EURASIP Journal on Applied Signal Processing 2003:9, 878–889

Face Detection Using a First-Order RCE Classifier

Byeong Hwan Jeon

Signal Processing Laboratory, School of Electrical Engineering, Seoul National University, Seoul 151-742, Korea

Institute of Intelligent Systems, Mechatronics Center, Samsung Electronics Co., Ltd. Suwon, Gyeonggi-Do 442-742, Korea
Email: jeon@samsung.com

Kyoung Mu Lee

Department of Electronics and Electrical Engineering, Hong-Ik University, Seoul 121-711, Korea
Email: kmlee@wow.hongik.ac.kr

Sang Uk Lee

Signal Processing Laboratory, School of Electrical Engineering, Seoul National University, Seoul 151-742, Korea
Email: sanguk@diehard.snu.ac.kr

Received 9 September 2002 and in revised form 9 April 2003

We present a new face detection algorithm based on a first-order reduced Coulomb energy (RCE) classifier. The algorithm locates frontal views of human faces at any degree of rotation and scale in complex scenes. The face candidates and their orientations are first determined by computing the Hausdorff distance between simple face abstraction models and binary test windows in an image pyramid. Then, after normalizing the energy, each face candidate is verified by two subsequent classifiers: a binary image classifier and the first-order RCE classifier. While the binary image classifier is employed as a preclassifier to discard nonfaces with minimum computational complexity, the first-order RCE classifier is used as the main face classifier for final verification. An optimal training method to construct the representative face model database is also presented. Experimental results show that the proposed algorithm yields a high detection ratio while yielding no false alarm.

Keywords and phrases: face detection, face model, Hausdorff distance, clustering algorithm, RCE classifier.

1. INTRODUCTION

In recent years, due to the potential applications in many fields, including surveillance, authentication, video indexing, and so forth, face detection and recognition problems have gained much attention in computer vision society. The face detection problem is to locate human faces in a scene or a sequence of images, and the face detection technique not only can be used as a key preprocessing step for face recognition but also has its own importance in several applications, such as tracking, video indexing, and so on. In general, the face detection problem is known to be very difficult due to the variations in race, gender, pose, expressions, adornments, illumination, and scale.

Face detection can be considered as a pattern recognition problem and can be solved by statistical pattern classification techniques [1, 2], yielding the Boolean output: face or nonface. Functionally, well-organized parametric classifiers, such as Bayesian classifier [3], artificial neural network [4, 5], support vector machine [6, 7], have been used to classify the feature vectors by supervised classification techniques in the feature space. These parametric classifiers for face detection use high degree of data abstraction such as a set of trained weights, coefficients, or probabilities. Usually, those parameters are extracted from the training sample faces.

Alternately, nonparametric clustering-based approaches for face detection or pattern classification have been also proposed [8, 9, 10]. A model-based clustering algorithm [11] tries to describe the face subspace using both representative face models and nonface models which are selected during the training step. In general, a clustering algorithm is effective when the distribution of the feature vectors is not known in advance.

Note that the performance of the pattern classification or recognition can be improved substantially by selecting appropriate features [12], combining the multiple classifiers [13], or also by defining multiple similarity measures [14]. Various similarity measures and properties are analyzed in [15].

In this paper, we present a new clustering-based face detection algorithm that locates frontal views of human faces with arbitrary in-plane rotation and scale in complex scenes.

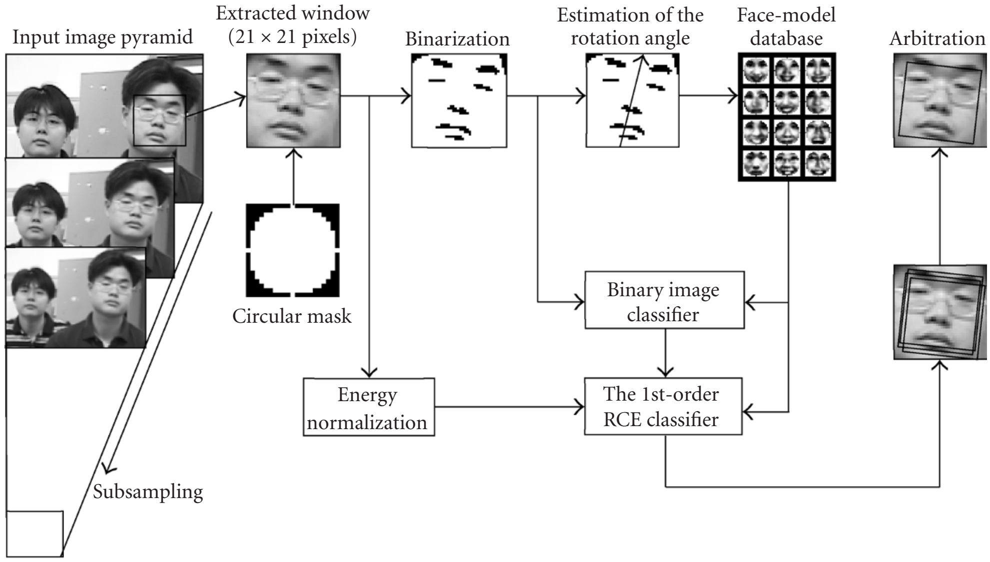


Figure 1: The overview of the proposed face detection system.

