23

Face Recognition Using Optimized 3D Information from Stereo Images

Changhan Park[1] and Joonki Paik[2]
[1]Advanced Technology R&D Center, Samsung Thales Co., Ltd., [2]Graduate School of Advanced Imaging Science, Multimedia, and Film Chung-Ang University, Seoul Korea

1. Introduction

Human biometric characteristics are unique, so it can not be easily duplicated [1]. Such information includes; facial, hands, torso, fingerprints, etc. Potential applications, economical efficiency, and user convenience make the face detection and recognition technique an important commodity compared to other biometric features [2], [3]. It can also use a low-cost personal computer (PC) camera instead of expensive equipments, and require minimal user interface. Recently, extensive research using 3D face data has been carried out in order to overcome the limits of 2D face detection and feature extraction [2], which includes PCA [3], neural networks (NN) [4], support vector machines (SVM) [5], hidden markov models (HMM) [6], and linear discriminant analysis (LDA) [7]. Among them, PCA and LDA methods with self-learning method are most widely used [3]. The frontal face image database provides fairly high recognition rate. However, if the view data of facial rotation, illumination and pose change is not acquired, the correct recognition rate remarkably drops because of the entire face modeling. Such performance degradation problem can be solved by using a new recognition method based on the optimized 3D information in the stereo face images.

This chapter presents a new face detection and recognition method using optimized 3D information from stereo images. The proposed method can significantly improve the recognition rate and is robust against object's size, distance, motion, and depth using the PCA algorithm. By using the optimized 3D information, we estimate the position of the eyes in the stereo face images. As a result, we can accurately detect the facial size, depth, and rotation in the stereo face images. For efficient detection of face area, we adopt *YCbCr* color format. The biggest object can be chosen as a face candidate among the candidate areas which are extracted by the morphological opening for the *Cb* and *Cr* components [8]. In order to detect the face characteristics such as eyes, nose, and mouth, a pre-processing is performed, which utilizes brightness information in the estimated face area. For fast processing, we train the partial face region segmented by estimating the position of eyes, instead of the entire face region. Figure 1. shows the block diagram of proposed algorithm.

This chapter is organized as follows: Section 2 and 3 describe proposed stereo vision system and pos estimation for face recognition, respectively. Section 4 presents experimental, and section 5 concludes the chapter.

Source: Face Recognition, Book edited by: Kresimir Delac and Mislav Grgic, ISBN 978-3-902613-03-5, pp.558, I-Tech, Vienna, Austria, June 2007

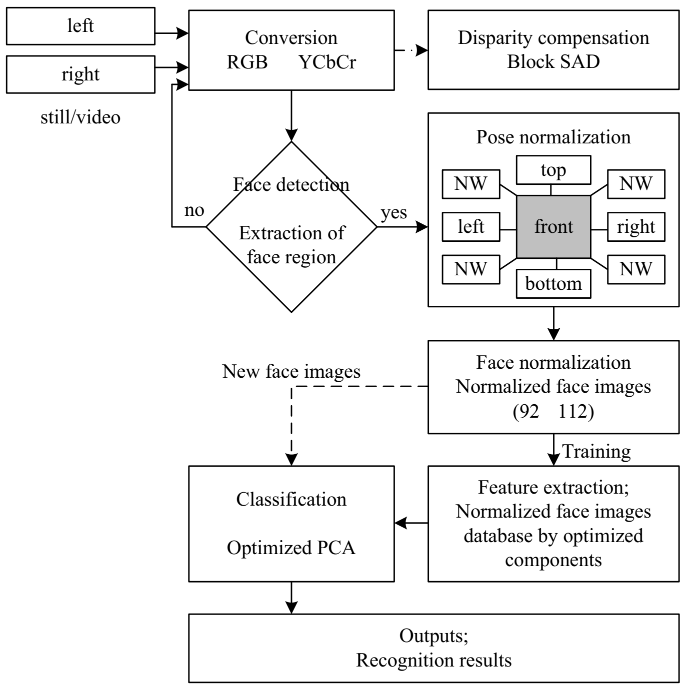


Figure 1. Block diagram of the proposed algorithm

2. Proposed stereo vision system

In order to acquire the distance and depth information, we use a parallel stereo camera as shown in Figure 2. From the stereo camera, we obtain the disparity between left and right images and estimate the distance by a stereo triangulation.

