Face Detection by Gabor Transform

Mohammadreza Noruzi; Mansour Vafadoost

ABSTRACT

A single reference image is used for detecting the face area in gray scale images. The algorithm uses Gabor transform as a powerful tool for feature extraction. Zero-crossing of Gabor transform is used for constructing the binary feature vectors and Hamming distance is used for measuring the similarity between feature vectors. Our developed method uses a pyramidal structure for finding the face images with different sizes. The algorithm has considerable robustness to the cluttered backgrounds and non-uniform illumination. More than 11,000 face images have been tested. The results show a very high detection rate for non-occluded face images with low facial expression. The algorithm can also be used for correcting the head rotation angle in input images. The algorithm can also be used for correcting the head rotation angle in input images. The proposed method is suitable for the cases that contain one face image.

KEYWORDS

Face detection, Gabor wavelet, Gabor transform, Hamming distance.

1. INTRODUCTION

Face detection in still images is a very important procedure which is used in Human Computer Interaction (HCI) and face recognition systems. The main problem of still image processing is the high computation time. The great number of pixels in an image, and consequently, the great size of feature vectors make the image processing be done slowly; therefore, most of the developed algorithms cannot be used for real applications.

In practice, the face detection is done by extracting a reference feature vector from the reference image and then finding an area in test image which has the most similar feature vector to the reference feature vector. The feature vector extraction is done by mathematical transforms and feature vectors comparison is done by classifiers. For feature extraction, mathematical tools like Gradient Vector [1], Principal Component Analysis (PCA) [2], [3], Kernel-PCA [4], Independent Component Analysis (ICA) [5], Discrete Cosine Transform (DCT) [6], Gabor Transform [7], [8], [9], Wavelet Transform [10] and Zernike Moments [11] are used. For feature comparison, classifiers like Bayesian [12], Neural Networks [13], Fuzzy Classifiers [14], Hidden Markov Model [15], Linear Discriminant Analysis (LDA) [16] and Support Vector Machine (SVM) [17], [18] have been exploited. The color, texture and structure of object, are the other important features which can be used in feature vector construction [19].

Gabor wavelet is one of the most powerful mathematical tools, which has been used widely for feature extraction. The main problem in this usage is the huge size of outputs. Most of the research activities have used an intermediate process between the output of filters and classifiers to reduce the dimensions of feature vector. A brief review of related works is presented in the following section.

Ben-Arie and Wang have used Gabor wavelet transform for initial feature extraction and Singular Value Decomposition (SVD) for classification. By using the SVD, in addition to reducing the dimension of feature vectors, the resulted feature vector is suited for 3D pose invariant object recognition [20]. Khalil has used Gabor wavelet transform and neural network for 2D objects recognition. Invariance to translation, rotation, and scale has been resulted by using of neural network [21].

Some characteristics of Gabor transform in object shape matching have been studied separately by Shan and Kyrki [22], [23]. Robustness to noise, scaling, rotation, translation and illumination are reported in these researches for Gabor transform. Pötzsch has used Gabor Jets for summarizing the feature vector. In that work, instead of applying the Gabor wavelet transform to the entire of image, some candidate points of object image are chosen and the Gabor wavelet transform is computed in these points. There is a tradeoff between the number of candidate points and the enrichment of feature vector. The less number of candidate points will produce shorter

M. Noruzi and M. Vafadoost are with the Faculty of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran (e-mail: noruzi@cic.aut.ac.ir, vmansur@aut.ac.ir).

feature vector with less strength but makes the program to run at higher speed [24].

A bottom-up structure is used by Park for object detection. It first detects object local parts such as corner points or edge segments. Then, the groups of detected local parts are processed to give evidence about unknown object pose. Detecting the parts of object is based on Gabor transform and Euclidean distance [25]. Gabor transform has also been used in Wavelet Networks for feature extraction. Wavelet Networks are similar to RBF networks, except that the radial basis function of RBF network is replaced by wavelet transforms [26].

Gabor transform and PCA have been used by Huang et al. for feature extraction and dimensionality reduction. They have used a polynomial neural network (PNN) for face detection [27], [28]. In another face detection application, Huang and Shimizu have utilized Gabor transform, 2D Haar wavelet and gradient direction for feature extraction and PCA for feature vector size reduction. Every feature vector is individually processed with a PNN classifier. The combination of each feature extraction method and PNN classifier is called an expert. Based on the fact that different detection methods offer complementary information about the face and non-face patterns, the results of multiple methods (experts) are combined to improve the final decision [29].

Imaoka and Okajima have developed a face detection system based on multiresolution searching algorithm with decision tree. The input image is divided into some sub blocks with different resolutions. Then, each sub block is convolved with a set of Gabor filters. The outputs are processed in two steps. In the fist step, a fast and rough classifier determines the faces from non faces. Then, the results of the first step are fed into a high accuracy classifier to improve the total classification results. The first classifier has a high false detection rate. The false detection rate is low for the second classifier [30].