Unlike conventional algorithms which describe the shape of the face subspace in the feature space by using parametric statistical pattern classifiers, the proposed algorithm tries to model the cluster covered by each training face model using a first-order reduced Coulomb energy (RCE) classifier with multiple distance threshold determined by false negatives in the feature space. The resultant shape of the face space is the union of those modelled clusters. As a result, the boundary of the modelled face space becomes more accurate so that the proposed face detection algorithm yields a high detection ratio while yielding a smaller number of false alarms than the conventional methods.

In order to cope with the rotation and scale problems, an image pyramid is constructed for an input image first. And then, candidate face regions are extracted and their orientations are estimated by using the Hausdorff distance [16, 17] between each test window in pyramidal images and a set of rotated versions of a binary face abstraction model. Then, the proposed algorithm classifies those face candidates using two subsequent face classifiers: a binary image classifier, and a first-order RCE classifier which is the extended version of the original RCE classifier explained in [18]. While the binary image classifier is employed as a preclassifier to reduce the computational burden for selecting appropriate candidate faces, the first-order RCE classifier is used for further and final verification. Experimental results demonstrate that the performance of the proposed face detection algorithm is quite satisfactory.

Section 2 describes the overview of the proposed face detection system. A method to obtain face candidates is explained in Section 3. Detailed description on the proposed first-order RCE classifier is given in Section 4. Section 5 presents the experimental results of the proposed algorithm on the Carnegie Mellon University (CMU) test images. The conclusions are drawn in Section 6.

2. THE SYSTEM OVERVIEW

Figure 1 shows an overview of the proposed face detection system. The system is composed of several key processing modules and a face model database. A set of pyramidal images of an input image is constructed first to cope with the scale problem of a face. In this image pyramid, the scale is reduced recursively by a factor of 1.2.

Every window, of size 21×21, in the pyramid is then examined from the top to the bottom. In order to reduce the effect of hair or background region and consider the face region exclusively, a circular mask is applied to each rectangular window. Then, a binary image in the circular mask is obtained and used for estimating the face orientation as well as measuring the binary similarity to the face models.

A simple binary face abstraction model and its rotated versions in 1 degree resolution are constructed to detect face candidates and determine their orientation. For a given binary image window, the matching to all the face abstraction models is performed, and by identifying the best matching model and its score, we can determine not only whether the image window is a face candidate or not but also what the orientation of it is. Once the image window is decided to be a face candidate, further verification using both the binary image classifier and the first-order RCE classifier is performed.

A binary image classifier is employed to eliminate possible nonfaces among the face candidates with less computational complexity. If the input binary image is similar to the one in the face model database, the energy normalization is performed on the windowed image, and then it is classified by the first-order RCE classifier presented in Section 4.

Actually, the first-order RCE classifier decides finally that a candidate is a face or nonface. Since the resolution of the orientation of a face model is 1°, multiple face candidates can occur at similar location. Even though some of those can be classified as nonface, many candidates will be classified as a face. In case of multiple detection for a face, the final face location is simply determined as the one yielding the minimum distance from (or the maximum similarity with) a face model in the database.

The face model database consists of representative faces which are selected optimally from a set of sample faces. Each face model has 360 rotated versions of a binary image and an energy-normalized gray-level image. In addition, each face model has one corresponding nonface image detected as a false positive during the training step and 180 distance thresholds trained during the training step. A nonface image of a face model is used for calculating the reference direction presented in Section 4.

FIGURE 2: Examples of face images and the binary images obtained by using the quasi-invariant property of face images.

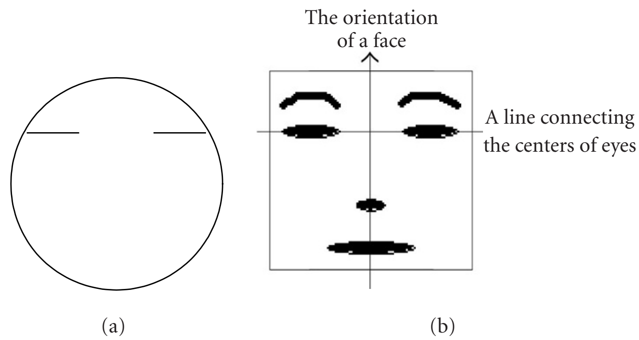


FIGURE 3: The face abstraction model and the face orientation. (a) The face abstraction model with two eyes. (b) The orientation of a face is perpendicular to the line connecting two eyes.

3. EXTRACTING FACE CANDIDATES

3.1. Obtaining binary face image

We observe that among face features, some features are always darker than the skin area, regardless of human races, facial expressions, head poses, or illumination conditions, except for the extreme cases. And it is also found that the proportion of the area of those facial features, such as eyes, eyebrows, mouth, and nostrils, to that of the circular face mask under normal illumination conditions does not change rapidly, yielding a quasi-invariant information of human faces. Therefore, except some extreme cases such as severely slanted illuminations or blurring, this quasi-invariant property can be utilized with most natural face images. From repeated experiments on various types of face images, it is empirically found that the area of those face features is approximately 20% of the total area of the circular mask. Thus, in this work, we use this value as the threshold for obtaining binary face images. Let N be the total number of pixels in the circular mask and n_i the number of pixels with gray level i. Then, a binary face image is obtained by segmenting the original image with the threshold value T which satisfies the following equation:

$$\frac{1}{N}\sum_{i=1}^{T} n_i = 0.2. \quad (1)$$

Figure 2 shows several examples of face images and the obtained binary face images.