2.1 Disparity compensation of stereo images

A block matching algorithm is used to extract the disparity in the stereo images, after applying 3×3 Gaussian noise smoothing mask.

In general, the block matching algorithm uses the mean absolute difference (MAD) or the mean square difference (MSD) as a criterion. However, the proposed method uses the sum of absolute difference (SAD) to reduce computational complexity as

$$SAD = \sum_{i=0}^{x=N} \sum_{j=0}^{y=N} \left| I_L(i, j) - I_R(i+k, j) \right|, \tag{1}$$

where I_L represents the $N_x \times N_y$ block of left image, I_R represents the $N_x \times N_y$ corresponding block of right image, and k represents the disparity between left and right images. In the stereo image matching, the disparity compensation between left and right images should be performed. When a point in the 3D space is projected on left and right images, the virtual line connecting two points is called an epipolar-line [9]. The corresponding blocks of the stereo images are matched on the epipolar-line with the same x-coordinate. The modified block matching algorithm based on 4×4 block is used for fast processing as shown in Figure 3.

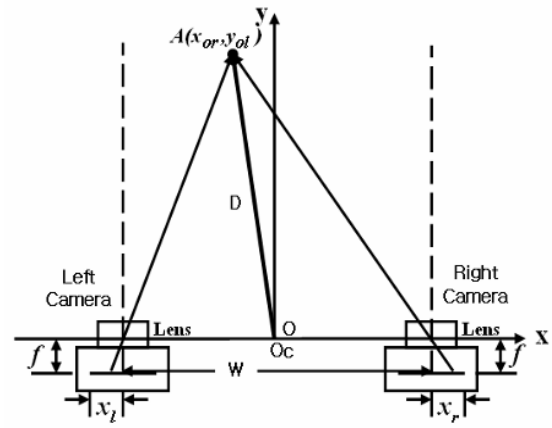


Figure 2. Structure of a parallel stereo camera

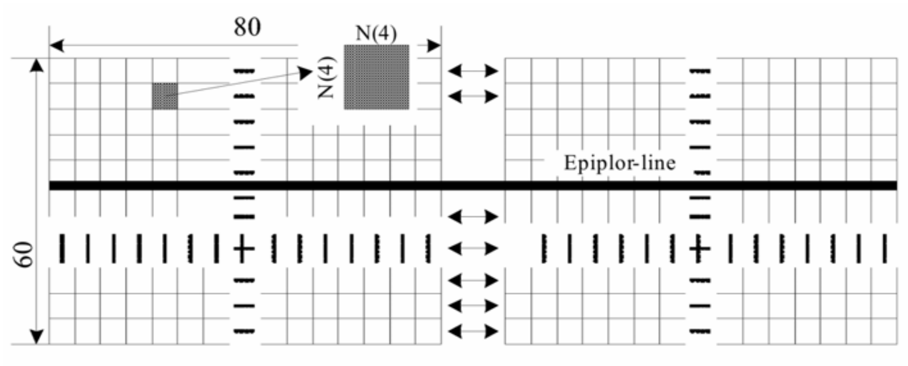


Figure 3. Disparity compensation of stereo images

The proposed block matching algorithm can remove unnecessary operations and the performance of the proposed block matching algorithm is as good as the one of the global searching algorithm. The process of the proposed algorithm is as following. First, SAD is calculated at each row and then the minimum value of SAD at the corresponding row is obtained as

$$SAD_{MIN}^{R} = MIN\left(\sum_{k}\left|\sum_{k=0}^{x=N}\sum_{j=0}^{y=N}\left|I_L(i,j)-I_R(i+k,j)\right|\right|\right), \qquad (2)$$

