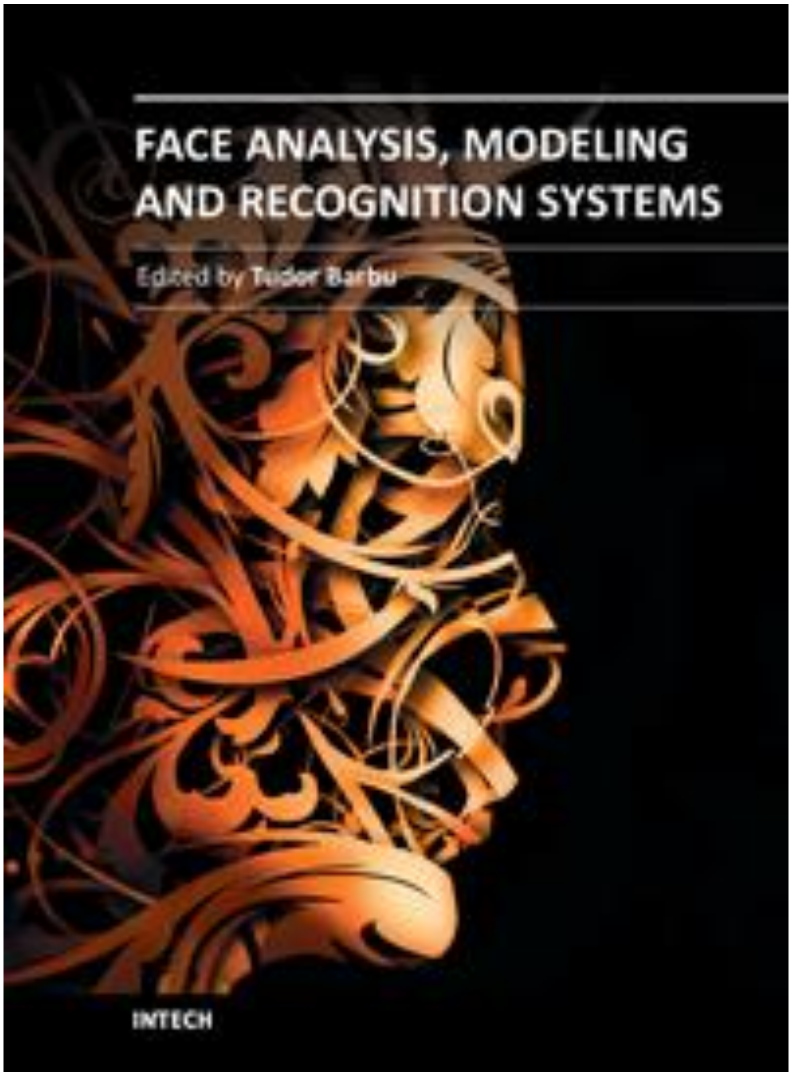


Edited by Tudor Barbu, ISBN 978-953-307-738-3, 224 pages, Publisher: InTech, Chapters published October 03, 2011 under CC BY-NC-SA 3.0 license
DOI: 10.5772/1837

Editor

Dr. Tudor Barbu

Romanian Academy, Iasi branch, Romania

Dr. Tudor Barbu is currently Senior Researcher II at the Institute of Computer Science of the Romanian Academy, Iași branch. He is the coordinator of an image and video processing research collective at this institute. Mr. Barbu has a PhD degree in Computer Science, awarded by the Faculty of Automatic Control and Computers of the University "Politehnica" of Bucharest. He possess a remarkable research profile. In the last decade he published two books and four book chapters as single or main author. Also, dr. Tudor Barbu published more than 60 articles in prestigeous international journals and volumes of international scientific events (conferences, symposiums and workshops). His prolific scientific activity also includes more than 30 research reports, elaborated with the institute research team coordinated by him or related to various research projects. His scientific publications have numerous citations, according to Google-Academic. In recent years he also coordinated various research directions in 6 projects based on contracts/grants. Dr. Tudor Barbu received also several awards for his research results, the most important being the Romanian Academy Prize "Gheorghe Cartianu", in the Information Science and Technology domain, awarded on December 18, 2008. He is a member of several conference scientific committees and also member of scientific and technical committee and editorial review boards of some journals. His main scientific areas of interest are: digital media (audio, video and image) signal processing and analysis, pattern recognition, computer vision, multimedia information storage, indexing and retrieval, and biometric authentication using voice, face and digital fingerprint recognition.

Book Content:

1

An Improved Face Recognition System for Service Robot Using Stereo Vision

Widodo Budiharto[1], Ari Santoso[2], Djoko Purwanto[2] and Achmad Jazidie[2]
[1]*Dept. of Informatics Engineering, BINUS University, Jakarta*
[2]*Dept. of Electrical Engineering, Institute of Technology Sepuluh Nopember, Surabaya*
Indonesia

1. Introduction

Service robot is an emerging technology in robot vision, and demand from household and industry will be increased significantly in the future. General vision-based service robot should recognizes people and obstacles in dynamic environment and accomplishes a specific task given by a user. The ability to face recognition and natural interaction with a user are the important factors for developing service robots. Since tracking of a human face and face recognition are an essential function for a service robot, many researcher have developed face-tracking mechanism for the robot (Yang M., 2002) and face recognition system for service robot(Budiharto, W., 2010).

The objective of this chapter is to propose an improved face recognition system using PCA(Principal Component Analysis) and implemented to a service robot in dynamic environment using stereo vision. The variation in illumination is one of the main challenging problem for face recognition. It has been proven that in face recognition, differences caused by illumination variations are more significant than differences between individuals (Adini et al., 1997). Recognizing face reliably across changes in pose and illumination using PCA has proved to be a much harder problem because eigenfaces method comparing the intensity of the pixel. To solve this problem, we have improved the training images by generate random value for varying the intensity of the face images.

We proposed an architecture of service robot and database for face recognition system. A navigation system for this service robot and depth estimation using stereo vision for measuring distance of moving obstacles are introduced. The obstacle avoidance problem is formulated using decision theory, prior and posterior distribution and loss function to determine an optimal response based on inaccurate sensor data. Based on experiments, by using 3 images per person with 3 poses (frontal, left and right) and giving training images with varying illumination, it improves the success rate for recognition. Our proposed method very fast and successfully implemented to service robot called Srikandi III in our laboratory.

