Detection of facial features based on the relaxation algorithm

Ho-Jin Lee, Dong-Gyu Sim, and Rae-Hong Park

Dept. of Electronic Eng., Sogang Univ.
C.P.O. Box 1142, Seoul 100-611, Korea

ABSTRACT

This paper proposes a relaxation-based algorithm for detection of facial features such as face outline, eyes, and mouth. At first, a number of candidates for each facial feature are detected. To select a correct set of facial features from the candidates, probabilities and geometric relationships of each candidate location are considered, in which a relaxation algorithm is used for implementation. Simulation results with various test images are presented.

Keywords: Relaxation, Facial features, Geometric relationship, Probability.

1. INTRODUCTION

Detecting human facial features is one of important issues in computer vision. It can be applied to face recognition and video communication. Man can easily distinguish human faces by means of their features, however, it is not easy to implement this human ability into a machine.

Many algorithms have been proposed to extract human facial features. Lee *et al.*[1] used horizontally and vertically projected edge maps. Eleftheriadis and Jacquin [2] used shape information and symmetric property of binary edges of a facial image. Kothari and Mitchell's algorithm[3] detected eye locations using the optical flow gradient vectors. Yuille *et al.*[4] used deformable templetes to describe shapes and locations of eyes and mouth. Deformable templates can accurately describe them, however deciding an initial point to fit templates to facial input image is a difficult problem and severely affects the overall performance. So an additional algorithm to effectively detect their locations is required. Lam and Yan [5] extracted corner points of eyes and mouth for face recognition. They used the coner angles of eyelid (lip) lines, orientation of the coner, and dissimilarity of gray levels between inner region and outer region of the corner. This scheme is based on that proposed by Xie *et al.*[6]. Jeng *et al.*[7] detected facial features using the geometrical face model with the relative distances of each feature. The most important feature to be detected first is the *base line* which passes through the centers of both eyes. After finding it, others are detected sequentially by minimizing the defined energy functions. Lin and Wu[8] detected facial features using a feature template and cost function at the preprocessing stage of their algorithm. They assumed that there are high frequency signals around each feature point. They defined feature searching regions using a region growing method and *a priori* knowledge about human face. The location at which the cost function has the largest value is detected and considered as a feature.

This paper proposes a relaxation-based method for detection of facial features, in which face outline is fitted to an ellipse. Ellipse candidates for face outline are detected by the fitting algorithm. Candidates for eye locations and mouth location are detected by bin incrementing and projection profile, respectively. Finally, the most desirable set of facial features is selected by the proposed relaxation algorithm.

2. PROPOSED RELAXATION-BASED ALGORITHM FOR DETECTION OF FACIAL FEATURES

It is difficult for machines to detect the human facial features as man does. However, if a system utilizes *a priori* information of the facial image, the performance of the system can be improved. So, the proposed algorithm considers geometrical relationships of facial features with appropriate constraints.

The proposed algorithm consists of three parts. The first part detects candidates for primitive facial features, the second part detects candidates for additional facial features, and the last part selects the correct feature set