Finally, the minimum SAD of entire image can be obtained as

$$SAD_{MIN} = MIN\left(\sum SAD_{MIN}^{R}\right). \qquad (3)$$

Also, the disparity value between left and right images can be calculated as [2]

$$right^{*} = right_{t-k}\,,\; left^{*} = left_{t+k}\,. \tag{4}$$

2.2 Scaling of the face images according to the distance

320×240 RGB color images including face region are used as an input image. For fast processing and reducing the effect for illumination changes, the RGB input image is converted to *YCbCr* image. By defining the color range for Asian's face skin as $R_{Cb} = [77,127]$ and $R_{Cr} = [133,173]$, a color-based image segmentation [10] is performed as

$$S(x,y) = \begin{cases} 1, & if\,[Cb(x,y) \in R_{Cb}] \cap [Cr(x,y) \in R_{Cr}] \\ 0, & otherwise \end{cases}. \tag{5}$$

By using the camera characteristics as given in Table 1, the distance can be measured as

$$D = \frac{bf}{x_l - x_r} \times 86.80 \times 10^{3}[m]\,, \tag{6}$$

where b represents the width between cameras, f represents the focal length, and x_l and x_r respectively represent the distances of left and right images. Also, the constant of 86.80×103 represents the effective distance per pixel.

Item		Characteristic
Camera setting method		binocular
Camera setting width		65(㎜)
Camera focus length(f)		3.6(㎜)
Size	1 pixel	7.2×5.6(㎛)
Resolution	width	512(dots)

Table 1. Camera's component elements

For the 320x240 input images, the maximum distance of the disparity, $x_l - x_r$ is equal to 320, and the minimum distance is equal to 1. The scaling according to the change of distance [11] is performed as

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{7}$$

where x', y' represent the position after scaling processing, s_x, s_y represent the scaling factor, and x, y represent the current position. From the obtained distance in (6), the scaling factor of face image can be calculated as

$$V_x = (B_{dist} \times V_{dist}) / A_{dist}\,, \tag{8}$$

where B_{dist}, V_{dist}, and A_{dist}, and represent the basic distance, the established value by distance, and the obtained distance, respectively.

2.3 Range-based pose estimation using optimized 3D information

In order to solve the problem of the low recognition rate due to the uncertainty of size, distance, motion, rotation, and depth, optimized 3D information from stereo images is used. By estimating the position of eyes, the proposed method can estimate the facial size, depth, and pose change, accurately. The result of estimation of facial pose change is shown in Figure 4.

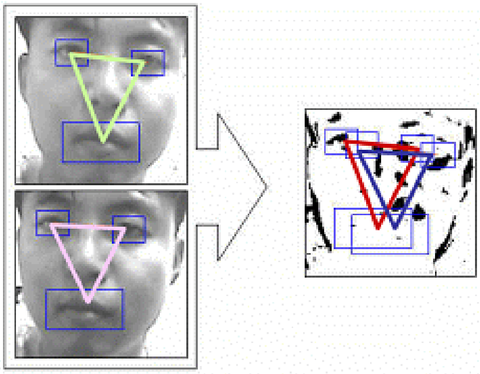


Figure 4. Estimation of face rotation

In Figure 4, the upper and lower images respectively represent the right image and the left image of frontal face. In Figure 5, the range of 9 directions for face images is defined to estimate the accurate facial direction and position of stereo images.

