

# An Automatic Face Recognition System in the Near Infrared Spectrum

Shuyan Zhao and Rolf-Rainer Grigat

Technical University Hamburg Harburg  
Vision Systems, 4-08/1  
Harburger Schloßstr 20  
21079 Hamburg, Germany  
{s.zhao, grigat}@tu-harburg.de

**Abstract.** Face recognition is a challenging visual classification task, especially when the lighting conditions can not be controlled. In this paper, we present an automatic face recognition system in the near infrared (IR) spectrum instead of the visible band. By making use of the near infrared band, it is possible for the system to work under very dark visual illumination conditions. A simple hardware enables efficient eye localization, thus the face can be easily detected based on the position of the eyes. This system exploits the feature extraction capabilities of the Discrete Cosine Transform (DCT) which can be calculated very fast. Support Vector Machines (SVMs) are used for classification. The effectiveness of our system is verified by experimental results.

## 1 Introduction

Face recognition has a wide variety of applications in commercial and law enforcement. Due to its nonintrusive characteristics, it emerges as a vivid research field in biometrics. Illumination is one of the challenges for face recognition. Most face recognition algorithms are associated with visible spectrum imagery, thus they are subject to changeable lighting conditions. For systems that have to work in the daytime and at night, infrared is a solution. However, thermal infrared is not desirable because of the higher cost of thermal sensors and poorer quality of the thermal images. Therefore near infrared is preferable and common silicon sensors can be used, since they are sensitive from the visible band to near infrared band (up to 1100 nm).

An automatic face recognition system must detect the face first. It requires either face detection as the first module, or localization eyes without face detection and cropping the face region accordingly. Since alignment is very significant for face recognition, it is advantageous to detect eyes first. When the eyes are localized, the face region can be segmented and aligned in accordance with the reference faces, in which case good recognition performance can be expected. On the other hand, precise face detection is a very tough task, so if recognition relies on the uncorrect face region, the system performance will be degraded. Moreover, when the lighting source is placed close to the camera axis oriented toward

the face, the interior of the eyes reflects the light and pupils appear very bright. This is the well-known “bright pupil” (red-eye) effect [1] and can be exploited to detect eyes.

Chellappa [2] and Zhao et al. [3] presented nice surveys of face recognition algorithms. Face recognition technology falls into three main categories: feature-based methods, holistic methods as well as hybrid methods. Feature-based approaches depend on the individual facial features, such as the eyes, nose and mouth, and the geometrical relationships among them. A representative of feature-based methods is *Elastic Bunch Graph Matching* [4]. Holistic methods take the entire face into account. Among global algorithms, the appearance-based methods are the most popular, for example, *Eigenface* [5], *Fisherface* [6], and *Independent Component Analysis* (ICA) [7]. Hybrid methods combine the feature-based and holistic methods, for instance, the algorithm presented in [8] is a combination of *Eigenface* and *Eigenmodule*. Recently, Huang et al. [9] proposed a hybrid method, which incorporated the component-based recognition with 3D morphable models.

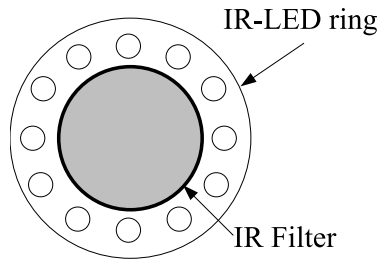
In this paper, we propose a face recognition system for access control in the near infrared spectrum. A simple and low cost hardware has been built up and used for data collection, which is presented in the next section. In Section 3, the algorithms for automatic face recognition are introduced. “Bright pupil” effect is utilized to localize the eyes, based on which the face is detected. DCT coefficients are then selected as features, Support Vector Machines (SVM) are employed to identify faces. Experimental results are shown in Section 4. In the last section conclusions are drawn and an outlook is given.