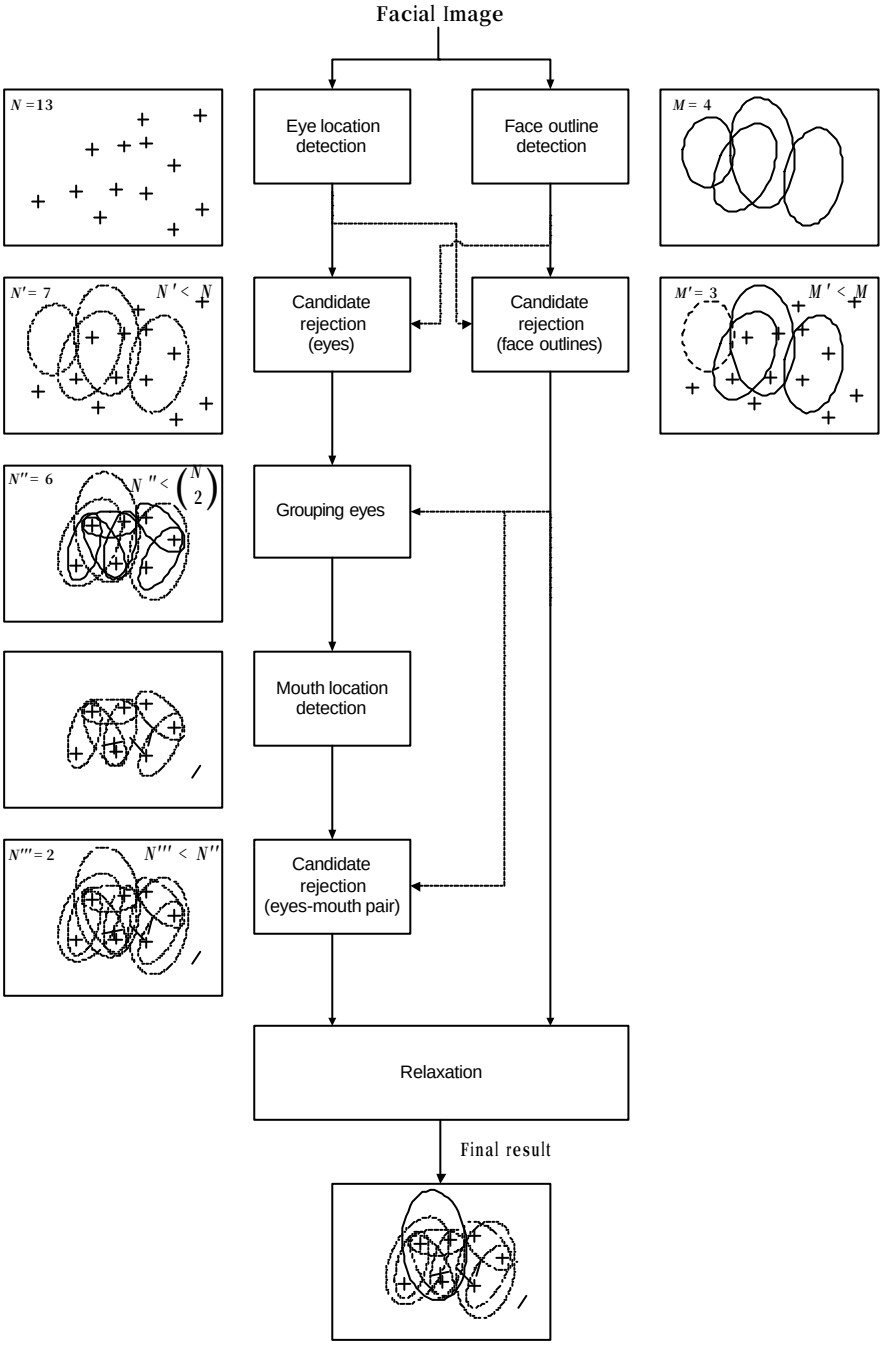


Figure 1. Block diagram of the proposed system.

among the candidates using the relaxation algorithm, where primitive facial features represent visually dominant characteristics among various features. Thus they can be detected by using their characteristics in a facial image. Additional facial features are secondary features that can be easily detected with or derived from the information of primitive facial features. Fig. 1 shows the block diagram of whole system.

In the proposed algorithm, face outline and eye locations are defined as primitive facial features, whereas a set of eye and mouth locations is regared as an additional facial feature. At the end of each detection part of primitive and additional facial features, candidate rejection steps reject inappropriate candidates to reduce the computational complexity.

2.1. Detection of primitive facial features

This part is composed of three steps: detection of candidates for face outline, detection of candidates for eye locations, and rejection of inappropriate candidates. The fitness of the ellipse candidate to a binary edge of an input facial image is considered for selection of the ellipse candidates[2] at the first step. The shape and contrast between candidate points and their neighborhood are considered to select the candidates of eye locations[3] at the second step. At the third step, the results of the first (second) step are used to confirm validity of the results of the second (first) step, and inappropriate candidates are rejected.

2.1.1. Ellipse candidates for face outline

Eleftheriadis and Jacquin[2] used the shape of a binary edge image as an important feature to select the candidates for face outline. An ellipse model was used for approximation of human face outline from binary edges. It has five parameters (a, b, x_c, y_c, θ) for describing its shape and position.