3.2. Finding face orientation

In this research, in order to extract face candidates and their orientations in an input image pyramid, we employ a face abstraction model. By measuring the Hausdorff distance between each rotated version of the face abstraction model and the binary image window, candidate faces along with their orientations can be determined. Note that the orientation information is very important to the subsequent binary image classifier as well as to the first-order RCE classifier.

3.2.1. Face abstraction model

Eyes play an important role in determining the orientation of a face. Once the positions of two eyes are determined precisely, the face orientation can be obtained easily. In terms of intensity characteristics, eyes and eyebrows are relatively static features, compared with the nose or mouth. Although eyes and eyebrows actually move due to facial expression, the movement is unnoticeable in a small face patch of size 21×21.

A face abstraction model is a simple binary sketch of a face with only two horizontal line segments representing the two eyes. It is noted that the orientation of a face in a frontal view is always perpendicular to the line connecting the two eye centers as depicted in Figure 3a. Figure 3b shows the orientation of a face which is perpendicular to the line connecting the centers of two eyes. The orientation or angle of the upright frontal view of a face is defined to be $0°$, and it increases counterclockwise. To cope with the orientation, 360 rotated versions of face abstraction models are constructed.

3.2.2. Hausdorff distance measure

Once a binary image patch in an input image is obtained, the existence and the orientation of a face in that patch are determined by matching it to all the face abstraction models using the Hausdorff distance measure. Note that by employing the simplified face abstraction models, the computational complexity of the Hausdorff distance can be greatly alleviated as depicted in Figure 4.

Given two sets of points $A = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$ and $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$, the directed Hausdorff distance from A to B

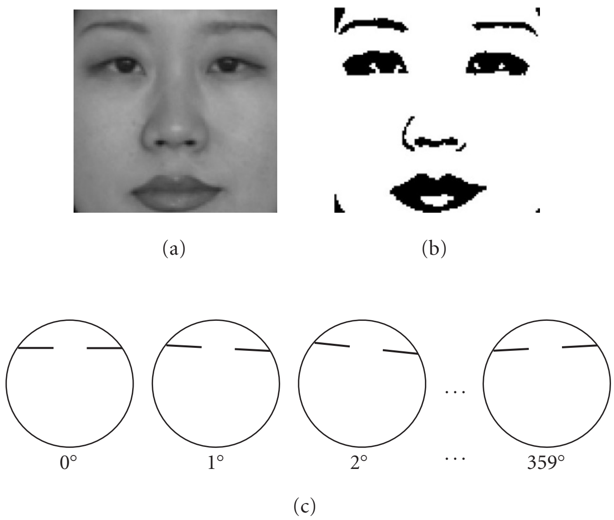


FIGURE 4: An example of the Hausdorff distance measurement between a binary face and the abstraction model with two eyes. (a) A gray-level face image, (b) a binary face image, and (c) the rotated versions of the face abstraction model to be matched.

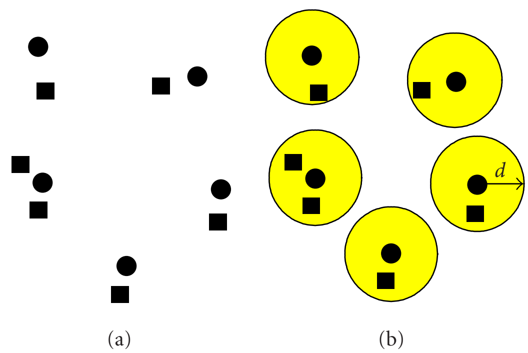


FIGURE 5: Illustration of the Hausdorff distance measure. (a) Two sets to be matched, A (dots) and B (squares), in a multidimensional space, (b) matching by the directed Hausdorff distance $h(A, B)$ with a threshold d is to check whether each circle of the radius d centered at each point in A includes at least one point in B or not.

is defined as

$$h(A, B) = \max_{\mathbf{a}\in A} \min_{\mathbf{b}\in B} \|\mathbf{a} - \mathbf{b}\|. \quad (2)$$

The directed Hausdorff distance measures the similarity between pattern A and any part of pattern B by identifying the point that is farthest from any point in B. Another way is to interpret it as the smallest radius d such that every point in A is within the distance d of some point in B [16, 17].

For test of face candidates, we use the directed Hausdorff distance from the face abstraction models to a binary input image. By the definition of Hausdorff distance, if there are m points in the face abstraction model and n points in a binary input image, then it is necessary to calculate the Euclidean distance $m \cdot n$ times. However, if the Hausdorff distance is given by d, then there should be at least one point in the circle of radius d centered at each face abstraction model, as shown in Figure 5. Thus, only Boolean operations are sufficient to calculate the Hausdorff distance, resulting in a significant saving in the computational cost.

3.3. A binary image classifier

Once the face candidates are identified, each of them is then examined by measuring the similarity of it to the faces in the face model database in binary mode. We define the distance D_b between two binary images to be the number of pixels that do not match. Then, the similarity between an input binary face candidate $\mathbf{u}$ and the mth binary face model $\mathbf{v}_m$ can be defined by the following binary image distance:

$$D_b^m = n(\mathbf{u} \oplus \mathbf{v}_m), \quad (3)$$

where the symbol $\oplus$ is the bitwise XOR operator, and $n(\cdot)$ is a function that counts the number of logic 1 (Boolean true).