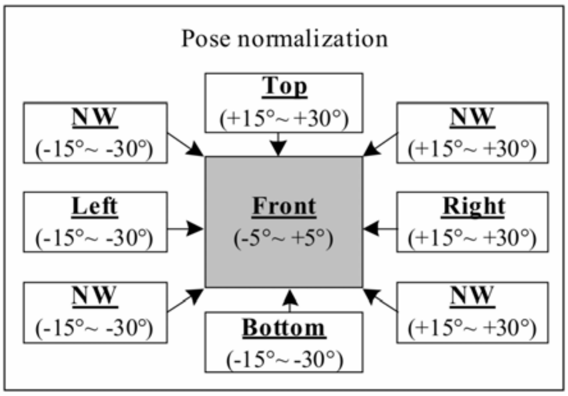
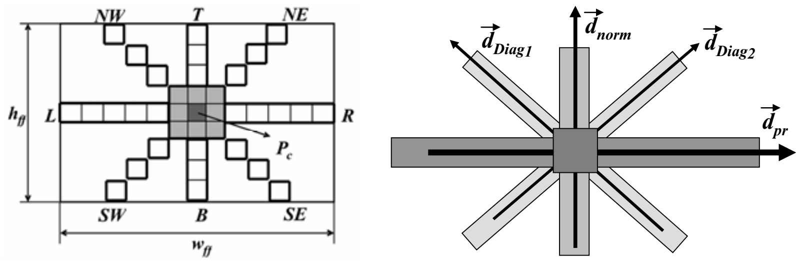


Figure 5. Range of face position according to direction

3. Pose estimation and face recognition

Face recognition rate is sensitive to illumination change, pose and expression change, and resolution of image. In order to increase the recognition rate under such conditions, we should consider the pose change as well as the frontal face image. The recognition rate can be increased by the 3D pose information as presented in Figure 5. In order to detect face region and estimate face elements, the multi-layered relative intensity map based on the face characteristics is used, which can provide better result than the method using only color images. The proposed directional blob template can be determined according to the face size. In detail, to fit for the ratio of the horizontal and vertical length of eyes, the template should be defined so that the length of horizontal axis is longer than that of vertical one as shown in Figure 6 (a). The central pixel of a template in a $W \times H$ image is defined as $P_c = (x_c, y_c)$. By using $W_{ff} \times H_{ff}$ directional template for face components, the average intensity $\overline{I_{Dir}}$ of *8*-neighborhood pixels is calculated in the central pixel, P_c. As a result, the brightness value at P_c, $\overline{I_C}$ and the brightness difference value can be obtained. The principal direction, $\overline{d_{pr}}$, and its magnitude, $|\overline{d_{pr}}|$, are determined as the direction including the biggest brightness difference as shown in Figure 6 (b).



(a) Directional template (b) New direction for map

Figure 6. Directional template for estimation of position for eyes and mouth

Figure 7 shows the result of the face region divided by the multi-layered relative intensity map. We can build the database including 92×112 face images at each direction. The directional range of face image can be classified into 9 groups as shown in Figure 6.

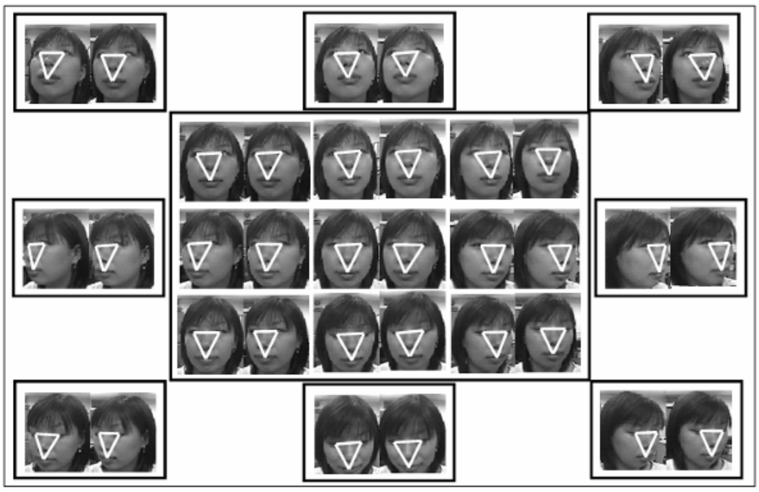


Figure 7. Face area division of multi-layered relative intensity map

The classified images are trained by PCA algorithm using optimized 3D information component. The block diagram of the proposed optimized PCA algorithm is shown in Figure 8.