## 2 Hardware Configuration and Data Collection

### 2.1 Hardware

In order to make use of the “bright pupil” effect, a lighting source along the camera axis is necessary. In the literature [10] [1], the camera was equipped with two lighting sources, one along the camera axis and the other far from the camera axis. When the on-axis illuminator is switched on, the bright pupil image is generated; when the off-axis illuminator is on, the dark pupil image is produced. The difference of the two images gives clues about the position of the eyes. However, such a system needs a switch control which has to be synchronized with the frame rate so that the even and odd frames correspond to the bright and dark images respectively.

In our system a simplified hardware is used. An illuminating ring, consisting of 12 IR-LEDs, is placed along the axis of an inexpensive CCD camera. Thus only bright images can be generated by our system. The dark image will be constructed by using software rather than hardware. In order to obtain stable illumination conditions, an IR filter is used to block the visible light. The hardware is shown in Fig. 1.

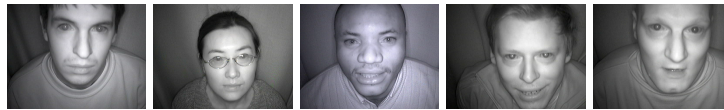


**Fig. 1.** The IR-LED ring and the IR filter

## 2.2 Data Collection

There are a couple of available face databases, but as far as we know, all of them contain face images taken under day light conditions. Therefore it is necessary for us to collect data ourselves in order to develop algorithms for face recognition in the near IR spectrum.

Reference data and test data were collected respectively. Our motivation is to realize access control in the situation where a camera is set above the door under protection. Therefore video sequences were captured under such a condition as test data, examples can be found in Fig. 2. Still images of frontal faces were taken as reference data, see Fig. 3.



**Fig. 2.** Some examples of test data



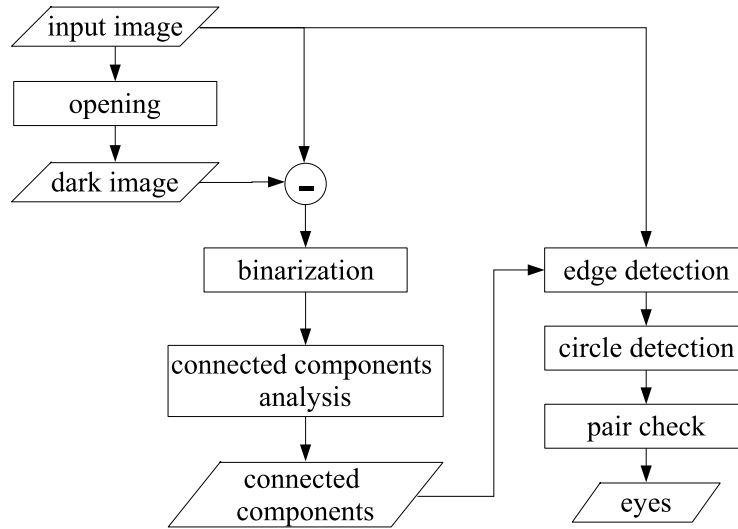
**Fig. 3.** Some examples of training data

In order to include variations of facial expressions, the subjects were requested to speak vowels a, e, i, o, and u. The resolution of all the images is  $320 \times 240$ . Images of 10 subjects from Europe, Asia and Africa have been collected so far.

### 3 The Automatic Face Recognition System

In the automatic face recognition system, the first step is eye localization. Based on the eye positions, the face region will be segmented and then normalized to a standard size. In the normalized face image, two eyes are located at the predefined positions. DCT features are then extracted, finally SVM is employed to recognize the face.