The new method, which is proposed in this paper, has two brilliant characteristics. First, it depends on Gabor wavelet transform and uses the powerful abilities of this transform. Second, it has decreased the computation time by reducing the feature vectors complexity. The feature vectors' simplicity enables us to use the simplest and fastest form of classifier. By applying a zero-crossing procedure to Gabor coefficients, binary feature vectors are created. Simple form of this complexity reduction has been utilized by Daugman for comparing the textures of iris [31]. We have extended this technique and introduced a new multipurpose feature vector which is useful for fast face detection. The method uses grayscale images with no restriction on the size of input image. By using a pyramidal processing, the detection of different size faces is possible. A single common face image is used as a reference for detecting the face images from different persons.

2. Gabor Transform

Gabor transform is a time-frequency transform. Two dimensional form of this filter can be defined by (1).

$$g(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\{-0.5(x'^2/\sigma_x^2+y'^2/\sigma_y^2)\}\exp\{i2\pi\omega_0 x'\} \qquad (1)$$

where:

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

The function $g(x,y)$ depicts a filter with center of (x_0, y_0), $i=\sqrt{-1}$ is a constant, ω_0 is modulation frequency, θ_0 is angle of modulation and σ_x, σ_y are the effective length of filter in x and y axes [32], [33], [28].

It can be shown that the Gabor transform is a Short Time Fourier Transform (STFT) with a Gaussian window and a modulation in frequency domain. Therefore, some characteristics of Gabor transform are similar to its parent. By applying the Gabor transform to an image, the frequency contents of windowed image will be extracted along the θ_0 direction. There is a reverse relation between the time resolution and frequency resolution. The wider window in time domain will extract the more exact frequency content but the position of these contents will be more ambiguous [34].

The Gabor transform not only extracts the borders but also illuminates the expansion angles of them. This important property of Gabor filter has made it as a powerful tool for texture analysis [35], [36].

Gabor transform extracts the spectral information of borders, textures and also the orientation of them. Gabor transform has also physiological evidence. It has been shown that the mechanism of the mammals' vision system cellules in extracting the information is like to the Gabor wavelet transform [37]. Although the Gabor transform produces a rich and powerful feature vector but the complexity of feature vector is very high. This complexity arises from type of outputs which are real-imaginary numbers. High complexity of feature vectors results heavy processing time.

Gabor filter includes two parts: real part and imaginary part. The real part has even symmetry and the imaginary part has odd symmetry. The even part has nonzero mean values. The effect of this DC term becomes negligible when the ratio of the Gaussian window width to wavelength, has sufficiently large values [28]. In other words, by increasing the values of σ_y and σ_x, the DC term is decreased [4]. In general case and for different filter parameters, we found that centralizing the filter by omitting the mean value from the even part, improves the results. This omission changes the Gabor filter to a band-pass filter; therefore the background illumination of image has the minimum effect on filter output.

Gabor wavelet transform is formed by using a Gabor filter in different resolutions and scales. It is generated by applying dilation, translation and rotation to a Gabor elementary filter; the resulted basis functions are biorthogonal and there is redundancy in output of filters [33].

There is a tradeoff between the number of filters and the run time of process. One suitable choice which is also used in our algorithm is 3 consecutive scales and 4 angles of rotation for Gabor wavelet. Our experiments have shown that the fine and course resolutions have the same degree of importance in classification process. Center frequency of filters have been chosen with 1.3 octave range, while the highest center frequency is equal to $\pi/2$, the value of σ_x is equal to $\pi/15$ for the first filter and increase 1 octave in each scale, and finally σ_y is equal to half of σ_x.

3. Description of Algorithm

Before describing the details of algorithm, suppose that there is no coding procedure; and the output Gabor filter is used directly for making the feature vector. This feature vector is going to be used for face detection; so, suppose that there is a second feature vector related to test image. If the first feature vector becomes similar to the second feature vector, we can conclude that the images are similar. There is a big problem about the way of measuring the similarity between the two feature vectors. The Gabor transform result is a real-imaginary value. By encountering a little deviation in one of the images, the values of its feature vector will deviate. The dimension of feature vector is equal to the dimension of image multiplied by the number of filters in Gabor wavelet transform, so it is very high. It is obvious that comparing the feature vectors needs a special tool and cannot be done by a simple subtraction. This special tool is a classifier.

Neural Networks and Euclidean distance are the simplest tools which are suitable for this purpose. It is noticeable that comparing the feature vectors is time-consuming, even for the simplest classifier. This time-consumption arises from the complexity of feature vectors. Indeed, every element of feature vectors is a complex number and must be defined by some bytes in computer memory. Processing of such a numbers by any classifier needs mathematical operations like summation and multiplication, but these mathematical operations are time-consuming. If these operations are omitted, a higher speed of processing will be gained.

The basic idea of our proposed algorithm is reducing the complexity of feature vectors. This purpose is attained by zero-crossing the outputs of Gabor wavelet filter and coding the results into binary format. In zero-crossing procedure, only the sign of filters' output is considered for constructing the feature vector.