Now, once the distances to all the binary model faces are calculated, the face candidate $\mathbf{u}$ is decided to be a binary face if the minimum value of them is less than a prespecified threshold, otherwise not.

4. THE FIRST-ORDER RCE CLASSIFIER

4.1. Modelling the face space

We assume that a multidimensional feature space is composed of two subspaces: face space and nonface space. The face space is considered as the set of all the individual human faces with possible variations including poses, expressions, aging, adornments, and illumination changes.

Let F be the face space in a multidimensional feature space. Note that, although the exact shape of F cannot be described visually, it will be very complex. In this research, instead of modelling the boundary of the face space in a parametric form, we attempt to represent it by the union of clusters of finite representative face samples. Let $\mathbf{f}_m$ ($m = 1, \ldots, M$) be the M representative face models selected from the K ($K > M$) training samples in the face space F, and F_m the individual cluster covered by $\mathbf{f}_m$. Then, the whole face space F can be modelled by the union of each cluster, given by

$$F' = \bigcup_{m=1}^{M} F_m, \quad (4)$$

where F' denotes the modelled face space.

Note that, in general, since face images are highly correlated, the volume of the face space is much smaller than that of nonface space.

4.2. Several model-based approaches

For simplicity, we assume that an arbitrarily shaped region in a 2-dimensional space, shown in Figure 6, is a face space made by the union of clusters corresponding to a finite

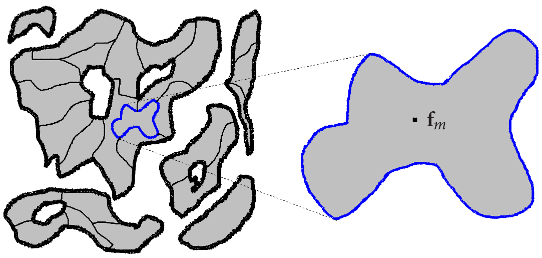


FIGURE 6: An arbitrarily shaped 2-dimensional space composed of several clusters.

number of representative face models. Note that the shapes of the clusters are not the same, and each representative face model may not be located at the center of the corresponding cluster. Now, our goal is to find an efficient way to model each face cluster so as to represent the whole face space accurately with a finite number of representative face models.

There are several model-based clustering algorithms to model a cluster covered by a representative face model. Jeon et al. [11] proposed a clustering algorithm in which the face cluster of a representative face model is initially considered as a hyperball with a relatively large specified diameter ($\mathbf{r}_f$), and then trimmed out by the cluster of nonface samples with smaller diameter ($\mathbf{r}_q$) as in Figure 7a. The nonface samples are false positives ($\mathbf{q}$), detected during a bootstrapping step using many nonface images. This method requires a larger number of nonface models than that of the face models, and the nonfaces located close to the face cluster can erode it, resulting in the degradation of the representation. The same problem also occurs in the 1-NN (nearest neighbor) method. By the 1-NN method, the face cluster is represented by the region where the distance to the representative face model is shorter than that to nonface samples, as shown in Figure 7b. Thus, if a nonface is located close to the true face cluster, the boarder of the face cluster can be altered severely by the nonface.

The RCE classifier [17] is an alternative way of the model-based clustering algorithm. The original RCE classifier employs a modifiable threshold for the radius of a hyperball corresponding to a pattern. During training, the radius is adjusted so that it becomes as large as possible without containing patterns of another category. In face detection problem, each face model has a modifiable threshold, and the threshold, starting from a sufficiently large value, is adaptively reduced by false positives detected in a training step. We refer to the original RCE classifier as the zeroth-order RCE classifier since the classifier employs only one distance threshold for a model with no angular component. So, the zeroth-order RCE classifier models the cluster of a face model as a minimum-bound circle (hyperball in the multidimensional space) as shown in Figure 7c. As a result, too many representative face models are needed for the zeroth-order RCE classifier to represent the face space sufficiently. Figure 7d shows the ideal first-order RCE classifier which can represent the cluster more accurately.