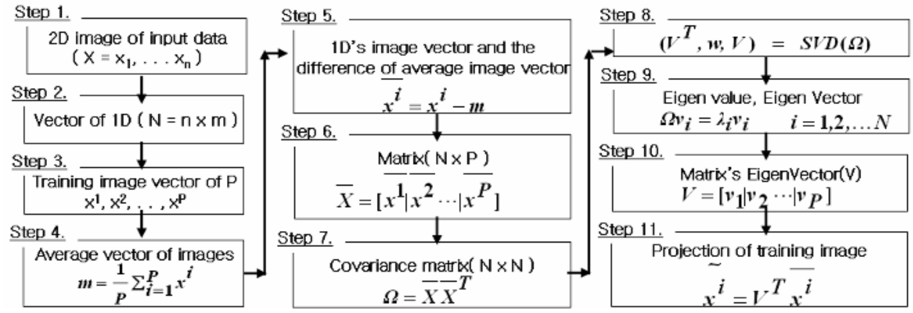


Figure 8. The block diagram of PCA algorithm

4. Experimental Results

For the experiment, we extracted 50 to 400 stereo pairs of face images of size 320×240. Figure 9 shows the matching result of the left and the right images captured in the distance of 43cm. Composed image shows Figure 9(c) which initializes 20×10 block in Figure 9(a), and is searched in the limited region of Figure 9(b). The disparity can be found in the most left and the top regions as shown in Figure 9(c). Facial pose estimation is performed with 9 directional groups at 100cm by using the proposed system as shown in Figure 10.

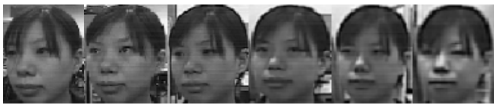
(a) left image (b) right image (c) matched image

Figure 9. The matching result of a stereo image pair

Figure 10. Detection results at stereo face images

Figs. 12 show the 92×112 scaled versions of the images captured at different distances. The scaling ratio of the captured face images was determined with respect to the reference image captured at the distance of 100cm. The scaling up ratios are respectively 1.2, 1.5, and 2.0 at the distances of 120cm, 150cm, and 200cm, while the scaling up ratios are 0.4 and 0.5 at the distances of 30cm and 50cm. The scaling factors were determined by experiment. Figs.13 show the samples of stereo image pairs used as input images. Figs. 14 show the some result images recognized by the proposed algorithm. The proposed algorithm can recognize the face as well as the pose of the face under pose changes.

(a) Left images



(b) Right images

Figure 11. The scaled version of the face images captured at the distance of 30, 50, 100, 120, 150, and 200cm

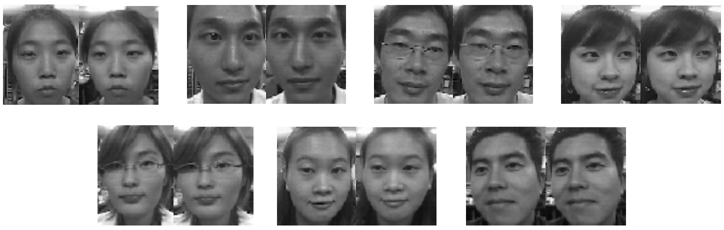


Figure 12. The samples of the input stereo image pairs

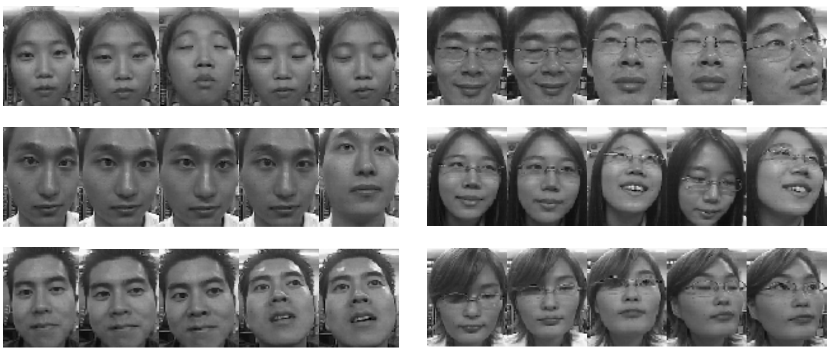


Figure 13. Various pose of the result images recognized by the proposed algorithm