In general operating systems, each real number is usually defined with 8 bytes, and a complex number is defined with 16 bytes (128 bits). After zero-crossing, only 2 bits are produced. So, the feature vector size is reduced 64 times. This property has two important benefits. First, the amount of allocated memory is reduced. Therefore, the memory management becomes very simple. Second, instead of heavy processing of complex numbers, the computations are done by bitwise operators; Hamming distance is used for comparing the binary feature vectors.

Hamming distance is based on logical operations like AND/OR. It is fast and simple. There is no need for any especial classifiers and consequently no lateral training process. Although, the zero-crossing omits some information from initial feature vector, our experiments will show that the remaining information is enough for developing a reliable face detection system.

Fig. 1 shows the result of coding for a face image processed by a Gabor filter with vertical orientation. The result is equal size to the input image and can be displayed like an image.

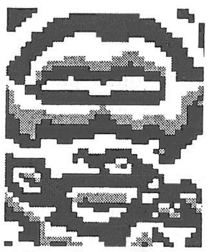

Figure 1: Result of coding for a Gabor transform of a face image. Filter parameters are: $\theta = 90$, $\sigma_x = 2\sigma_y = 0.21$, and $\omega_0 = \pi/2$.

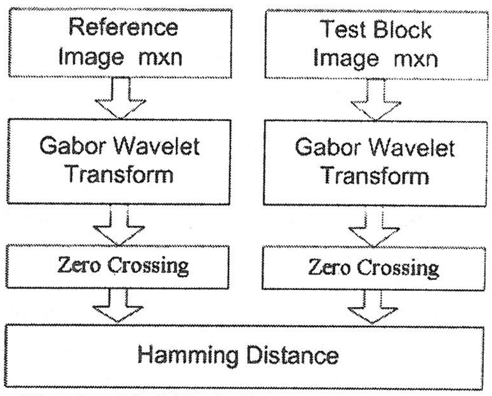

Figure 2: The simplified block diagram of matching algorithm for equal size images.

As shown, the Gabor transform with vertical orientation extracts the horizontal edges in image. In wavelet implementation of Gabor transform, there are different orientation and different scales. Thickness and orientation of extracted edges are related to the filter parameters. The simplified block diagram of matching algorithm is show in Fig. 2.

The inputs are two equal size images, one image is reference template image and another image is a selected portion of test image. This portion is determined by a searching algorithm to find the best matching between the template and input image.

A. Hamming distance

A special version of Hamming distance is used for comparing the binary feature vectors. This comparison is done by applying the logical operations to vectors elements. Suppose that V and R are two equal length vectors. These vectors are resulted from zero-crossing. So, each element of them is constructed from 2 bits. We calculate the Hamming distance between the two vectors by (2).

$$d_H = \left\{ \sum_{k=1}^{n} XOR(V[k], R[k]) \right\} / n \qquad (2)$$

where XOR represents the Exclusive OR logic function for both of bits simultaneously, k is the index of vectors and n is the length of vector as a normalization value.

The XOR function is defined as follows. If both of bits of V[k] are equal to decimal number of $R[k]$, the result of XOR is 0, otherwise the output is 1. This definition is a little different with the method of Daugman. He applied XOR function to each bit separately [31].This different definition has a noticeable reason. See section 4 for details.

If the two vectors are completely the same, the normalized Hamming distance will be equal to 0 and if the vectors are completely different, the normalized Hamming distance will be equal to 1.

After defining the template, a searching algorithm uses this template for finding the best matching between the template and the input image. The searching algorithm divides the test image into some overlapped blocks which are equal size to the reference template. Then a feature vector is extracted from each of these blocks. The extracted feature vectors are compared separately with the feature vector of template image. One of these comparisons has the least Hamming distance and its related block has the most similarity to the template, and it is the final result.

There is a problem in algorithm yet; if size of target object in test image differs from the size of template, then the searching algorithm will not work properly. The next section presents the solution.

B. Pyramidal algorithm

In real applications, the size of face image can differ from one input image to another input image. The size of extracted feature vector is equal to the size of face, so the resulted feature vectors are not equal size. In this case, computing the Hamming distances between these feature vectors and the reference template feature vector is not possible. To overcome this problem, we have used a pyramidal structure [38], [1]. The pyramidal algorithm, which is also known as multiresolution algorithm [19], produces some different scale images from each input image. In this case, the size of target face in one of these images will be near the size of reference object. An example of pyramidal image production is shown in Fig. 3.

The template image is slide over every image from pyramidal structure, and the similarity between the template and each portion is computed. Of course, pixel to pixel sliding of template will be very time consuming. Instead of one pixel displacement, we used $\Delta x > 1$ and $\Delta y > 1$ for this purpose.