This chapter is organized as follows. Improved method and a framework for face recognition system is introduced in section 2. In section 3, the system for face detection and depth estimation for distance measurement of moving obstacles are introduced. Section 4, a detailed implementation of improved face recognition for service robot using stereo vision is presented. Finally, discussions and future work are drawn in section 5.

2. Improved face recognition system using PCA

The face is our primary focus of attention in developing a vision based service robot to serves peoples. Unfortunatelly, developing a computational model of face recognition is quite difficult, because faces are complex, meaningful visual stimuli and multidimensional. Modelling of face images can be based on statistical model such as Principal Component Analysis (PCA) (Turk & Pentland, 1991) and Linear Discriminat analysis (LDA) (Etemad & Chellappa, 1997; Belhumeur et.al, 1997), and physical modelling based on the assumption of certain surface reflectance properties, such as Lambertian surface (Zoue et al., 2007). Linear Discriminant Analysis (LDA) is a method of finding such a linear combination of variables which best separates two or more classes. Constrasting ther PCA which encodes information in an orthogonal linear space, the LDA which also known as fischerfaces method encodes discriminatory information in a linear separable space of which bases are not necessary orthogonal. However, the LDA result is mostly used as part of a linear classifier (Zhao et al., 1998).
PCA is a standard statistical method for feature extraction by reduces the dimension of input data by a linear projection that maximizes the scatter of all projected samples. The scheme is based on an information theory approach that decomposes faces images into a small set of characteristic feature images called eigenfaces, as the principal components of the initial training set of face images. Recognition is performed by projecting a new image into the subspace spanned by the eigenfaces called face space, and then classifying the face by comparing its position in face space with the positions of known individuals. PCA based approaches typically include two phases: training and classification. In the training phase, an eignespace is established from the training samples using PCA and the training face images are mapped to the eigenspace for classification. In the classification phase, an input face is projected to the same eignespace and classified by an appropriate classifier (Turk & Pentland, 1991). Let a face image $I(x, y)$ be a two-dimensional N by N array of (8-bit) intensity values. An image may also be considered as a vector of dimension N^2, so that a typical image of size 256 by 256 becomes a vector of dimension 65,536 (a point in 65,536-dimensional space). If Φ is face images and M training set face images, we can compute the eigenspace u_i :

$$u_i = \sum_{k=1}^{M} v_{ik}\Phi_k \quad i = 1,2,\ldots,M \tag{1}$$

Where u_i and v_{ik} are the $\mathrm{i^{th}}$ eigenspace and the $\mathrm{k^{th}}$ value of the $\mathrm{i^{th}}$ eigenvector. Then, we can determining which face class provides the best description of an input face images to find the face class k by using the euclidian distance ε_k between the new face projection Ω, the class projection Ω_k and threshold θ using formula :

$$\varepsilon_k = \|\Omega - \Omega_k\| < \theta \tag{2}$$

The stereo camera used in this research is 640x480 pixels. The size of face image is cropped to 92x112pixels using Region of Interest method (ROI) as shown in figure below. These images also used as training images for face recognition system. We use histogram equalization for contrast adjustment using the image's histogram. This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values.

(a) (b) (c)

Fig. 1. Original image (a) , then applied preprocessing image to greyscale image (b) and histogram equalization (c).

The illumination variation is one of the main challenging problems that a practical face recognition system needs to deal with. Various methods have been proposed to solve the problem, named as face and illumination modeling, illumination invariant feature extraction and preprocessing and normalization. In (Belhumeur & Kriegman 1998), an illumination model illumination cone is proposed for the first time. The authors proved that the set of n-pixel images of a convex object with a Lambertian reflectance function, under an arbitrary number of point light sources at infinity, formed a convex polyhedral cone in IR^n named as illumination cone (Belhumeur & Kriegman 1998). In this research, we construct images under different illumination conditions by generate a random value for brightness level developed using Visual C++ technical Pack using this formula :

$$I_o(x, y) = I_i(x, y) + c \tag{3}$$

Where I_o is the intensity value after brightness operation applied, I_i is the intensity value before brightness operation and c is a brightness level. The effect of brightness level shown at histogram below:

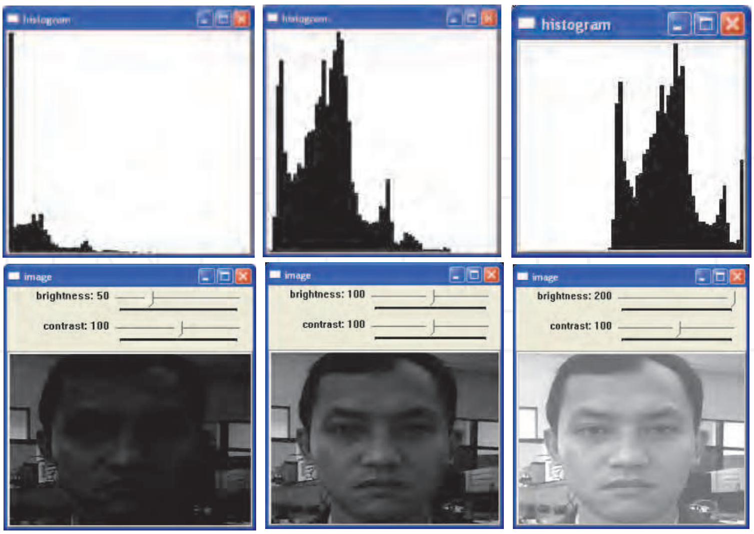


Fig. 2. Effect of varying the illumination for a face.

We have developed a Framework of Face recognition system for vision-based service robot. This framework very usefull as a information for robot to identify a customer and what items ordered by a customer. First, to storing training faces of customers, we have proposed a database for face recognition that consists of a table faces, products and order. An application interface for this database shown below :

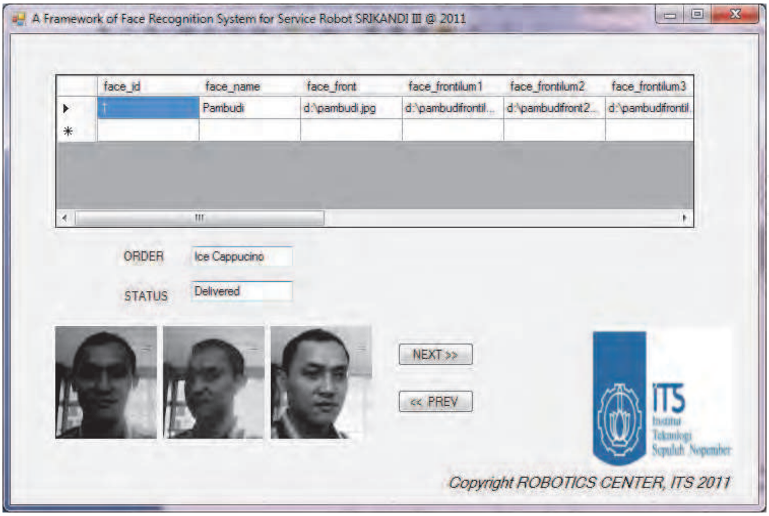


Fig. 3. We have proposed face databases using 1 table, 3 images used for each person (frontal, left and right poses).