In Table 2, the recognition rate is compared according to the distance. As shown in the Table 2, the highest recognition rate can be obtained at the reference distance of 100cm. After training 200 stereo images, the recognition rates of the proposed methods were compared to those of the existing methods with respect to 120 test images. The recognition rate of the proposed method based on optimized 3D information is provided in Figure 14. Experiment 1 and 2 respectively used frontal face images and images with various pose change. Figure 14 shows that the recognition rate using the conventional PCA or HMM drops in inverse proportion to the distance. From the experiments, the proposed method can increase the recognition rate.

No. of training images (L/R)	No. of test images	Recognition rate according to distance (%)					
		30 (cm)	50 (cm)	100 (cm)	120 (cm)	150 (cm)	200 (cm)
200/200	120	90.00	93.33	95.83	91.67	90.00	87.50

Table 2. The recognition rate according to the distance

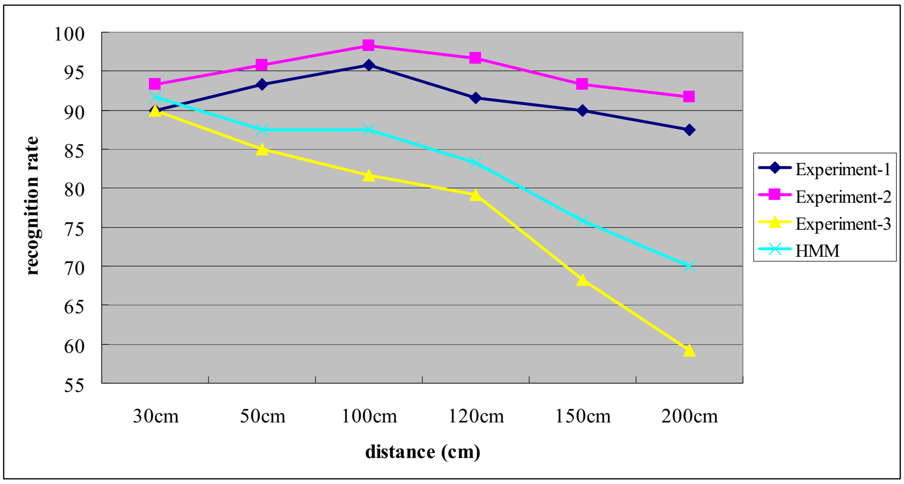


Figure 14. Recognition rates versus distance comparison for the proposed and various existing methods

5. Conclusions

This paper proposed a new range-based face detection and recognition method using optimized 3D information from stereo images. The proposed method can significantly improve the recognition rate and is robust against object's size, distance, motion, and depth using the PCA algorithm. The proposed method uses the *YCbCr* color format for fast, accurate detection of the face region. The proposed method can acquire more robust information against scale and rotation through scaling the detected face image according to the distance change. Experiments were performed in the range of 30~200cm and we could get the recognition rate up to 95.8% according to the scale change. Also, we could get the high recognition rate of 98.3% according to the pose change. Experimental results showed that the proposed method can increase the low recognition rate of the conventional 2D-based algorithm.