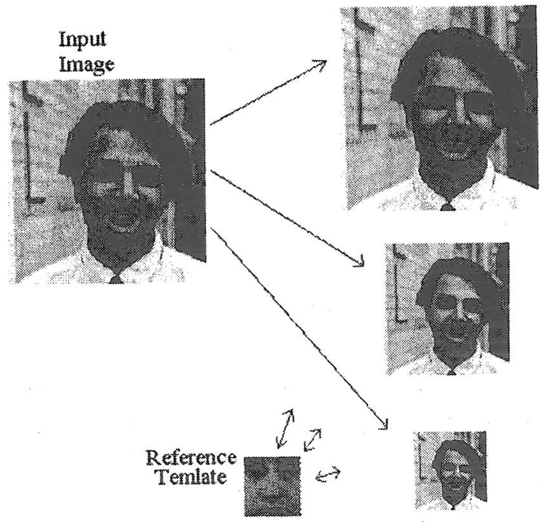

Figure 3: In pyramidal algorithm, the input image is reproduced with different sizes. The template is searched in each image separately.

After different experiments, we found (3) as a good estimation for Δx and Δy.

$$\Delta x = \Delta y = 1 + \mathrm{int}(\kappa \min(x, y)) \qquad (3)$$

where x and y are the length and height of template, κ is a coefficient equal to 0.1, and int( ) is a function that computes integer value of its input. The constant number 1 is used to assure that the result value will not become 0 in any situation. Suppose that input image is with size of 140×140 and the template is with size of 40×40. Sliding the template by one pixel, needs 10,000 comparisons. But the latter method reduces it to 400 comparisons.

Although, choosing a greater value for κ increases the speed of processes, in practice, the errors of detection are not acceptable.

The experiments have shown that about 5% difference between the sizes of two successive scales is suitable for gaining a good result. For example, if the size of input image is 1000×1000 and the number successive scales is 3, then we produce the second similar image with size of 950×950 and the third similar image with size of 903×903. For real applications, the number of pyramidal images is about 12.

C. Rotation effect

By applying the rotation to the object image, the matching percentage is decreased. It was observed that the rotations below 5 degrees have no significant effects on the results. Fig. 4 shows the Hamming distance between a reference face pattern image and two different persons' face image with various rotations.

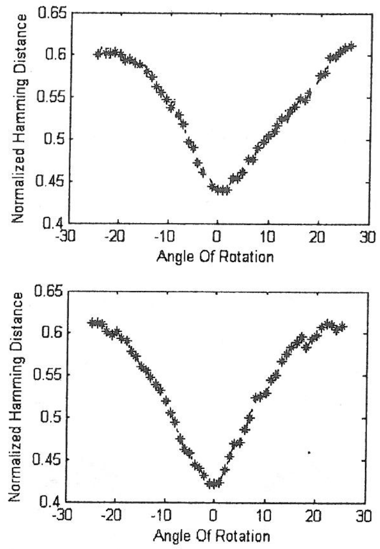

Figure 4: Hamming distance between a face pattern and the rotated face image from 2 different persons. Horizontal axis shows the rotation value in degree and vertical axis shows the value of Hamming distance.

As shown, the minimum distance is occurred when the relative angle between the template image and test image is zero. We have used this property for correcting the rotation angle of head in input images. This aim is achieved by applying some different manual rotations to input image and comparing the rotated images with reference template. The one with minimum Hamming distance is selected as output. The experiments have shown that the rotation step between two consecutive images must not be greater than 5 degrees.

4. THE PROPOSED METHOD VS. THE DAUGMAN'S METHOD

Because of some overlaps between our proposed algorithm and Daugman's method, it is preferred to describe these overlaps, clearly. Daugman applied Gabor transform to the segmented iris image and used its quantized phase value to produce a binary feature vector. He utilized Hamming distance for comparing the feature vectors.

Daugman method compares the texture of iris images. But before any comparison, the position of iris must be determined in input image with a localization algorithm. Each eye image is initially processed by this localization algorithm for finding the interior and exterior boundaries of iris; the localization algorithm is usually based on gradient method or Hough transform. The circular area of segmented iris is then mapped onto a constant size rectangular area. After all of these processes, the resulted image is convolved with Gabor filters, and the binary feature vector is extracted from it. Note that the constant size mapping area, causes to produce a constant length feature vector [31], [39].

We have previously developed an iris recognition system and used the Daugman's method for comparing the texture of irises. In that research, we also proposed a new method for detecting and segmenting the iris in input image. We used the phase value of Gabor transform directly, and without any coding procedure for iris detection and segmentation [40]. Now, we have extended that method and used it for developing a face detection algorithm.

Of course, Daugman's method is a motivation for our method, but there are some basic differences between them.

