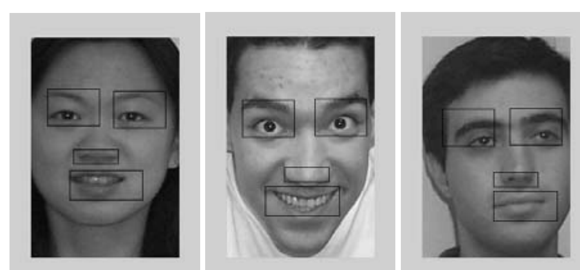

Figure 7. Facial features extracted from Figure 6

Figure 8. Feature extraction of face with complex background

0-7803-7282-4/02/$10.00 ©2002 IEEE

Figure 9. Feature extractions of faces with different facial expression

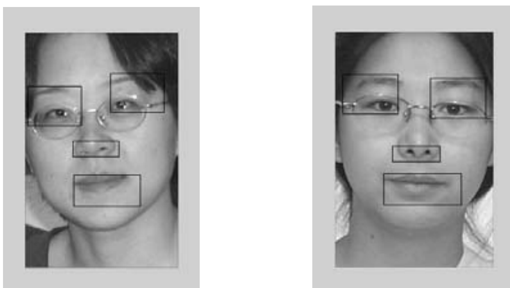

Figure 10. Feature extractions of faces with glasses

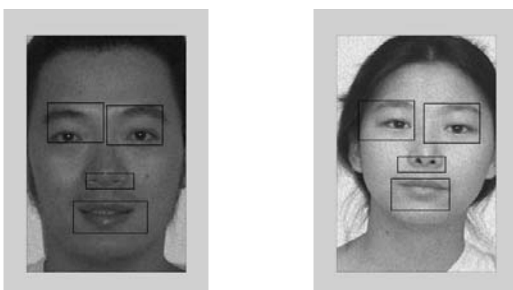

Figure 11. Feature extractions of faces with injected noise

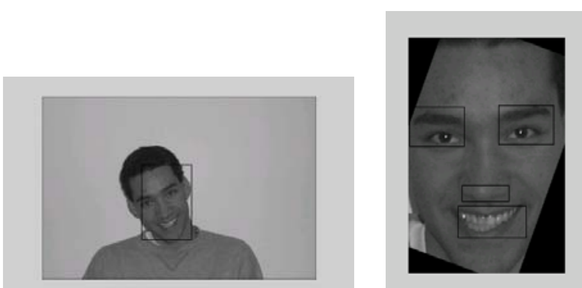

Figure 12. Feature extractions of faces oriented at an angle

## IV. CONCLUSION

An automatic facial feature extraction from video images has been developed in this paper. In the preprocessing stage the face is located by approximating face to ellipsoid. The facial features are extracted from horizontal edge map of the image. Different sizes of templates are defined for different facial features and the features are searched in the predefined sub regions of the image. The cost function is based on edge density distribution.

The proposed facial feature extraction approach has been validated with a large number of images. Some of the images contained more than one person, while others with person oriented at an angle. Simulation results show that the facial features were extracted successfully. Genetic algorithm was able to search effectively and reduce computational complexity, therefore reduce the search time. The facial features were extracted even in the presence of artificial noise. The algorithm can be improved so that it can be applied to the real world problems, by incorporating more characteristics in the fitness function during the evolutionary process and also to extract features on faces containing either beard or mustache.

## REFERENCES


1. Chellappa, R., Wilson, C.L. and Sirohey, S., “Human and machine recognition of faces: a survey,” *Proceedings of the IEEE*, Vol. 83, No. 5, pp. 705 -741, 1995.
2. Li, H. and Chutatape, O., “Fundus image features extraction,” *Proceeding of the IEEE International Conference on Engineering in Medicine and Biology*, pp. 3071-3073, 2000.
3. Fitton, N. and Cox, S., “Linear feature extraction in geoscientific data,” *Proceedings of Digital Image Computing Techniques and Applications Conference*, pp. 104-109, 1995.
4. Brunelli, R. and Poggio, T., “Face recognition: features versus templates,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 15, No. 10, pp. 1042 -1052, 1993.
5. Chuang, M., Chang, R. and Huang, L., “Automatic facial feature extraction model based coding,” *Journal of Information Science and Engineering*, Vol. 16, pp. 447-458, 2000.
6. Lin, C. and Ling, W., “Automatic facial feature extraction by genetic algorithms,” *IEEE Transactions on Image Processing*, Vol. 8, No. 6, pp. 834-845, 1999.
7. Aizawa, K. and Huang, T.S., “Model based image coding: advanced video coding techniques for low bit-rate applications,” *Proceedings of the IEEE,* Vol. 83, No. 2, pp. 259-271, 1995.
8. Henry, V., “Biometrics: face recognition technology,” in http://www.sans.org/infosecFaQ/authentic/face_rec.htm.
9. USA Today, “America on alert: many Britons smile on spread of candid cameras,” *USA Today*, September 26, 2001, pp. 5.
10. Shah, D. and Marshall, S., “Statistical coding method for facial feature,” *IEE Proceedings of the Visual Image Signal Processing,* Vol. 145, pp. 187-192, 1998.
11. King, I. and Xu, L., “Localized principal component analysis learning for face extraction and recognition,” *Proceedings of Workshop on 3D Computer Vision,* pp. 124 -128, 1997.


0-7803-7282-4/02/$10.00 ©2002 IEEE

12. Baron, R.J., “Mechanisms of human facial recognition,” *International Journal on Man Machine Studies,* Vol. 15, pp 137-178, 1981.
13. Turk, M. and Pentland, A., “ Eigenfaces for recognition,” *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
14. Wu, H., Chen, Q. and Yachida, M., “Feature extraction and face verification,” *Proceedings of the IEEE International Conference on Pattern Recognition*, pp. 484-488, 1996.
15. Yullie, A.L., Cohen, D.S. and Hallinan, P.W., “Feature extraction from faces using deformable templates,” *Proceedings of the IEEE International Conference on Pattern Recognition,* pp. 104-109, 1989.
16. Huang, L.C. and Chen, C.W., “Human face feature extraction for face interpretation and recognition,” *Proceedings of the IEEE International Conference on Pattern Recognition*, pp. 204-207, 1996.
17. Goldberg, D.E., *Genetic Algorithm in Search, Optimization, and Machine Learning*, Reading, M.A: Addison-Wesley, 1989.
18. Tang, K.S., Man, K.F., Kwong, S. and He, Q., **“** Genetic algorithms and their applications,” *IEEE Signal Processing Magazine,* Vol. 13, No. 6, pp. 22 -37, 1996.
19. Yokoo, Y. and Hagiwara, M., “Human face detection method using Genetic algorithm,” *Proceedings of the IEEE Congress on Evolutionary Computation,* pp.113-118, 1996.


0-7803-7282-4/02/$10.00 ©2002 IEEE