6. References

M. Yang, D. Kriegman, and N. Ahuja, Detecting faces in images: a survey, *IEEE Trans. Pattern Analysis, Machine Intelligence*, vol. 24, no. 1, pp. 34-58, January 2002. [1]

C. Park, I. Paek, and J. Paik, Improved face recognition using extended modular principal component analysis, *Proc. Int. Symposium, Visual Computing, LNCS*, vol. 4291. pp. 599-607, November 2006. [2]

Z. Sun, G. Bebis, X. Yuan, and S. J. Louis, Genetic feature subset selection for gender classification: A comparison study, *Proc. 6th IEEE Workshop Applications of Computer Vision (WACV 2002)*, pp. 165-170. December 2002. [3]

H. Rowley, S. Baluja, and T. Kanade, Neural Network-based face detection, *IEEE Trans. Pattern Analysis, Machine Intelligence*, vol. 20, no. 1, pp. 203-208, January 1998. [4]

E. Osuna, R. Freund, and F. Girosi, Training support vector machines: An application to face detection, *Proc. IEEE Computer Vision, Pattern Recognition*, pp. 130-136, June 1997. [5]

F. Samaria and S. Young, HMM based architecture for face identification, *Image, Vision Computing*, vol. 12, no. 8, pp. 537-543, October 1994. [6]

P. Belhumeur, J. Hespanha, and D. Kriegman, Eigenfaces vs fisherfaces: Recognition using class specification linear projection, *IEEE Trans. Pattern Analysis, Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997. [7]

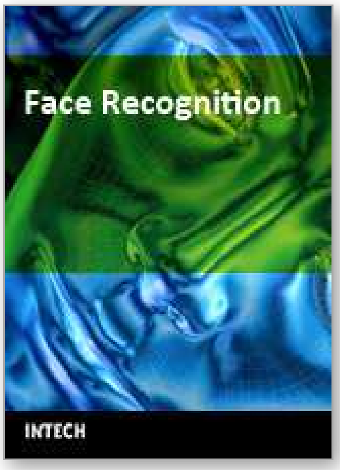
B. McLindin, Baselining illumination variables for improved facial recognition system performance, *Proc. 4th EURASIP Conf., Video/Image Processing, Multimedia Communications*, vol. 1, pp. 417-422, July 2003.[8]

L. Chen and W. Lin, Visual surface segmentation from stereo, *Image, Vision Computing*, vol. 15, no. 2, pp. 95-106, February 1997. [9]

D. Chai and K. Ngan, Face segmentation using skin-color map in videophone applications, *IEEE Trans. Circuits, Systems for Video Technology*, vol. 9, no. 4, pp. 551-564, June 1999, [10]

O. Faugeras, *Three-Dimensional Computer Vision*, 4th edition, MIT Press, 2001. [11]

W. IJsselsteijn, H. Ridder, and J. Vliegen, Subjective evaluation of stereoscopic images: effects of camera parameters and display duration, *IEEE Trans. Circuits, Systems for Video Technology*, vol. 10, no. 2, pp. 225-233, March 2000. [12]



Face Recognition

Edited by Kresimir Delac and Mislav Grgic

ISBN 978-3-902613-03-5
Hard cover, 558 pages
Publisher I-Tech Education and Publishing
Published online 01, July, 2007
Published in print edition July, 2007

This book will serve as a handbook for students, researchers and practitioners in the area of automatic (computer) face recognition and inspire some future research ideas by identifying potential research directions. The book consists of 28 chapters, each focusing on a certain aspect of the problem. Within every chapter the reader will be given an overview of background information on the subject at hand and in many cases a description of the authors' original proposed solution. The chapters in this book are sorted alphabetically, according to the first author's surname. They should give the reader a general idea where the current research efforts are heading, both within the face recognition area itself and in interdisciplinary approaches.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Changhan Park and Joonki Paik (2007). Face Recognition Using Optimized 3D Information from Stereo Images, Face Recognition, Kresimir Delac and Mislav Grgic (Ed.), ISBN: 978-3-902613-03-5, InTech, Available from: http://www.intechopen.com/books/face_recognition/face_recognition_using_optimized_3d_information_from_stereo_images

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821