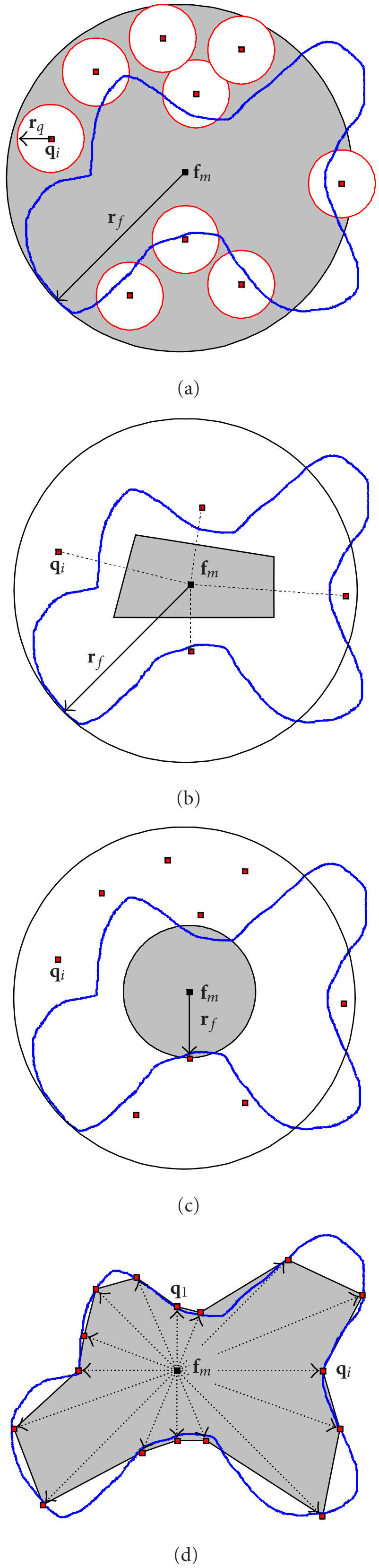


FIGURE 7: Several model-based clustering algorithms. (a) A distance threshold clustering algorithm, (b) 1-NN classifier, (c) the zeroth-order RCE classifier, and (d) the *ideal* first-order RCE classifier.

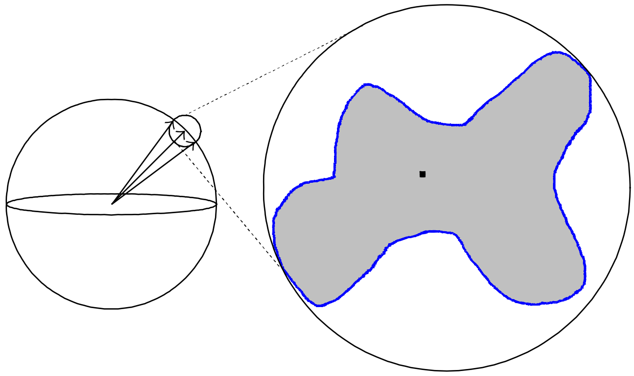


FIGURE 8: A 3-dimensional case of the energy-normalized feature space.

4.3. *Higher order RCE classifiers*

In an N-dimensional feature space, an Mth ($N \geq M > 0$)-order RCE classifier is defined by a distance threshold function of some M angular components centered at a certain vector such that

$$r(\mathbf{\Theta}), \quad \mathbf{\Theta} = (\theta_1, \theta_2, \ldots, \theta_N)^T, \tag{5}$$

while the zeroth-order RCE classifier has a single distance threshold value which is the same for all angular directions.

For simplicity, we consider the representation of a cluster using RCE classifier in a 3-dimensional feature space. If we normalize the feature vectors so that they have unit energy in the sense of L_2 norm, they are all projected onto the surface of the unit sphere as shown in Figure 8. If we assume that the cluster corresponding to each representative model is relatively small, we can approximate the cluster as a 2-dimensional region. To describe the boundary of the 2-dimensional region with respect to the given representative model in 1° angular resolution, 360 different distance values are needed in 360 angular directions. Thus, we can represent the 3-dimensional cluster shape by precisely using the distance (threshold) function of one angular variable, which is the first-order RCE classifier. Notice that a training procedure is required to get those 360 different distance values.

We extend this notion to the N-dimensional case. If the feature vectors are normalized, then they are projected onto the surface of the unit hyperball. We assume that the cluster of each face model is relatively small, then the feature vectors in the cluster lie in $(N-1)$-dimensional space. In the polar coordinate system, the $(N-1)$-dimensional space can be represented by one distance component and $(N-2)$ angular components. Thus, to represent the $(N-1)$-dimensional cluster ideally, we need an $(N-2)$th-order RCE classifier. However, this representation is impractical since, for large N as in the face vector case and sufficiently small angular resolution of m degree, there should be as large as $(360/m)^{N-2}$ threshold values for each face model.

Note that if we use a zeroth-order RCE classifier to describe the $(N-1)$-dimensional cluster, the cluster is modelled by a hyperball since it assigns the same threshold for all angular directions.

4.4. *The first-order RCE classifier*

The goal of the first-order RCE classifier is to model the face cluster more accurately by assigning multiple distance thresholds for some specified directions as shown in Figure 7d. Those distance thresholds are also trained by false negatives.

In contrast to the conventional zeroth-order RCE classifier, the proposed first-order RCE classifier has not only one distance component but also one angular component to describe an N-dimensional space. If we set the angular resolution to 1°, there are 360 distance threshold values for the first-order RCE classifier.

We assume that all the normalized face images are located close to each others on the surface of the N-dimensional hyperball, which can be approximated by $(N-1)$-dimensional space. Now, we denote $\mathbf{f}_m$ to be the mth representative face model and $\mathbf{q}_1$ to be the first false positive of it. Then, the reference direction vector becomes

$$\mathbf{r}_m = \mathbf{q}_1 - \mathbf{f}_m. \tag{6}$$

During the training stage, if a new false positive $\mathbf{q}$ occurs, then the angle θ between $\mathbf{r}_m$ and $\mathbf{q} - \mathbf{f}_m$ is calculated in 1° resolution by

$$\theta = \operatorname{acos}\left(\frac{\mathbf{r}_m \cdot (\mathbf{q} - \mathbf{f}_m)}{||\mathbf{r}_m||||\mathbf{q} - \mathbf{f}_m||}\right), \tag{7}$$

and the distance threshold for the mth representative face model along this angle, $T_g^{m,\theta}$, is obtained by