We have identified the effect varying illumination to the accuracy of recognition for our database called ITS face database as shown in table 1 :

Training images	**Testing images**	**Success rate**
No varying illumination		
6	6	100%
12	6	100%
24	6	100%
Varying Illumination		
6	6	50.00%
12	6	66.00%
24	6	100%
24	10	91.60%

Table 1. Testing images without and with varying illumination. Results shows that by giving enough training images with variation of illumination generated randomly, the success rate of face recognition will be improved.

We also evaluate the result of our proposed face recognition system and compared with ATT and Indian face database using Face Recognition Evaluator developed by Matlab. Each of face database consists of 10 sets of people's face. Each set of ITS face database consists of 3 poses (front, left, right) and varied with illumination. ATT face database consists of 9 differential facial expression and small occlusion (by glass) without variation of illumination . The Indian face database consists of eleven pose orientation without variation of illumination and the size of each image is too small than ITS and ATT face database. The success rate comparison between 3 face databases shown below:

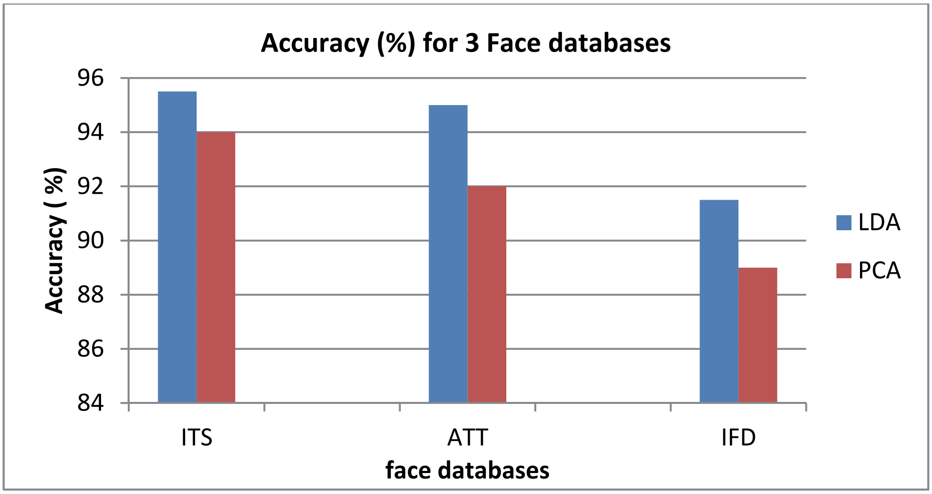


Fig. 4. Success rate comparison of face recognition between 3 faces databases, each using 10 sets face. It shown clearly that ITS database have highest success rate than ATT and Indian face database when the illumination of testing images is varied. The success rate using PCA in our proposed method and ITS face database is 95.5 %, higher than ATT face database 95.4%.

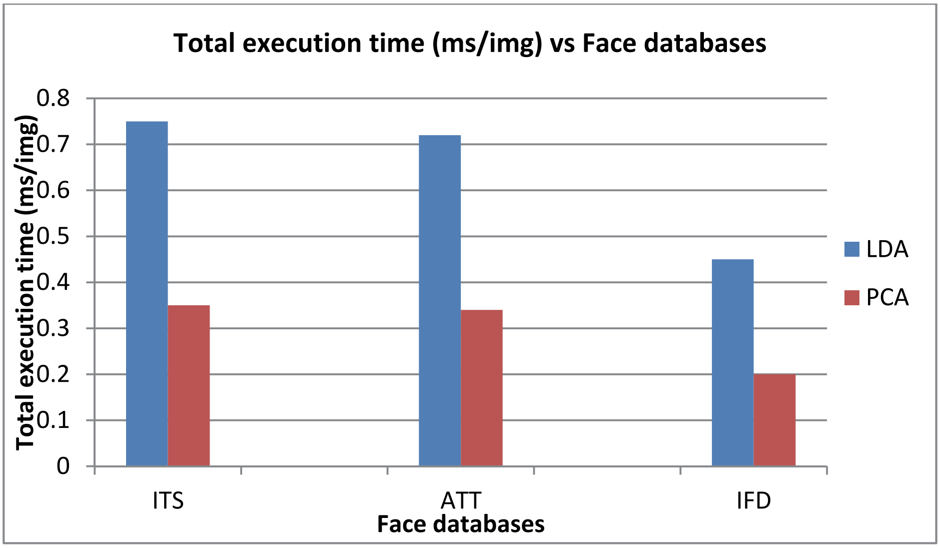


Fig. 5. For total execution time, we can see the Indian face database (IFD) is shortest because the size of each image is lowest then ITS and ATT.

3. Face detection and depth estimation using stereo vision

We have developed a system for face detection using Haar cascade classifier and depth estimation for measuring distance of peoples as moving obstacles using stereo vision. The camera used is a Minoru 3D stereo camera. The reason for using stereo camera in order robot able to estimate distance to obstacle without additional sensor(only 1 ultrasonic sensor infront of robot for emergency), so the cost for development can be reduced. Let's start froma basic conce where a point q captured by camera, the point in the front image frame ${}^{I}p({}^{I}p_x,{}^{I}p_y)$ is the projection of the point in camera frame ${}^{C}p({}^{C}p_x,{}^{C}p_y,{}^{C}p_z)$ onto the front image frame. Here, f denotes the focal length of the lens. Fig. 6 shown is the projection of a point on the front image frame.