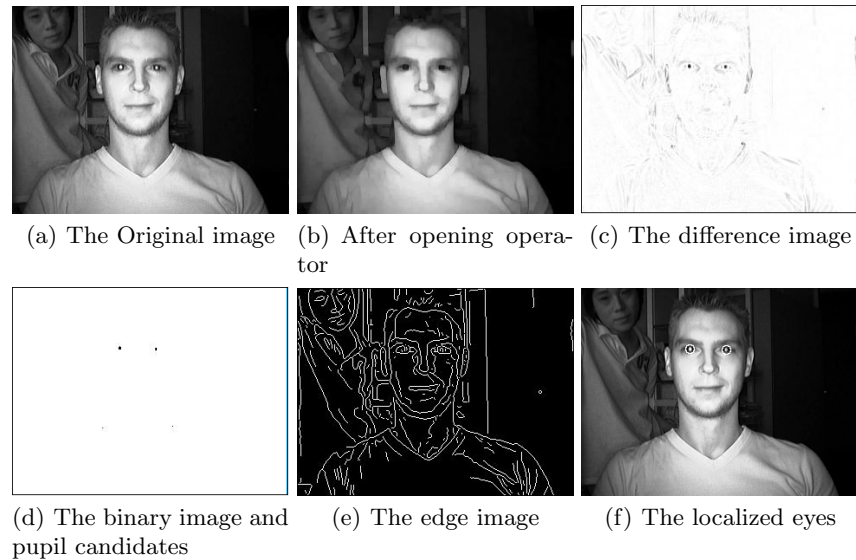
#### 3.1 Eye Localization



**Fig. 4.** The procedure of eye localization

The algorithm for eye localization is shown in Fig. 4. The morphological operator, opening, is at first applied to the original image. This operation removes the bright pupils. Then a difference image is obtained by subtracting the dark image from the original (bright) image, which contains only the bright areas. This difference image is then binarized. The connected components in the binary image are considered as pupil candidates that will be further verified.

The iris of the eye has a circular boundary, which can be detected due to the contrast between the iris and the area around it. Edges of the original image are detected using Canny's algorithm [11]. In the edge image, a window surrounding each pupil candidate is chosen, then Hough transform [11] is exploited to detect circles inside the window. Those candidates without a circular boundary are noise. Finally pair check is performed, and the two eyes are localized. The stepwise results are illustrated in Fig. 5.



**Fig. 5.** Stepwise results of eye localization

### 3.2 Face Recognition

Discrete Cosine Transform shares a closely related mathematical background with Eigenfaces. However, its merits over Eigenface are: it needs less training time; it is deterministic and does not require the specification of the data set. It seems more desirable when new faces are added into the database frequently. Besides, fast algorithms are available to calculate DCT.

2-dimensional DCT is performed on the normalized face. Only a subset of the DCT coefficients at the lowest frequencies is selected as a feature vector. These DCT coefficients have the highest variance and are sufficient to represent a face. To further reduce variations of illumination, these DCT coefficients are normalized to the DC component. SVMs [12] are used as the classifier.

## 4 Experiments

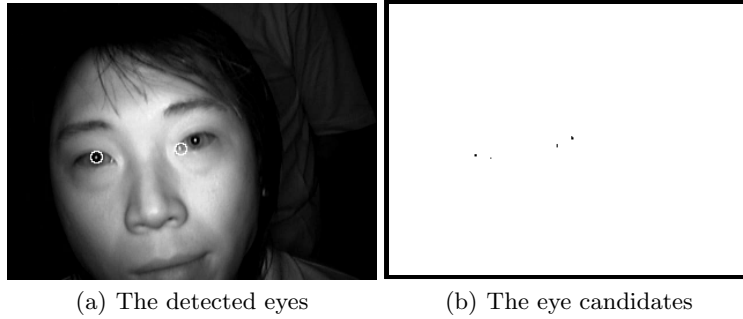
Three experiments are carried out. The first two experiments test the performance of eye localization and face recognition independently, and the third experiment tests the performance of the complete system.

*Eye Localization* 300 images are chosen to evaluate the eye localization performance. The criterion for correct localization is that any estimated eye position within a circle of radius 3 pixels from the true position is counted as a correct detection. In our test, 87.7% (263/300) accuracy is obtained, some results are shown in Fig. 6. The algorithm works when the horizontal distance of two eyes

is larger than 30 pixels, and the in-plane rotation of the face is less than  $25^\circ$ . If the noise is too close to the true eye, the algorithm may fail, see Fig. 7.