$$T_g^{m,\theta} = ||\mathbf{q} - \mathbf{f}_m||. \tag{8}$$

If a new false negative gives smaller distance from $\mathbf{f}_m$ than the old one, the threshold value for that angle is replaced by the new one. Thus, in this fashion, through the training process, the distance threshold values of the angular directions for each representative face model are repeatedly replaced with the new minimum value so that the boundary of each representative face model is specified by $T_g^{m,\theta}$, $m = 1, \ldots, M$, and $\theta = 0, \ldots, 179$. Note that there exist an infinite number of vectors that have the same θ degree as the reference vector $\mathbf{r}_m$, which lie on a hypercone in the $(N-1)$-dimensional space. Thus, in the above mentioned fashion, the first-order RCE classifier represents this whole family of vectors associated with θ by a single vector whose length is the minimum. Figure 9 shows a 3-dimensional case example. Note that when the system starts training the thresholds for 180 directions, a default initial threshold is given. So, if no threshold for a certain angle is trained by the nonface training samples through the training process, a default threshold value is set to that angle.

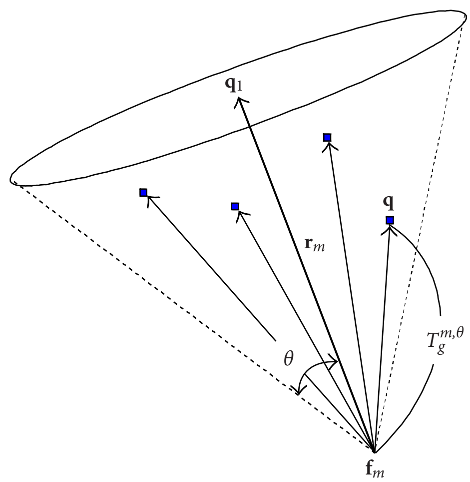


FIGURE 9: A family of vectors and its representative threshold associated with an angle θ with respect to the reference vector.

However, the first-order RCE classifier has two shortcomings. Since the arccosine function generates angles from 0 to π (not from 0 to 2π), the shape of the modelled face cluster becomes symmetric. Moreover, since the angle θ of each false positive in (9) is defined with respect to the reference direction vector $\mathbf{r}_m$, different choices of it (equivalently, the initial false positive $\mathbf{q}_1$ detected during the training step) may result in different shapes of modelled face clusters. As a result, the modelled cluster may lose parts of the original shape. Two examples are shown in Figure 10. Nevertheless, empirical study shows that the first-order RCE classifier is good enough to yield satisfactory results in face detection, which will be discussed in Section 5.

In the classification stage, when a normalized input image patch vector $\mathbf{p}$ is given, similar to the training stage, for all the representative face models $\mathbf{f}_m$, $m = 1, \dots, M$, the angles θ_m between $\mathbf{r}_m$ and $\mathbf{p} - \mathbf{f}_m$, given by

$$\theta_m = \operatorname{acos}\left(\frac{\mathbf{r}_m \cdot (\mathbf{p} - \mathbf{f}_m)}{||\mathbf{r}_m||||\mathbf{p} - \mathbf{f}_m||}\right), \quad m = 1, \dots, M, \tag{9}$$

and the distances between $\mathbf{p}$ and $\mathbf{f}_m$ along this direction

$$D_g^{m,\theta_m} = ||\mathbf{p} - \mathbf{f}_m||, \quad m = 1, \dots, M, \tag{10}$$

are calculated.

Then, the input image patch $\mathbf{p}$ is decided to be a face candidate if there exists an $\mathbf{f}_i$, $i \in \{1, \dots, M\}$, which satisfies

$$D_g^{i,\theta_i} < T_g^{i,\theta_i}, \tag{11}$$

where T_g^{i,θ_i} is the prespecified threshold for the ith face model.

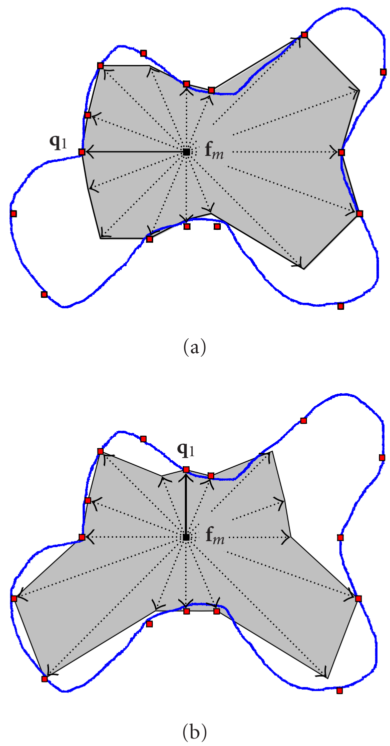


FIGURE 10: Symmetry and initial point dependency of the modelled shapes by the first-order RCE classifier. (a) A case where the initial point is located at the 9 o'clock direction. (b) Another case where the initial point is located at the 12 o'clock direction.