Fig. 2 shows the ellipse model with specified parameters: a center (x_c, y_c), the lengths a and b of minor and major axes, respectively, and a tilt-angle θ of the major axis with respect to the vertical direction. By changing the five parameters (a,b,x_c,y_c,θ) of an ellipse model and by considering the fitness of the ellipse model to a binary edge image, the ellipse candidates for face outline are selected. The model parameters yielding a large fitness measure R are selected[2]. N candidates for face outline are selected according to the fitness measure value $R(l)$, where l, $0 \le l \le N-1$, is an index of the face outline candidate.

2.1.2. Candidates for eye locations

To detect eye locations, Kothari and Mitchell[3] considered the fact that an eye region shows a large contrast and the shape of iris is ellipsoidal. Optical flow on the edge of iris radiates outwards. If lines are extrapolated along the opposite direction of the optical flow on the edge points of iris, lines intersect at a point. They defined *bin*s which

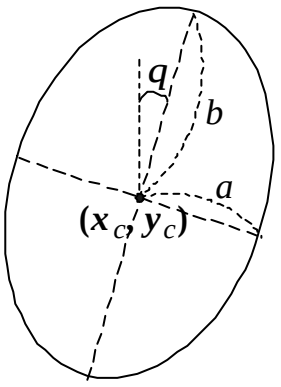


Figure 2. Ellipse model.

are accumulators in two-dimensional array. If a large number of lines pass through a pixel, the bin corresponding to this pixel has a large value, which is equal to the number of lines that pass through the pixel. If the value of a bin is large, the point corresponding to the bin usually satisfies the contrast and ellipsoidal conditions for eye location detection.

M pixels having large bin values are selected as the candidates for eye locations, and the fitness $B(l)$ is defined by the bin value, where l, $0 \le l \le M-1$, is an index of the eye location candidate.

2.1.3. Candidate rejection

After detection of primitive features, some candidates are rejected because they are redundant or topologically inappropriate. Rejection of these candidates reduces the computational complexity of the next parts and helps the selection of the correct feature set (one face outline, two eye locations, and one mouth location).

The rejection procedures are performed based on the fact that *eyes and mouth lie inside the face outline*. They reject candidates that do not satisfy the minimum requirements, which mean that a valid face outline candidate contains at least one eye candidate inside its region, and a valid eye candidate has at least one face outline candidate that encompasses the eye candidate. These minimum requirements are represented as follows: the m'th face outline candidate is valid if it satisfies the condition

$$\sum_{n=0}^{N-1} I_E^{m'}(x^n, y^n) > 0 \tag{1}$$

where I_E^m denotes the index function indicating whether the set of pixels is on (or inside) the mth face outline ellipse or not. The value of $I_E^m(x^n, y^n)$ is 1 if pixel (x^n, y^n) is inside the mth face outline ellipse, otherwise it is 0. The superscript n (m) represents the index of the eye (face outline) candidate and (x^n, y^n) denotes the position of the nth eye candidate. Similarly, the n'th eye candidate is valid if it satisfies the condition

$$\sum_{m=0}^{M-1} I_E^m(x^{n'}, y^{n'}) > 0. \tag{2}$$

These two inequalities check the minimum requiremenets for valid feature candidates. The positive value in Eq. (1) denotes that the m'th face outline ellipse encompasses at least one eye candidate whereas the positive value in Eq. (2) signifies that the n' th eye candidate has at least one face outline candidate that encompasses it.

2.2. Detection of additional facial features

The purposes of this part is to detect secondary facial features that are not detected in the primitive feature detection part, and to make the results of the previous part more reliable. For example, we detect mouth location and construct eye groups. This part is composed of three steps: generating eye groups from eye candidates, detecting mouth locations, and rejecting inappropriate candidates.

2.2.1. Grouping of eyes

The results from the previous part become more useful if some simple operation such as grouping is added. In this algorithm, among all the combinations of eye candidates, only some combinations can be chosen as eye groups. The constraint of grouping eyes is that *two eye candidates which have at least one common face outline candidate that encompasses them can be defined as a group.* After grouping, right and left eye candidates are defined in a group. In the next step, these groups are used as eye candidates. The eye group is a derived feature from eye primitive features.