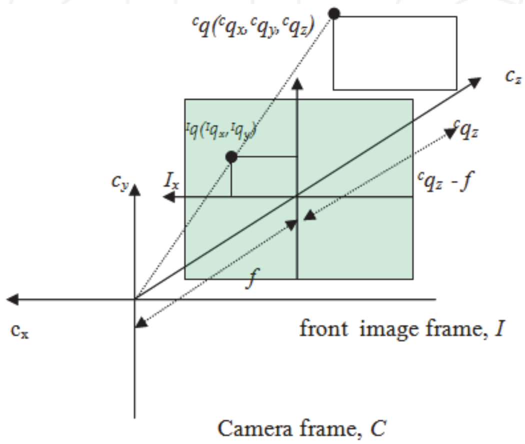


Fig. 6. Projection of point on front image frame.

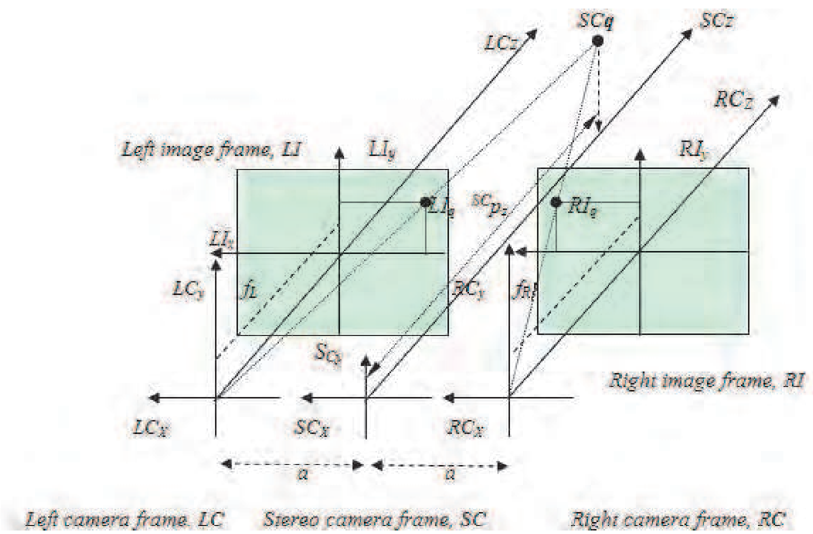


Fig. 7. Stereo Imaging model

In the stereo imaging model, the tree-dimensional points in stereo camera frame are projected in the left and the right image frame. On the contrary, using the projection of the points onto the left and right image frame, the three-dimensional points positions in stereo camera frame can be located. Fig. 7 shows the stereo imaging model using the left front image frame *LF* and right front image frame *RF* (Purwanto, D., 2001).
By using stereo vision, we can obtain the position of each moving obstacle in the images, then we can calculate and estimate the distance of the moving obstacle. The three-dimensional point in stereo camera frame can be reconstructed using the two-dimensional projection of point in left front image frame and in right front image frame using formula:

$$
{}^{SC}\mathbf{q} = \begin{bmatrix} {}^{SC}q_x \\ {}^{SC}q_y \\ {}^{SC}q_z \end{bmatrix} = \frac{2}{{}^{RI}q_x - {}^{LI}q_x} \begin{bmatrix} \frac{1}{2} a\left({}^{RI}q_x + {}^{LI}q_x\right) \\ a\,{}^{RI}q_y \\ f\,a \end{bmatrix} \tag{4}
$$

Note that ${}^{LI}q_y = {}^{RI}q_y$

Figure shown below is the result of 2 moving obstacle identification using stereo vision, distance of obstacle obtained using depth estimation based on eq. 4. State estimation is used for handling inaccurate vision sensor, we adopted it using Bayesian approach for probability of obstacle denoted as *p(Obstacle)* and probability of direction *p(Direction*) with the *value* between 0-1.

Fig. 8. Robot succesfully identified and estimated distance of 2 moving obstacles in front of robot.

4. Implementation to vision-based service robot

4.1 Architecture of service robot

We have developed a vision-based service robot called Srikandi III with the ability to face recognition and avoid people as moving obstacles, this wheeled robot is next generation from Srikandi II (Budiharto, W. 2010). A mobile robot involving two actuator wheels is considered as a system subject to nonholonomic constraints. Consider an autonomous wheeled mobile robot and position in the Cartesian frame of coordinates shown in fig. 10, where x_R and y_R are the two coordinates of the origin **P** of the moving frame and θ_R is the robot orientation angle with respect to the positive x-axis. The rotation angle of the right and left wheel denoted as φ_r and φ_l and radius of the wheel by R thus the configuration of the mobile robot c_R can be described by five generalized coordinates such as :

$$c_R = (x_R, y_R, \theta_R, \varphi_r, \varphi_l)^T \tag{5}$$

Based on fig, 10, v_R is the linear velocity, ω_R is the angular velocity, r_R and λ_R are radial and angular coordinate of the robot (Mahesian, 2007). The kinematics equations of motion for the robot given by :

$$\dot{x}_R = v_R \cos\theta_R \tag{6}$$

$$\dot{y}_R = v_R \sin\theta_R \tag{7}$$

$$\dot{\theta}_R = \omega_R \tag{8}$$

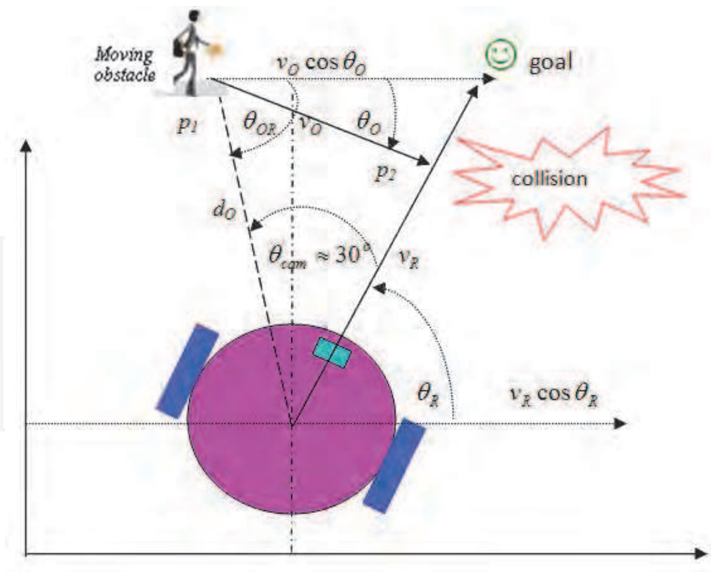


Fig. 9. Proposed Cartesian model of mobile robot with moving obstacle

The angular velocity of the right and left wheel can be obtained by :

$$\omega_r = \frac{d\varphi_r}{dt} \text{ and } \omega_l = \frac{d\varphi_l}{dt} \tag{9}$$