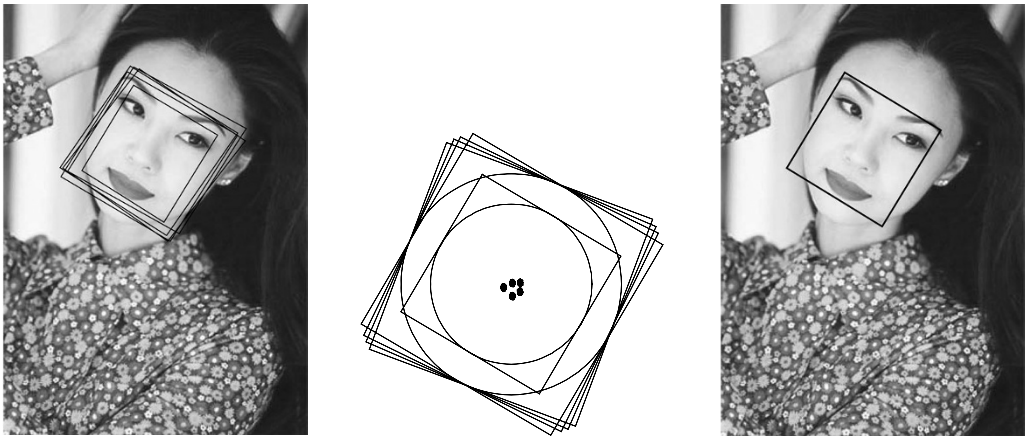
5. EXPERIMENTAL RESULTS

5.1. Constructing the face model database

To evaluate the performance of the proposed face detection algorithm, we have first constructed a face database for training. The face database was composed of 4,100 sample face images obtained from various sources including internet websites, academic face databases, such as Yale face database and Stirling face database, and some photo albums. From these images, each face region was manually cropped and normalized into the size of 21×21. Then, the rotated versions of the binary and energy-normalized faces of each face region at 1° resolution were obtained and stored.

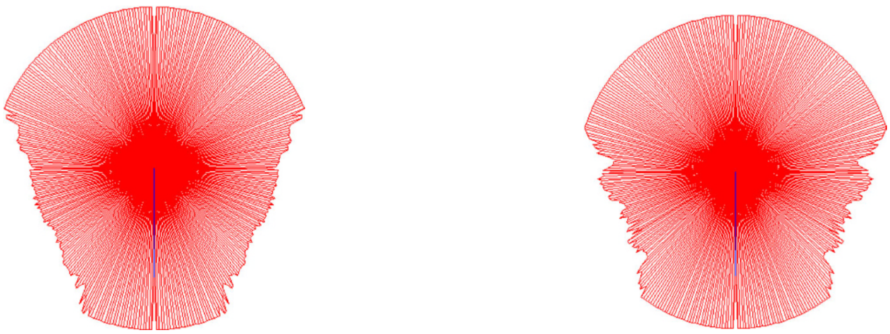
In order to optimize the number of representative face models, we applied the sequential forward selection (SFS) algorithm [2] for selecting representative face models among face samples. The SFS algorithm is a feature-selection method which selects the best single feature first, and then add one feature at a time which, in combination with the selected features, maximizes a criterion function. After the distance thresholds for each sample face model are determined in the first-order RCE training step, the represented face and nonface models for the experiments.

We have constructed a representative face model database which was composed of 227 representative faces extracted from the face samples using the proposed optimization and



(a) (b) (c)

Figure 11: Removing multiple detections. (a) An example of multiple detections on a face. (b) Selecting the best match among the cluster. (c) The final detection result.



(a) (b)

Figure 12: Two examples of face clusters trained by the first-order RCE classifier. (The trained threshold values are plotted in polar form.)

Table 1: The optimized number of the representative face and nonface models.

Methods	No. of faces	No. of nonfaces
The model-based clustering [11]	258	356
The first-order RCE	227	0

Table 2: The detection results on the rotated set of CMU test images and a comparison with the results of Rowley.

Proposed method		Rowley [5]	
Detection rate	False alarms	Detection rate	False alarms
91.0%	0	85.7 ∼ 90.1%	15 ∼ 303

training method, which means that the proposed optimization method removes 94.5% of the sample face images. Table 1 shows the number of the represented face and nonface models for the experiments.

We have tested about 4,200 nonface images to train 180 threshold values for each representative face model. Figures 12a and 12b show the training results of the 180 distance thresholds corresponding to angles from 0° to 179° for two representative face models in polar form, respectively. They are symmetric as expected. We can see that the thresholds for some angles are not trained, and thus set to be the default one.

5.2. *Experimental results with a test set*

The proposed face detection algorithm has been tested on the CMU face image database http://vasc.ri.cmu.edu/idb/html/face/profile_images/index.html, which consists of 50 images, containing 223 faces of arbitrary scales and rotations. It was observed that the proposed algorithm could detect 203 correct faces while yielding no false alarm. Table 2 summarizes the performance of the proposed algorithm, along with that of [5]. The sample experimental results are shown in Figures 13 and 14. These results demonstrate

Figure 13: Sample experimental results by the proposed method.

FIGURE 14: Sample experimental results by the proposed method.

that the proposed system yields a performance superior to the CMU's not only in the detection rate but also in false alarms.

Moreover, unlike conventional algorithms [3, 5, 6] which enforce the same operations to every window regardless of being face or nonface, since the proposed algorithm employed sequential face classifiers in the order of computational complexity, most nonface windows were rejected in the early stages through simple binary tests so that the overall complexity was greatly reduced. Generally, the computational complexity is directly proportional to both the number of the representative face models and the rotational resolution. In our experiments, it was showed that the proposed algorithm with 227 representative face models and 1° of rotational resolution run less than 1 minute for a 640×480 image with a 933 MHz Pentium 3 PC.