- Our proposed method is not for texture comparison. It is a face detection algorithm. It tries to find the similar structures. Note that the structure of an object is related with its main edges rather than to its texture.
- In Daugman's method, before applying the Gabor transform, the iris is localized by a segmentation algorithm. But in our method, there is no information about the position of face. The purpose of algorithm is to find the position of object by using the Gabor transform.
- If the rotation of images is not considered, the feature vector comparison is done once in Daugman's method. But in our method, this comparison is done many times. In this case, the binary format of feature vectors and also the fast Hamming distance comparator, are very significant for achieving the best computation time.
- We have used different type of Hamming distance, in comparison with Daugman's method. Daugman's method is used for comparing the texture of irises. The elastic texture of iris has random characteristics [41]. So, the probability of small deviations between two iris images from one person is very high. These small deviations produce small changes in Gabor filter output. In this case, after zero-crossing, the probability of one bit changing is very greater than 2 bits changing. So, in iris texture comparison, it is better to process each bit of zero-crossing separately. By this technique, most of the information of texture is preserved. In our method, the used images belong to different category. Random characteristics are not interesting and the reference feature vector deals with the deterministic structures, e.g., the edges of target object. We found that our new variant of Hamming distance, results in more robustness in cluttered background.
- A very important difference between two methods is as follows. In iris recognition, the Hamming distance value between the reference iris image and the second input image, can only determine that the two images belong to one person or not. In other words, in the case of dissimilar images, it cannot distinguish that the

second image is a random image or is an iris image of a different person [41]. In our method, a single reference image is used to find the similar objects in input images. Note that, Daugman's method deals with the identity of textures, but our method deals with the similarity of structures and objects.

5. EXPERIMENTAL RESULTS

A computer program has been developed based on the proposed algorithm. The program has a pyramidal structure that searches for a template in input image. The template is constructed from the reference face image at the starting of the program. The template is defined once at the beginning of the program and does not change again.

After determining the template, its feature vector is extracted. Then the program uses searches the template feature vector inside the different scales of test images with different resolution.

The computation time is dependent on the size of reference pattern and the size of input image. For example, for a Gabor wavelet with 3 levels and 4 orientations, a template with size of 30×30 pixels and a test image with size of 160×120 pixels, the computation time of searching the template inside the image was 350 milliseconds in a Pentium-IV 2400MHz with 512MB RAM. Note that by applying the pyramidal searching, the number of processed images and also the total computation time are increased.

Some results are shown in Fig. 5. In each column of these figures, the upper image shows the reference image and the rectangle in it shows the used reference template. The test image is shown below the reference image. The result of searching the reference template is shown by a rectangle in test image.

Figure 5: Detecting a reference pattern in test image. The reference template is shown with a rectangle in upper image and the result of detection is shown in the below image. Notice to the cluttered background in left result and non-regular illumination in right images.

By applying the rotation to the target image, the matching percentage is decreased. It was observed that the rotations below 5 degrees have no significant effects on the results. Robustness to the cluttered background and non-regular illumination is another characteristic of the proposed algorithm. Fig. 5 is a good example of this robustness. As shown, the bright window and light sources behind the person have not effected the detection result.

For evaluating the algorithm, we have used AR face database from Robot Vision Lab of Purdue University [42]. We have selected 550 face images randomly. The AR database includes more than 3000 face images which are taken in frontal view from 125 persons. The images are taken in simple background, with different facial expression and different conditions of illumination. The images include neutral expression, smile, anger, scream, wearing sunglasses, wearing scarf, normal illumination, left side illumination, right side illumination and bright frontal illumination. Furthermore, there are some persons that have worn lasses all the time. Some sample images and the results of face detection are shown in Fig. 6.

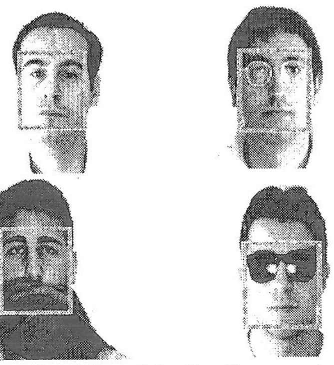

Figure 6: The results of face detection for some images from AR database.

A constant and single template is used for all of the images. This template is depicted in upper left corner of Fig. 5, and the results are shown in Table 1.

TABLE 1
RESULT OF DETECTION FOR AR FACE DATABASE. A SINGLE TEMPLATE IS USED FOR ALL THE IMAGES

Image Type	Frontal Illumination	Lateral Illumination
Without Beard	99%	98%
With Beard	98%	96%
Wearing Scarf	98%	96%
Screaming Expression	98%	No Data
Wearing Sunglasses	80%	50%

Based on the type of illumination, the images are divided into two groups. The images with normal illumination and bright frontal illumination are labeled by 'Frontal Illumination'. The images with other types of illumination are labeled by 'Lateral Illumination'. In 'Screaming Expression' case, the images of database are taken only with normal illumination, so there is no result for 'Lateral Illumination' case. The algorithm has shown great performance for face images with slight facial expression, all of the errors in this case are resulted from glasses in images.

In another experiment, we have used Cohn-Kanade face images database. The Cohn-Kanade database includes more than 8,700 face images which are taken in frontal view from 128 persons [43]. They range in age from 18 to 30 years. Sixty-five percent are female, 15 percent are African-American, and three percent are Asian or Latino. The images are taken in simple background, with different facial expression. A single and constant template is used for detecting the face position in all of the images. Some results of face detection are shown in Fig. 7.

Figure 7: The results of face detection for some images of Cohn-Kanade database.