Finally, the linear velocity v_R can be formulated as :

$$v_R = R(\omega_r + \omega_l) / 2 \tag{10}$$

Camera become important sensor if we want to identify specific object such as face, small object, shape etc) that could not identified by other sensor such as ultrasonic sensors. Camera as a vision sensor have limitation in angle area for capturing object. We defined θ_{cam} as a maximum angle that moving obstacle can be detected by camera used in this research. Based on the fig. 1, we defined angle between moving obstacle and robot θ_{OR} as:

$$\theta_{OR} = 180^\circ - (\theta_R + \theta_{cam}) \tag{11}$$

$\theta_O, \theta_{OR}, \theta_{Cam}, \theta_R, v_R$ and v_O are very important properties for calculating whether robot will collides or not with moving obstacle. To calculate the speed of moving obstacle v_O based on vision is a complex task, we propose the model for calculate the v_O that moving with angle θ_O detected by the camera, whereas at the same time the robot moving with speed v_R to the goal with angle θ_R, we need 2 point of tracked images with interval t = 1 second, then the difference of pixel position obtained.

Based on the fig. 10, the equation for estimates v_O when moving obstacle and robot are moving is :

$$v_O \cos\theta_O = \frac{|p_2 - p_1| s}{t} + v_R \cos\theta_R \tag{12}$$

Finally, we can simplified the eq. 12 as :

$$v_O = \frac{|p_2 - p_1| s}{t \ \cos\theta_O} + \frac{v_R \cos\theta_R}{\cos\theta_O} \tag{13}$$

Where p_1 and p_2 are the position of the obstacle in pixel and s is the scaling factor in cm/pixel. We proposed mechanism for predicting collision using time t needed for robot to collides the moving obstacle that move with orientation θ_O as shown in fig, 1 and should be greater than threshold T for robot to allowing moving forward, can be calculated by formula:

$$t = \frac{d_O \sin\theta_{OR}}{(v_R \sin\theta_R + v_O \sin\theta_O)} \tag{14}$$

Note: if $t <= T$ then robot stop
if $t > T$ then robot moving forward

Fig. 10 shown below is an architecture of service robot Srikandi III that utilizing stereo camera, compass and distance sensors. Because this robot need to recognizes and tracks people, many supporting functions developed and integrated such as face recognition system, static and moving obstacles detection and moving obstacle tracking, to make the robot robust and reliable. We developed efficient Faces database used by face recognition system for recognizing customer. There is interface program between Laptop for coordinating robot controller. 1 controller using Propeller used for coordinating actuator and communication with the and used for distance measurement. Srikandi III implements path planning based on the free area obtained from the landmark by edge and smoothing operation.

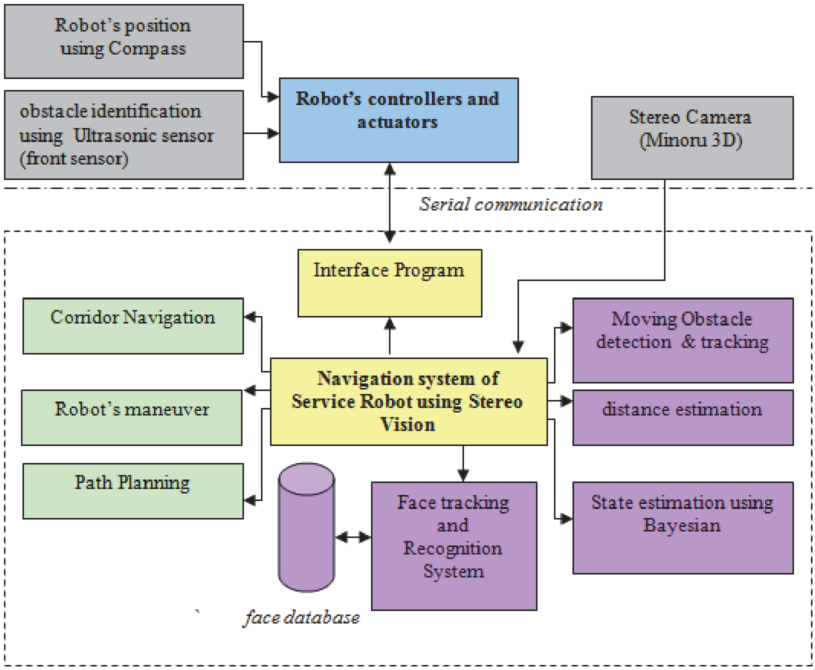


Fig. 10. General architecture of service robot called Srikandi III. Hardware and software parts are separated by the dashed line. All arrows indicate dataflow.

Because of the limitation of stereo camera used for distance measurement, Kalman filtering applied to make sure the measurement of distance between robot and obstacle more stable. The prototype of service Robot Srikandi III that utilized a low cost stereo camera using Minoru 3D is shown in fig. 11:

4.3 Proposed navigation system of vision-based service robot

4.3.1 Flow chart of a navigation system

The service robot should navigates from start to goal position and go back to home savely. We assumed the when robot running, people as moving obstacle may collides with the robot. So we proposed a method for obstacles avoidance for Service robot in general as shown in Fig. 12. The model of experiment for customer identification is using stereo camera, the advantage is we can estimate the distance of customer/obstacles and direction's movement of obstacles. There is no map or line tracking to direct a robot to an identified customer. Image captured by stereo camera used as testing images to be processed by Haar classifier to detect how many people in the images, and face recognition by PCA. We

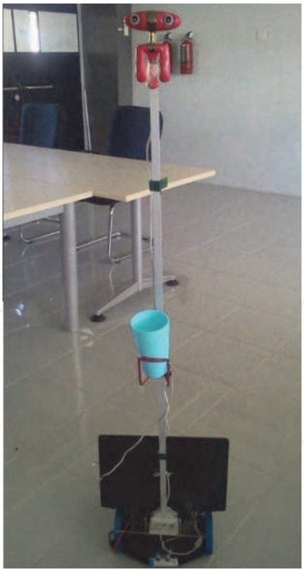


Fig. 11. Prototype of Service robot Srikandi III using stereo camera

implementing visual tracking to heading a robot to a customer. Robot continuously measures the distance of obstacle and send the data to Laptop. The next step is multiple moving obstacle detection and tracking. If there is no moving obstacle, robot run from start to goal position in normal speed. If moving obstacle appeared and collision will occurred, robot will maneuver to avoids obstacle.