6. CONCLUSIONS

In this paper, a new face detection algorithm has been presented. Face candidates and their orientations were extracted by using binary face abstraction models first and then two subsequent face classifiers: binary and first-order RCE classifiers were followed to verify those face candidates. The binary classifier was adopted to filter out nonfaces from the face candidates by simple and fast operations, while the first-order RCE classifiers were used for final verification.

The novel first-order RCE clustering technique models the face and nonface subspaces in a feature space effectively with a finite number of the representative face models so that it classifies the face candidate robustly and accurately.

Experimental results showed that the proposed face detection algorithm could yield 91.0% of detection ratio and no false alarm with the CMU's rotated face image database, demonstrating better performance than that of Roweley's [5].

The proposed first-order RCE classifier-based face detection technique can be extended for nonfrontal views if the out-of-plane rotational views are sampled and modelled properly, which will be one of our future works.

REFERENCES

[1] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, Mass, USA, 2nd edition, 1990.

[2] A. K. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

[3] H. Schneiderman and T. Kanade, "Probabilistic modeling of local appearance and spatial relationships for object recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '98)*, pp. 45–51, Santa Barbara, Calif, USA, June 1998.

[4] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23–38, 1998.

[5] H. A. Rowley, S. Baluja, and T. Kanade, "Rotation invariant neural network-based face detection," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '98)*, pp. 38–44, Santa Barbara, Calif, USA, June 1998.

[6] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pp. 130–136, Puerto Rico, USA, June 1997.

[7] J.-C. Terrillon, M. N. Shirazi, M. Sadek, H. Fukamachi, and S. Akamatsu, "Invariant face detection with support vector machines," in *Proc. IEEE International Conference on Pattern Recognition (ICPR '00)*, pp. 210–217, Barcelona, Spain, September 2000.

[8] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 39–51, 1998.

[9] T. D. Rikert, M. J. Jones, and P. Viola, "A cluster-based statistical model for object detection," in *Proc. IEEE International Conference on Computer Vision (ICCV '99)*, pp. 1046–1053, Corfu, Greece, September 1999.

[10] S. Pittner and S. V. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 83–88, 1999.

[11] B. H. Jeon, K. M. Lee, and S. U. Lee, "Rotation invariant face detection using a model-based clustering algorithm," in *Proc. IEEE International Conference on Multimedia and Expo (ICME '00)*, pp. 1149–1152, New York, NY, USA, July–August 2000.

[12] W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1991.

[13] J. Kittler, M. H. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

[14] I. Craw, N. Costen, T. Kato, and S. Akamatsu, "How should we represent faces for automatic recognition?," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 725–736, 1999.

[15] S. Santini and R. Jain, "Similarity measures," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 871–883, 1999.

[16] D. P. Huttenlocher, G. A. Klanderman, and W. A. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[17] W. A. Rucklidge, *Efficient Visual Recognition Using the Hausdorff Distance*, Springer-Verlag, Berlin, Germany, 1996.

[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, New York, NY, USA, 2001.

Byeong Hwan Jeon was born in Kyungnam, Korea in 1965. He received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University, Seoul, Korea in 1988 and 1990, respectively. In February 1990, he joined Samsung Electronics Co., Ltd., Korea, where he is now a Principal Engineer. He is working toward his Ph.D. degree at the School of Electrical Engineering, Seoul National University since 1998. His research interests include computer vision, image understanding, and man-machine interface.

Kyoung Mu Lee received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University, Seoul, Korea in 1984 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles in 1993. From 1993 to 1994, he was a Research Associate at the Signal and Image Processing Institute at the University of Southern California. He was with the Samsung Electronics Co., Ltd., Suwon, Korea as a Senior Researcher from 1994 to 1995, where he worked on developing industrial real-time vision systems. In 1995, he joined the Department of Electronics and Electrical Engineering of the Hong-Ik University in Seoul, Korea as an Assistant Professor, where he is currently an Associate Professor. Dr. Lee is currently a member of the Editorial Board of the EURASIP Journal on Applied Signal Processing. His current primary research interests include computational vision, shape from X, 2D, and 3D object recognition, human-computer interface, and visual navigation.

Sang Uk Lee received the B.S. degree from Seoul National University, Seoul, Korea in 1973, the M.S. degree from Iowa State University, Ames in 1976, and Ph.D. degree from the University of Southern California, Los Angeles in 1980, all in electrical engineering. From 1980 to 1981, he was with the General Electric Company, Lynchburg, Va, working on the development of digital mobile radio. From 1981 to 1983, he was a member of the technical staff of M/A-COM Research Center, Rockvill, Md. In 1983, he joined the Department of Control and Instrumentation Engineering at Seoul National University as an Assistant Professor, where he is now a Professor at the School of Electrical Engineering. Currently, he is also affiliated with the Automation and System Research Institute and the Institute of New Media and Communications at Seoul National University. His current research interests are in the areas of image and video signal processing, digital communication, and computer vision. He served as an Editor-in-Chief for the Transaction of the Korean Institute of Communication Science from 1994 to 1996. Currently, he is a member of the Editorial Board of the Journal of Visual Communication and Image Representation, EURASIP Journal on Applied Signal Processing, and an Associate Editor for IEEE Transactions on Circuits and System for Video Technology. He is a member of Phi Kappa Phi.