2.2.2. Detection of mouth location

In human facial images, mouth has not visually dominant characteristics as eyes. With *a priori* information about a detected group of eyes, mouth location candidates can be detected more easily and reliably. In the proposed algorithm, *a priori* knowledge of geometrical relationships is used to detect mouth location candidates. The geometrical relationship that *mouth is on the perpendicular bisector line of the line segment which connects two*

eye points (right and left eyes) is assumed. To detect mouth location candidates, projection profiles of an edge map is used. In the proposed algorithm, Lin and Wu's method[8] is employed to define mouth region, in which mouth region is defined depending on the positions of and diatance between right and left eyes. Projection profile along the direction of eye-line (a line which passes through the two eye locations) is considered to detect mouth location candidates in the region R_c'. The vertical position at which the value of projection profile is the maximum is considered to be a mouth location candidate. In this scheme, one mouth location candidate is detected as a new group of eyes and added to the group of eyes.

2.2.3. Candidate rejection

After adding a mouth candidate to a candidate group of eyes, some inappropriate eye-mouth candidate groups are rejected. The assumption that *eyes and mouth lie inside the face outline* is used for candidate rejection. The assumption is similar to the one used for rejection of inappropriate eye location candidates. Similarly, the minimum requirement is represented as follows: the n'th eye-mouth candidate group is valid if it satisfies the condition

$$\sum_{m=0}^{M-1} I_E^m(x_r^{n'}, y_r^{n'})\, I_E^m(x_l^{n'}, y_l^{n'})\, I_E^m(x_m^{n'}, y_m^{n'}) > 0 \tag{3}$$

where the superscript n (m) signifies the index of the eye-mouth group (face outline) candidate. ($x_r^{n'}$, $y_r^{n'}$), ($x_l^{n'}$, $y_l^{n'}$), and ($x_m^{n'}$, $y_m^{n'}$) denote the positions of right eye, left eye, and mouth candidates of the n'th eye-mouth candidate group, respectively. The minimum requirement says that a valid eye-mouth candidate group has at least one face outline candidate that encompasses the group. The positive value in Eq. (3) denotes that the n'th eye-mouth candidate group has at least one face outline candidate encompassing all of its elements, which satisfies the minimum requirements.

2.3. Relaxation algorithm for detecting facial features

With the candidates of face outline and eye-mouth locations, a relaxation algorithm[9] is employed to select a correct set of face outline, eye locations, and mouth location. In the relaxation algorithm, each candidate (face outline and eye-mouth group) has the probability value supporting that it is a correct feature. The probability value is updated iteratively. If a candidate satisfies the relational condition for correct features, the probability of this candidate will be larger as iteration proceeds, in which relational condition denotes that there are other features satisfying the geometric relationship with this feature.

Fig. 3 shows the geometric relationships of the face model that are used for relaxation, where an ellipse (face outline) and three squares (eye locations and mouth location) signify facial elements of the model. d_0 and d_1 denote distances between model eyes and candidate eyes, whereas d_2 denotes the distance between model mouth and candidate mouth. Two dotted lines are major and minor axes of the ellipse. Filled boxes lie on the minor axis, and the distances from the center of ellipse to them are half the length of the minor axis. Empty box lies on the

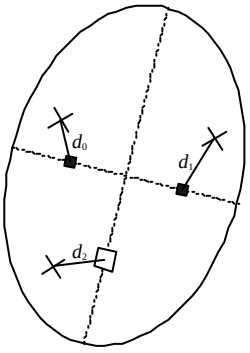


Figure 3. Face model.

major axis of the ellipse, and the distance from the center to the empty box is half the length of the major axis. In the relaxational iterations, if there is a candidate set for facial features (one face outline, two eyes, and one mouth) satisfying the geometric relationships, the probability for the set is large at the end of iterations.

Fitnesses R and B cannot be used directly for assigning the initial probability, because relaxation needs the probability constraint in adopting the compatibility measure of the candidates[9]. Thus, the initial probability is defined by

$$P_i^0(l) = \frac{F_i(l)}{\sum_m F_i(m)} \tag{4}$$

where $F_i(l)$ denotes the fitness of the lth candidate, the subscript i signifies the feature index, with i=0 and 1 representing the face outline and eye-mouth group, respectively, and the superscript denotes the iteration index. $F_0(l)$ and $F_1(l)$ are defined by $R(l)$ and $B(l)$, respectively. The probability at the $(k+1)$th iteration is given by

$$P_i^{k+1}(l) = \frac{P_i^k(l)[1+q_i^k(l)]}{\sum_m P_i^k(m)[1+q_i^k(m)]} \tag{5}$$

and the probability update term $q^k{}_i(l)$ at the kth iteration is represented by

$$q_i^k(l) = \sum_{j \neq i} w_j(l') P_j(l') \tag{6}$$

$$w_i(l) = \frac{c_i}{h_i + d_i(l)} \tag{7}$$

where P represents probability, w is a weight, d denotes a diatance, and h and c are constants.

3. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the proposed algorithm, computer simulations are performed. The Miss America sequence images and a number of face images from MIT face database† are used as test images. It is assumed that test images contain a single face image.

Fig. 4 shows an example of face outline detection result. Eleftheriadis and Jacquin's method[2] is used for detection and selection of best fitted ellipse. The best fitted ellipse is superimposed on the edge map in Fig. 4. The binary edge map is generated by series of operations: downsampling, Sobel edge operation, and Thresholding of the edge map. A downsampled image is desirable in the point of view that it is still contains global facial features and can greatly reduce the computational complexity. The downsampling factor is set to eight in the proposed algorithm.

Fig. 5 shows an example of bin incrementing result to detect eye location, where a test image is the first frame of the Miss America sequence. The darkest pixel represents the largest bin value. As mentioned above, if the value of a bin is large, the point corresponding to the bin usually satisfies the contrast and ellipsoidal conditions. In Fig. 5, the points that have large bin value locate around eye and mouth region.

In experiments, N=10 and M=30 are used, where N and M represent the initial numbers of candidates for face outline and eye locations, respectively. These candidates are chosen based on the fitness measure of each feature. After rejection of topologically inappropriate candidates, relaxation is employed to select a correct set of features.

† ftp://whitechapel.media.mit.edu.

Figure 4. An example of fitted ellipse.

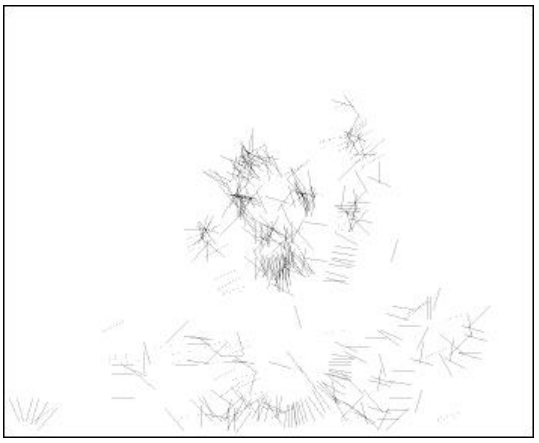
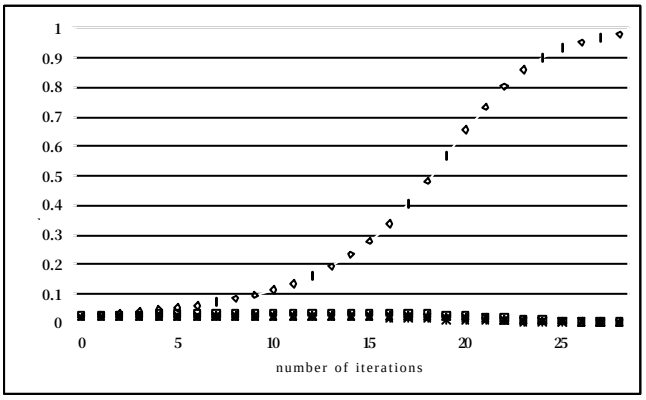
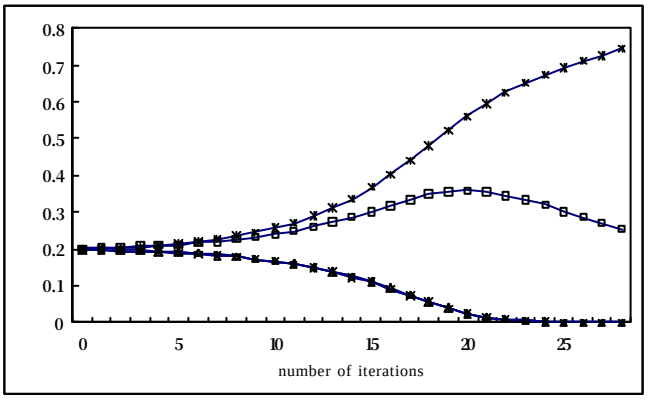


Figure 5. An example of bin incrementing.