Figure shown below is a flowchart that describes general mechanism for our method for detecting multiple moving obstacle and maneuver to avoids collision with the obstacles.

To implement the flowchart above for service robot that should recognize the customer and have the ability for multiple moving obstacle avoidance, we have developed algorithm and programs consist of 3 main modules such as a framework for face recognition system, multiple moving obstacle detection and Kalman filtering as state estimator for distance measurement using stereo camera.

4.3.2 Probabilistic robotics for navigation system

Camera as vision sensor sometimes have distortion, so Bayesian decision theory used to state estimation and determine the optimal response for the robot based on inaccurate sensor data. Bayesian decision rule probabilistically estimate a dynamic system state from noisy observations. Examples of measurement data include camera images and range scan. If x is a quantity that we would like to infer from y, the probability $p(x)$ will be referred to as prior probability distribution. The Bayesian update formula is applied to determine the new posterior $p(x, y)$ whenever a new observation is obtained:

$$p(x,y) = \frac{p(y|x,z)p(x|z)}{p(y|z)} \tag{15}$$

To apply Bayesian decision theory for obstacle avoidance, we consider the appearance of an unexpected obstacle to be a random event, and optimal solution for avoiding obstacles is obtained by trading between maneuver and stop action. If we want service robot should stay on the path in any case, strategies to avoid moving obstacle include:

- Maneuver, if service robot will collides.
- stop, if moving obstacle too close to robot.

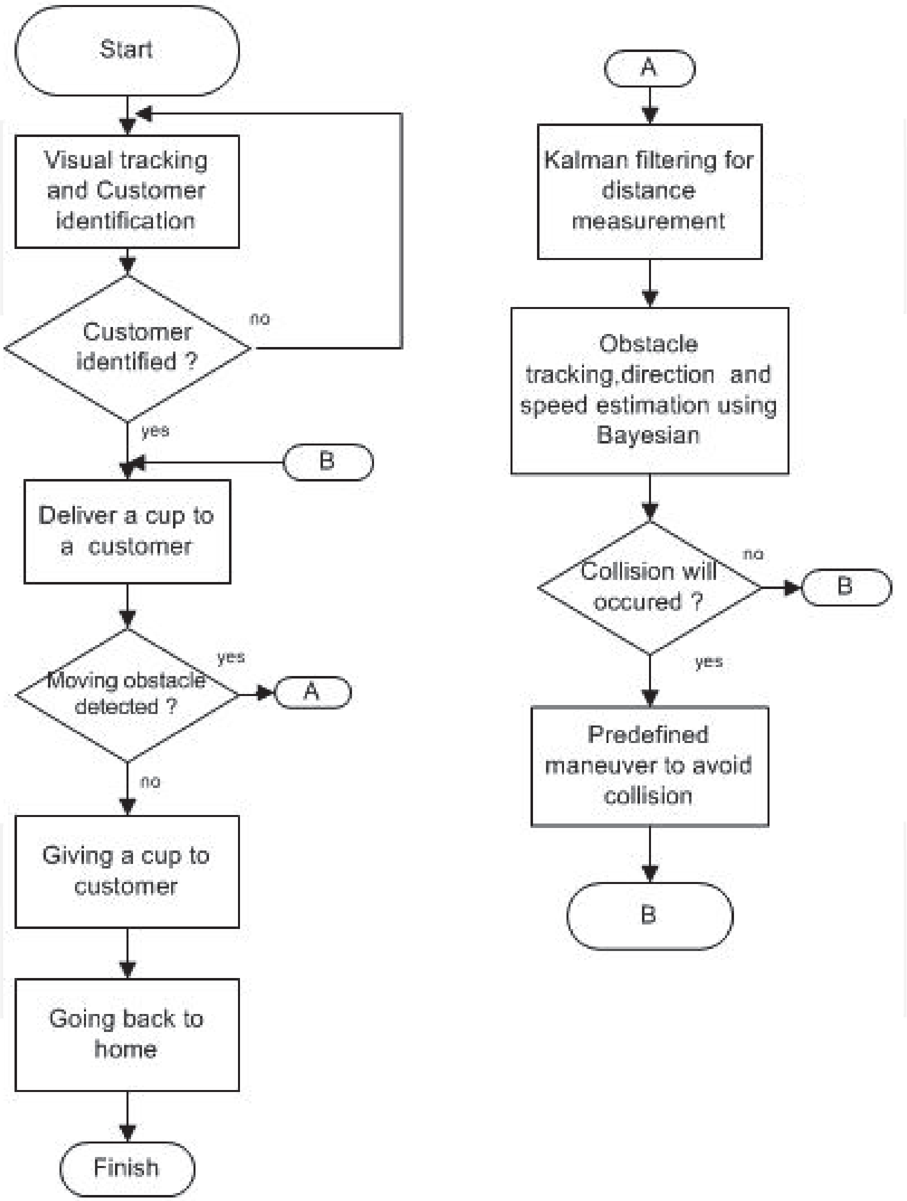


Fig. 12. Flow chart of Navigation System from start to goal position for service robot.

Then, we restrict the action space denoted as **A** as :

$$\mathrm{A} = (a_{1,} a_2, a_3) \tag{16}$$

$$= \text{maneuver to left, maneuver to right, stop} \tag{17}$$

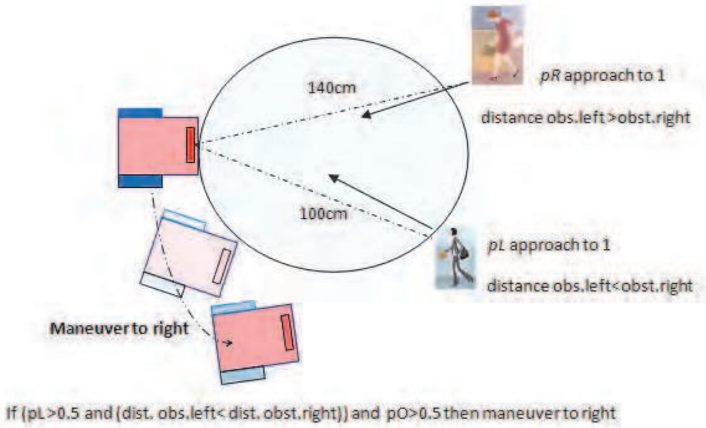
We define a loss function *L(a, θ)* which gives a measure of the loss incurred in taking action *a* when the state is θ. The robot should chooses an action a from the set **A** of possible actions based on the observation **z** of the current state of the path θ. This gives the posterior distribution of θ as:

$$p(\theta \mid z) = \frac{p(z \mid \theta) p(\theta)}{\sum p(z \mid \theta) p(\theta)} \tag{18}$$