All of the images are evaluated visually and two types of errors are found in results. It is possible to divide the errors into 'high importance' and 'low importance' errors. In high important errors, the position of face is found incorrectly. There are just 3 images with this error in whole of database, so the error rate is below 0.04%. Fig. 8 shows two samples with this kind of error.

Figure 8: High important errors of face detection in Cohn-Kanade database.

In low important errors, the position of face is found correctly but the circumference rectangle has incorrect size. This kind of error is generally produced by occluding the eyebrows with hairs and also is produced from facial expressions. All of the images of database have been evaluated visually. There are 70 errors in whole of database, so the error rate is below 0.9%. Fig. 9 shows two samples with this error.

In the third experiment, we have used BioID [44] database. It includes 1521 images from 23 persons. The total detection rate is 92.1% for this database. Some samples of this database are shown in Fig. 10, and the results of all databases are indicated in Table 2.

Another application of method is correcting the rotation angle of head in input image. As mentioned before, the minimum Hamming distance between the reference template and test image is produced at zero rotation. Usage of this characteristic is shown in Fig. 11.

Figure 9: Some samples of low important errors for face detection in Cohn-Kanade database.

Figure 10: Some results of BioID face database.

TABLE 2
THE RESULTS OF DIFFERENT DATABASES

Database name	Number of tested images	Correct detection
Cohn-Kanade	8700 (128 Persons)	99.0%.
AR (Excluding the sunglass images)	550 (125 Persons)	97.8%
BioID	1521 (23 Persons)	92.1%

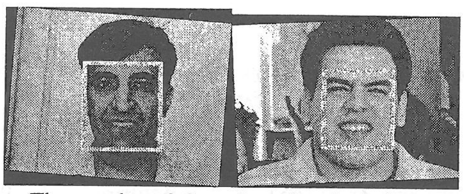

Figure 11: The results of face detection including automatic rotation correction.

The input test image is rotated automatically by program and the best matching between the reference template and each of the rotated images is computed. The rotated image with the best matching value is final result. We have tested it with 400 images; the error value of this correction was below 2.5 degrees for 385 images.

6. Comparison with Other Algorithms

The proposed algorithm has been tested on BioID face database. This database includes images with a large variety of cluttered backgrounds and different types of illuminations. Some researchers have used this database in their work. For example, Wu and Zhou tested their face detection algorithm with this database. They used the intensity of eye images for estimating the position of eyes. This goal was attained by a rule-based inference system. After finding some candidate points as eye image in input image, some different face templates were used to extract the face images around the candidate points. They reported 94.5% detection rate for BioID database [45]. Ramirez and Fuentes used intensity of eye image and the special shape of its histogram for estimating the position of eyes in input image. In the next stage, different types of classifiers like SVM and Neural Network were used to detect the probable faces around the estimated eye positions. They reported 95.13% detection rate for BioID face database [46]. Huang et al. used Cencus Transform for feature extraction. In the next stage, the feature vector was processed by a 22 levels inference system. He used Boosting method for training the inference system. They reported 90.0% detection rate for BioId database [47].

All of above systems are based on multiple training images and a high level decision system at their outputs. Note that these properties do not exist in our method.

Some methods may use a model for face detection. For example, Jesorsky et al. used a gradient based method for extracting the edges of input image in binary mode. He used Hausdorff distance for measuring the similarity between a face model and testing area in input image. They reported 91.8% detection rate for BioId database [48]. This method is very similar to our method in basics. Both of them work on single reference. They extract the feature vector in binary space and finally they use distance criterion for measuring the similarity between feature vectors.

The results of different methods for BioID database are presented inTable 3.

TABLE 3
COMPARISON WITH OTHER ALGORITHMS FOR BIOID FACE DATABACE

Algorithm	Reference images	Detection rate
Wu and Zhou [45]	Multiple	94.5%
Ramirez and Fuentes [46]	Multiple	95.13%
Huang et al. [47]	Multiple	90.0%
Jesorsky et al. [48]	Single model	91.8%
The Proposed Method	Single	92.1%

7. Conclusion

A new method for face recognition has been introduced that uses the Gabor wavelet transform for feature extraction. By zero-crossing and coding the value of Gabor transform, a binary feature vector is constructed. This simple form of feature vector enables us to use the Hamming distance for comparing the feature vectors. The Hamming distance is a very fast method which is based on logical operations. The results have shown that this simple form of feature vector decreases the rate of computation time 10 times in comparison with the case that uses non-quantized feature vector directly with Euclidean distance measurement.

By changing the distance between the target face and the camera, the size of face in input image will change. A pyramidal structure is used to overcome this problem. The pyramidal algorithm produces some images which are similar to the input image but have different resolution. It tries to find the best matching between the template image and these different scale images.

The proposed algorithm uses just one image as template. This functionality makes the algorithm very brilliant against the other algorithms. The characteristics of algorithm are explained as follows.

- There is no training procedure in proposed algorithm. So, by changing the template, the algorithm is adapted to our desired application. This change can be done online.
- Our preliminary experiments have shown that it is possible to choose the template from a non face object. In other words, the proposed method can be extended to an object detection algorithm just by changing the template. This new idea is under development.
- The output feature vector has binary format. Due to this property, it is possible to implement the algorithm by Hardware Description Language (HDL) into an Application Specific Integrated Circuit (ASIC) design.