**Fig. 6.** Results of eye localization



(a) The detected eyes

(b) The eye candidates

**Fig. 7.** A false acceptance

*Face Recognition* In this experiment, only the performance of face recognition is taken into account, thus the centers of the eyes are manually marked. 250 face images (25 images per subject) are selected from our reference data set, collected in Section 2, to evaluate the algorithm. These images were captured at different photo sessions so that they display different illumination and facial expressions, even slight pose variations, see Fig. 8. Among them 120 images (12 images per subject) are randomly chosen for training and the left images for testing.



**Fig. 8.** Examples used to test the performance of face recognition

All the faces are scaled to the size  $48 \times 48$ , aligned according to the eye positions, and histogram equalized. 64 ( $8 \times 8$ ) DCT coefficients are extracted as features. LIBSVM [12], a library for Support Vector Machines, was used to performance recognition, where RBF kernel was employed. Recognition accuracy of 96.15% (125/130) has been achieved.

*Automatic Face Recognition* The same face recognition algorithm as the last experiment is used. By integrating with the eye localization module, the automatic face recognition system is tested. 500 images selected from the test data set collected in Section 2 are used for evaluation. These images demonstrate a lot of difference from the training data, because they were captured in different sessions and by a camera mounted above the door, as shown in Fig. 2. Although it is assumed that the subjects are cooperative, pose variations are not impossible. 398 faces are correctly recognized, corresponding to 79.6% accuracy. The reasons for the degraded performance are: 1) The imprecise eye localization results in uncorrect alignment; 2) perspective distortion exists because of the position of the camera, i. e. above the door.

## 5 Conclusion and Future Work

We present a robust and low cost system for automatic face recognition in the near infrared spectrum. By using the near infrared band, a stable illumination condition is obtained. Face images in the near infrared spectrum have been collected in our laboratory. The system utilizes the "bright pupil" effect to detect eyes. Relying on a series of simple image processing operations, eye localization is efficient. The face region is segmented and aligned according to the position of the eyes. DCT coefficients are selected as features, and the powerful machine learning algorithm Support Vector Machines is used for classification. The experimental results show that good performance is achieved by using our system.

*Future work* Pose variation is the toughest problem for face recognition. In the near future, the functionality of the system will be extended to be able to recognize faces with pose variations.

## References

1. Morimoto, C.H., Koons, D., Amir, A., Flickner, M.: Pupil detection and tracking using multiple light sources. *Image and Vision Computing* **18** (2000) 331–335
2. Chellappa, R., Wilson, C., Sirohey, S.: Human and machine recognition of faces: A survey. *IEEE*. **83** (1995) 705–740
3. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.: Face recognition: A literature survey. Technical Report CAR-TR-948, University of Maryland (2000)
4. Wiskott, L., Fellous, J.M., Krüger, N., Malsburg, C.B.: Face recognition by elastic bunch graph matching. In: Proc. 7th Intern. Conf. on Computer Analysis of Images and Patterns, CAIP'97, Kiel. (1997) 456–463
5. Turk, M., Pentland, A.: Eigenfaces for Recognition. *Journal of Cognitive Neuroscience* **3** (1991) 77–86
6. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **19** (1997) 45–58
7. Bartlett, M.S., Lades, H.M., Sejnowski, T.J.: Independent component representations for face recognition. In: Proc. of SPIE Symposium on Electronic Imaging. (1998) 528–539
8. Pentland, A., Moghaddam, B., Starner, T.: View-based and modular eigenspaces for face recognition. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'94), Seattle, WA (1994) 84–91
9. Huang, J., Heisele, B., Blanz, V.: Component-based face recognition with 3D morphable models. In: Proc. 4th Audio-Video-Based Biometric Person Authentication, Guildford, UK (2003) 27–34
10. Haro, A., Flickner, M., Essa, I.: Detecting and tracking eyes by using their physiological properties, dynamics, and appearance. In: Proc. of IEEE. Conf. Computer Vision and Pattern Recognition (CVPR2000). (2000) 163–168
11. Pratt, W.K.: *Digital Image Processing*, Third Edition. Wiley-Interscience, New York (2001)
12. Chang, C.C., Lin, C.J.: libsvm (2005) <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.