Figs. 6(a) and (b) show examples of convergence characteristics of five candidates of eye-mouth locations and face outline, respectively with the first frame of the Miss America sequence, where x-axis denote the number of iterations and y-axis signifies the probability. Five candidates are chosen by the probability value of each candidate, especially highly ranked five candidates at the final iteration stage are chosen and the probability values of them are illustrated. Fig. 6(b) shows the convergence characteristic when there is a competitive candidate. Dominant candidate converges to one rather slowly relative to Fig. 6(a). After 20 iterations, the convergence to a correct solution is observed. The maximum iteration number is experimentally set to 30. It is noted that face outline candidates and eye-mouth location candidates are independently ranked and relaxation finds the most probable set of combined facial features. The possibility of the candidates ranked highly at the initial stage is lowered as iteration proceeds because they do not satisfy the geometric relationships with other candidates, which is implemented by the relaxation process. In Fig. 6(a), the dominant candidate for eye-mouth location is the 10^{th} ranked candidates at the initial stage, whereas in Fig. 6(b), the dominant candidate for face outline is the fifth ranked candidates at the initial stage. The dominant candidate satisfying the geometric relationships becomes a higher ranked candidate as iteration proceeds.



(a) eye-mouth candidates



(b) face outline candidates

Figure 6. Convergence characteristics of feature candidates in relaxation.

Figure 7. Experimental results.

Fig. 7 shows the experimental results of the first and 85th Miss America images and other test face images. To show the performance of the proposed algorithm, the most probable candidates for face outlines, eye locations, and mouth locations are superimposed on the input images. Experiments show that the proposed algorithm yields satisfactory performance.

4. CONCLUSION

In this paper, a relaxation-based algorithm to detect facial features is proposed. The proposed algorithm is composed of three parts: detecting primitive facial feature candidates, detecting additional facial feature candidates, and selecting the most probable set of facial features from a number of candidates that satisfy the geometric relationships of the face model. The proposed algorithm employs the fitnesses and geometric relationships of all feature candidates, yielding more reliable results. Further research will focus on the development of the algorithm to detect more reliable candidates that yield a higher face recognition rate.

ACKNOWLEDGMENTS

This work was supported by the Brain Korea 21 Project.

REFERENCES

1. S. Y. Lee, Y. K. Ham, and R.-H. Park, Recognition of human front faces using knowledge-based feature extraction and neuro-fuzzy algorithm, *Pattern Recognition* **29**(11), 1863-1876 (1996).
2. A. Eleftheriadis and A. Jacquin, Automatic face location detection for model-assisted rate control in H.261 compatible coding of video, *Signal Process.: Image Commun.* **7**(4-6), 435-455 (1995).
3. R. Kothari and J. L. Mitchell, Detection of eye locations in unconstrained visual images, in *Proc. Int. Conf. Image Processing 96* **3**, 519-522, Lausanne, Switzerland (1996).
4. A. L. Yuille, P. W. Hallinan, and D. S. Cohen, Feature extraction from faces using deformable templates, *International Journal of Computer Vision* **8**(2), 99-111 (1992).
5. K.-M. Lam and H. Yan, An analytic-to-holistic approach for face recognition based on single frontal view, *IEEE Trans. Pattern Anal. Machine Intell.* **20**(7), 673-686 (1998).
6. X. Xie, R. Sudhakar, and H. Zhuang, Corner detection by a cost minimization approach, *Pattern Recognition* **26**(8), 1235-1243 (1993).
7. S.-H. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, Facial feature detection using geometrical face model: An efficient approach, *Pattern Recognition* **31**(3), 273-282 (1998).
8. C.-H. Lin and J.-L. Wu, Automatic facial feature extraction by genetic algorithms, *IEEE Trans. Image Processing* **8**(6), 834-845 (1999).
9. A. Rosenfeld, R. A. Hummel, and S. W. Zucker, Scene labeling by relaxation operations, *IEEE Trans. Syst., Man, Cybern.* **6**(6), 420-433 (1976).