The algorithm has considerable performance in cluttered backgrounds. It has great robustness to the disturbances like non-uniform illumination and bright light sources in background.

One of the restrictions of algorithm is that the face images must be captured in frontal view. Another restriction of algorithm is about the number face images in test image. The algorithm assumes that there is only one face image in each input image. If there is no face image, then a portion of input image will be determined incorrectly, and if there is more than one face image, the one which has the most similarity to the reference pattern, will be selected. Of course, these restrictions are not important for a face recognition system that its data acquisition process is done in a controlled environment.

8. References

[1] J. MacLean and J. Tsotsos, ''Fast pattern recognition using Gradient-Descent search in an image pyramid," in Proc. 2000 ICPR, pp. 877-881.

[2] H. C. Kim, D. Kim, and S. Y. Bang, ''Face recognition using the mixture-of-eigenfaces method," Pattern Recognition Letters, vol. 23, no 13, pp. 1549-1558, 2002.

[3] S.-K. Sun, Z. Chen, and T.-L. Chia, ''Invariant feature extraction and object shape matching using Gabor filtering," Lecture Notes in Computer Science, vol. 2314, pp. 95-104, 2002.

[4] C. Liu, ''Gabor-based kernel PCA with fractional power polynomial models for face recognition," IEEE Trans. Pattern Anal. Mach. Intell, vol. 26, no 5, pp. 572-581, 2004.

[5] J. Kim, J. Choi, J. Yi, and M. Turk, ''Effective representation using ICA for face recognition robust to local distortion and partial occlusion," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no 12, pp. 1977-1981, 2005.

[6] X.-Y. Jing and D. Zhang, ''A face and palmprint recognition approach based on discriminant DCT feature extraction," IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 34, no. 6, pp. 2405-2415, 2004.

[7] D. Dunn, W. Higgins, and J. Wakeley, ''Texture segmentation using 2-d gabor elementary functions," IEEE Trans. Pattern Anal. and Mach. Intell., vol. 16, no 2, pp. 130-149, February 1994.

[8] M. Idrissa and M. Acheroy, ''Texture classification using gabor filters," Pattern Recognition Letters, vol. 23, no 9, pp. 1095-1102, 2002.

[9] R. Alterson and M. E. Spetsakis, ''Object recognition with adaptive gabor features," Image Vision Comput, vol. 22, no 12, pp. 1007-1014, 2004.

[10] J. Zhang, D. Wang, and Q. Tran, ''A wavelet-based multiresolution statistical model for texture," IEEE Trans. Image Processing, vol. 7, no 11, pp. 1621-1627, 1998.

[11] J. Haddadnia, K. Faez, and M. Ahmadi, ''An efficient human face recognition system using pseudo zernike moment invariant and radial basis function neural network," IJPRAI, vol. 17, no 1, pp. 41-62, 2003.

[12] C. Liu, ''A bayesian discriminating features method for face detection," IEEE Trans. Pattern Anal. Mach. Intell, vol. 25, no 6, pp. 725-740, 2003.

[13] L.-L. Huang, A. Shimizu, Y. Hagihara, and H. Kobatake, ''Face detection from cluttered images using a polynomial neural network.," Neurocomputing, vol. 51, pp. 197-211, 2003.

[14] A. Baraldi and P. Blonda, ''A survey of fuzzy clustering algorithms for pattern recognition.," IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 29, no. 6, pp. 778-801, 1999.

[15] M. Bicego, U. Castellani, and V. Murino, ''Using hidden markov models and wavelets for face recognition," in Proc. 2003 International Conference on Image Analysis and Processing, pp. 52-56.

[16] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, ''Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell, vol. 19, no. 7, pp. 711-720, 1997.

[17] C. J. C. Burges, ''A tutorial on support vector machines for pattern recognition," Knowledge Discovery and Data Mining, vol. 2, pp. 121-167. Kluwer Academic Publishers, 1998.

[18] D. Casasent and Y.-C. Wang, ''A hierarchical classifier using new support vector machines for automatic target recognition," Neural Networks, vol. 18, no. 5-6, pp. 541-548, 2005.

[19] R. J. Campbell and P. J. Flynn, ''A survey of free-form object representation and recognition techniques," Computer Vision and Image Understanding, vol. 81, no 2, pp. 166-210, 2001.

[20] J. Ben-Arie and Z. Wang, ''Pictorial recognition of objects employing affine invariance in the frequency domain," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no 6, pp. 604-618, 1998.

[21] M. I. Khalil and M. M. Bayoumi, ''Affine invariants for object recognition using the wavelet transform," Pattern Recognition Letters, vol. 23, no 1-3, pp. 57-72, 2002.

[22] S. Shan, W. Gao, Y. Chang, B. Cao, and P. Yang, ''Review the strength of gabor features for face recognition from the angle of its robustness to mis-alignment" in Proc. 2004 ICPR, pp. I: 338-341.