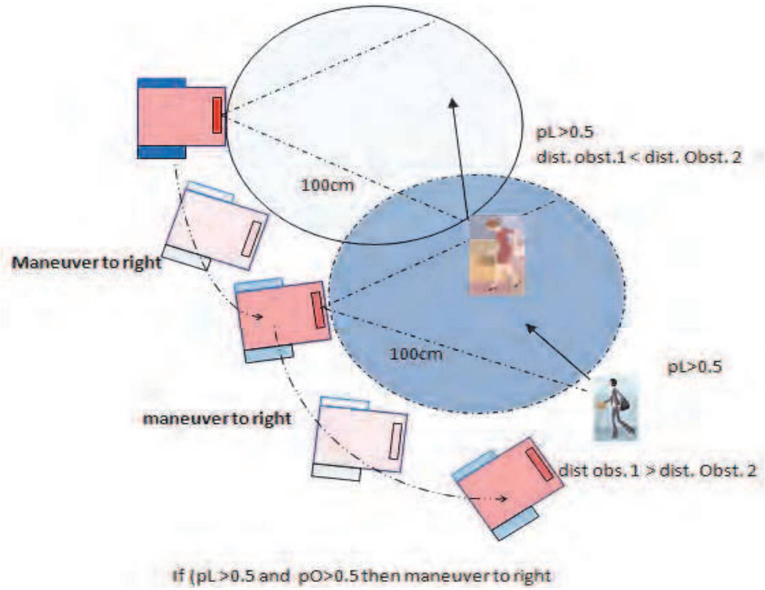
Then, based on the posterior distribution in (17), we can compute the posterior expected loss of an action (Hu, H et al., 1994):

$$B(p(\theta \mid z), a) = \sum_{\theta} L(\theta, a) p(\theta \mid z) \tag{19}$$

The figure below shows the proposed model of maneuvering on the service robot, *pL* which is the probability of moving obstacle leads to the left, and *pR* the probability of moving obstacle leads to the right. By estimating the direction of motion of the obstacle, then the most appropriate action to avoid to the right / left side can be determined, to minimize collisions with these obstacles. If there are more than 1 moving obstacle, then robot should identified the nearest moving obstacle to avoid it, and the direction of maneuver should be opposite with the direction of moving obstacle.



(a)



(b)

Fig. 13. A maneuvering model to avoids multiple moving obstacle using stereo vision, 2 multiple moving obstacle with the different direction (a) and the same direction (b)

Result of simulation using improved face recognition system and implemented to a service robot to identify a customer shown in figure 14. In this scheme, robot will track the face of a customer until the robot heading exactly to a customer, after that, robot will run to customer. If there are moving obstacles, robot will maneuver to avoid the obstacle.

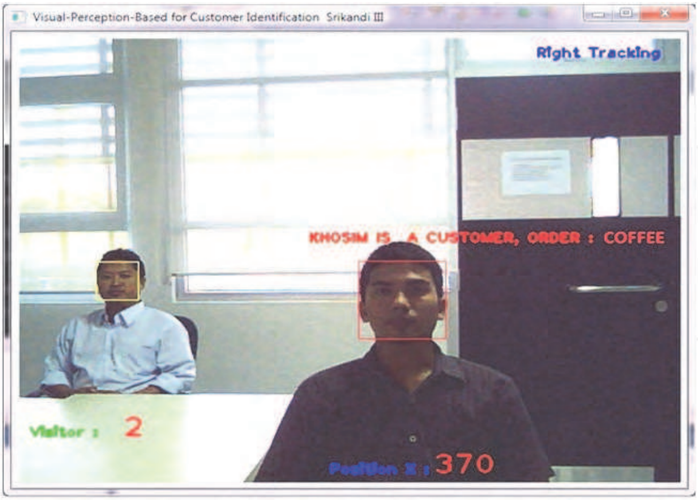


Fig. 14. Result of simulation using improved face recognition system and implemented to a service robot to identify a customer.

5. Discussion

This chapter presents an improved face recognition system using PCA and implemented to a service robot in dynamic environment using stereo vision. By varying illumination in training images, it will increase the success rate in face recognition. The success rate using our proposed method using ITS face database is 95.5 %, higher than ATT face database 95.4%. The simple face database system propsed can be used for the vision-based service robot. Experimental results with various situations have shown that the proposed methods and algorithms working well and robot reaches the goal points while avoiding moving obstacle. Estimation of distance of moving obstacle obtained by stereo vision. Bayesian decision rule implemented for state estimation makes this method more robust because the optimal solution for avoiding obstacles is obtained by trading between maneuver and stop action. In future work, we will implementing this system and develop a Vision-based humanoid service robot for serving customers at Cafe/Restaurants.

6. Acknowledgment

Research described in this paper was done in the Robotics Center, Institute of Technology Sepuluh Nopember (ITS) Surabaya, Indonesia. Part of this research also funded by JASSO Scholarship and conducted at Robotics Lab, Graduate school of Science and Technology, Kumamoto Universty-Japan.

7. References

Adini, Y.; Moses, Y. & Ulman, S. (1997). Face Recognition : the problem of compensating for changes in illumination direction, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 19, no. 7, 721-732, ISSN : 0162-8828

Belhumeur, P. & Kriegman, D. (1998). What is the set of images of an object under all possible illumination conditions, *International Journal of Computer Vision, Vol. 28, NO. 3, 245-260*, ISSN : 1573-1405 (electronic version).

Etemad, K. & Chellappa R (1997). Discriminant analysis for recognition of human face images*, Journal of the Optical Society of America A*, Vol. 14, No. 8, 1724-1733, ISSN : 1520-8532 (online).

Borenstein, J. & Koren, Y.(1991). Potential Field Methods and Their Inherent Limitations for Mobile Robot Navigation, *Proceeding IEEE Conf. on Robotics and Automation*, California, pp.1398-1404.