[23] V. Kyrki, J. K. Kamarainen, and H. Kalviainen, ''Simple gabor feature space for invariant object recognition," Pattern Recognition Letters, vol. 25, no 3, pp. 311-318, 2004.

[24] M. Pötzsch, N. Krüger, and C. von der Malsburg, ''Improving object recognition by transforming gabor filter responses," Network: Computation in Neural Systems, vol. 7, no. 2, pp. 341-347, 1996.

[25] H.-J. Park and H. S. Yang, ''Invariant object detection based on evidence accumulation and gabor features," Pattern Recognition Letters, vol. 22, no 8, pp. 869-882, 2001.

[26] R. Feris, J. Gemmell, K. Toyama, and V. Kruger, ''Hierarchical wavelet networks for facial feature localization," in Proc. 2002 AFGR, pp. 118-123.

[27] L.-L. Huang, A. Shimizu, Y. Hagihara, and H. Kobatake, ''Face detection from cluttered images using a polynomial neural network," in Proc. 2001 ICIP, pp. 669-672.

[28] L.-L. Huang, A. Shimizu, and H. Kobatake, ''Robust face detection using gabor filter features," Pattern Recognition Letters, vol. 26, no 11, pp. 1641-1649, 2005.

[29] L. Huang and A. Shimizu, ''A multi-expert approach for robust face detection," Pattern Recognition, vol. 39, no 9, pp. 1695-1703, 2006.

[30] H. Imaoka and K. Okajima, ''An algorithm for the detection of faces on the basis of gabor features and information maximization," Neural Comput., vol. 16, no 6, pp. 1163-1191, 2004.

[31] J. G. Daugman, ''High confidence visual recognition of persons by a test of statistical independence," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no 11, pp. 1148-1161, 1993.

[32] W. Ma and B. Manjunath, ''Texture features for browsing and retrieval of image data," IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no 8, pp. 837-842, 2006.

[33] T. S. Lee, ''Image representation using 2D gabor wavelets," IEEE Trans. Pattern Anal. Mach. Intell, vol. 18, no 10, pp. 959-971, 1996.

[34] S. Qian and D. Chen, Joint Time-Frequency Analysis: Method and Application, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.

[35] D. Clausi and M. Jernigan, ''Designing gabor filters for optimal texture separability," Pattern Recognition, vol. 33, no 11, pp. 1835-1849, 2000.

[36] O. Pichler, A. Teuner, and B. J. Hosticka, ''A comparison of texture feature extraction using adaptive gabor filtering, pyramidal and tree structured wavelet transforms," Pattern Recognition, vol. 29, no 5, pp. 733-742, 1996.

[37] D. J. Field, ''Relations between the statistics of natural images and the response properties of cortical cells," Journal of Optical Society of America, vol. 4, no 12, pp. 2379-2394, 1987.

[38] M.-H. Yang, D. J. Kriegman, and N. Ahuja, ''Detecting faces in images: A survey," IEEE Trans. Pattern Anal. Mach. Intell, vol. 24, no. 1, pp. 34-58, 2002.

[39] J. G. Daugman, ''How iris recognition works.," IEEE Trans. Circuits Syst. Video Techn., vol. 14, no 1, pp. 21-30, 2004.

[40] M. Noruzi, M. Vafadoost, and M. S. Moin, ''Iris recognition: Localization, segmentation and feature extraction based on gabor transform," Lecture Notes in Computer Science, vol. 3982, pp. 1180-1189, 2006.

[41] J. G. Daugman, ''The importance of being random: statistical principles of iris recognition," Pattern Recognition, vol. 36, no 2, pp. 279-291, 2003.

[42] AR Face Database. Purdue Robot Vision Lab [Online]. Available: http://cobweb.ecn.purdue.edu/ aleix/aleix_face_DB.html

[43] T. Kanade, Y. li Tian, and J. F. Cohn, ''Comprehensive database for facial expression analysis," in Proc. 2000 FG. pp. 46-53, IEEE Computer Society.

[44] BioID Face Database. BioID Technology Research [Online]. Available: http://www.bioid.com/downloads/facedb/index.php

[45] J. Wu and Z.-H. Zhou, ''Efficient face candidates selector for face detection.," Pattern Recognition, vol. 36, no 5, pp. 1175-1186, 2003.

[46] G. A. Ramirez and O. Fuentes, ``Face detection using combinations of classifiers," in Proc. 2005 CRV '05: Proceedings of the The 2nd Canadian Conference on Computer and Robot Vision (CRV'05), Washington, DC, USA, pp. 610-615, IEEE Computer Society.

[47] Y.-S. Huang, H. ying Cheng, P.-F. Cheng, and C.-Y. Tang, ``Face detection with high precision based on radial-symmetry transform and eye-pair checking," in Proc. 2006 AVSS '06: Proceedings of the IEEE International Conference on Video and Signal Based Surveillance, Washington, DC, USA, p. 62, IEEE Computer Society.