Budiharto, W., Purwanto, D. & Jazidie, A. (2011), A Robust Obstacle Avoidance for Service Robot using Bayesian Approach . *International Journal of Advanced Robotic Systems,* Vol. 8, No.1, (March 2011), pp. 52-60, ISSN 1729-8806.

Budiharto, W.; Purwanto, D. & Jazidie, A. (2010), A Novel Method for Static and Moving Obstacle Avoidance for Service robot using Bayesian Filtering, *Proceeding of IEEE 2nd International conf. on Advances in Computing, Control and Telecommunications Technology*, Jakarta-Indonesia, pp. 156-160. DOI: 10.1109/ACT.2010.51.

Fatih, K.; Binnur, K. & Muhittin, G. (2010). Robust Face Alignment for Illumination and Pose Invariant Face Recognition, Face Recognition, ISBN: 978-953-307-060-5.

Hu, H. & Brady, M. (1994). A Bayesian Approach to Real-Time Obstacle Avoidance for a Mobile Robot. Autonomous Robots, vol. 1, Kluwer Academic Publishers, Boston, pp. 69-92.

Khatib, O.(1986) Real-time Obstacle Avoidance for Manipulator and Mobile Robots, *International Journal of Robotics Research*, vol. 5 no. 1, pp.90-98.

Masehian, E. & Katebi, Y. (2007). Robot Motion Planning in Dynamic Environments with Moving Obstacles and Target, *Int. Journal of Mechanical Systems Science and Engineering*, 1(1), pp. 20-25.

Purwanto, D. (2001). Visual Feedback Control in Multi-Degrees-of-Freedom Motion System, Dissertation at Graduate School of Science and Technology - Keio University, Japan.

Turk, M. & Pentland A. (1991). Eigenfaces for recognition, *International Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86.

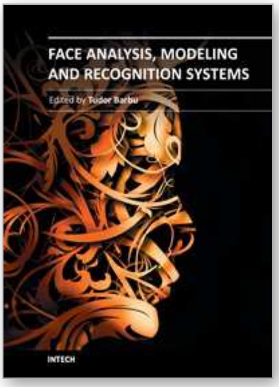
Yang M. (2002) Detecting faces images: A survey, IEEE Transactions on Pattern Analysis and Machine Inteligence, vol. 24 no. 1, pp.34-58

Zhichao, L. & Meng Joo E. (2010). Face Recognition Under Varying Illumination, *New Trends in Technologies: Control, Management, Computational Intelligence and Network Systems*, ISBN: 978-953-307-213-5.

OpenCV (2010). www.opencv.org.

ATT face database, http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

Indian face database (2010). http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase



Face Analysis, Modeling and Recognition Systems
Edited by Dr. Tudor Barbu

ISBN 978-953-307-738-3
Hard cover, 212 pages
Publisher InTech
Published online 30, September, 2011
Published in print edition September, 2011

The purpose of this book, entitled Face Analysis, Modeling and Recognition Systems is to provide a concise and comprehensive coverage of artificial face recognition domain across four major areas of interest: biometrics, robotics, image databases and cognitive models. Our book aims to provide the reader with current state-of-the-art in these domains. The book is composed of 12 chapters which are grouped in four sections. The chapters in this book describe numerous novel face analysis techniques and approach many unsolved issues. The authors who contributed to this book work as professors and researchers at important institutions across the globe, and are recognized experts in the scientific fields approached here. The topics in this book cover a wide range of issues related to face analysis and here are offered many solutions to open issues. We anticipate that this book will be of special interest to researchers and academics interested in computer vision, biometrics, image processing, pattern recognition and medical diagnosis.

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

2

Two Novel Face Recognition Approaches

Tudor Barbu
Institute of Computer Science of the Romanian Academy, Iasi branch, Romania

1. Introduction

Face recognition represents a very important biometric domain, the human face being a psychological biometric identifier that is widely used in person authentication. Also, it constitutes a popular computer vision domain, facial recognition being the most successful application of object recognition. Recognizing of faces is a task performed easily by humans but it remains a difficult problem in the computer vision area. Automated face recognition constitutes a relatively new concept, having a history of some 20 years of research. Major initiatives and achievements in the past decades have propelled facial recognition technology into the spotlight (Zhao et. al, 2003).

A facial recognition system represents a computer-driven application for automatically authenticating a person from a digital image, using the characteristics of its face. As any biometric recognition system, it performs two essential processes: identification and verification. Facial identification consists in assigning an input face image to a known person, while face verification consists in accepting or rejecting the previously detected identity. Also, facial identification is composed of a feature extraction stage and a classification step.

Face recognition technologies have a variety of application areas, such as: access control systems for various services, surveillance systems and law enforcement (Zhao et. al, 2003).. Also, these technologies can be incorporated into more complex biometric systems, to obtain a better human recognition. Facial recognition techniques are divided into two major categories: geometric and photometric methods. Geometric approaches represent feature-based techniques and look at distinguishing individual features, such as eyes, nose, mouth and head outline, and developing a face model based on position and size of these characteristics. Photometric techniques are view-based recognition methods. They distill an image into values and compare these values with templates. Many face recognition algorithms have been developed in the last decades. The most popular techniques include Eigenfaces (Turk & Pentland, 1991, Barbu, 2007), Fisherfaces (Yin et. al, 2005), Linear Discriminant Analysis (LDA), Elastic Bunch Graph Matching (EBGM), Hidden Markov Models (HMM) (Samaria et. al, 1994) and the neuronal model Dynamic Link Matching (DLM) (Wiskott et. al, 1996).

In this chapter we present two facial recognition approaches. The first one is an Eigenface-based recognition technique, based on the influential work of Turk and Pentland (Turk & Pentland, 1991). Proposed in 1991 by M. Turk and A. Pentland, the Eigenface approach was the first genuinely successful system for automatic recognition of human faces, representing a breakaway from contemporary research trend on facial recognition. The second technique, described in the third section, uses Gabor filtering in the feature extraction stage (Barbu, 2010). It applies a set of 2D Gabor filters, at various frequencies, orientations and standard deviations, on the facial images. A supervised classifier is used in the classification stage.

2. Eigenimage-based facial recognition approach

We have proposed an eigenimage-based face recognition technique based on the well-known approach of Turk and Pentland (Turk & Pentland, 1991). Their recognition method considers a large set of facial images that works as a training set.
Thus, each of these images is represented as a vector Γ_i, $i = 1,...,M$, then one computes the average vector Ψ. The covariance matrix is computed next as $C = A \cdot A^T$, where $A = [\Phi_1,...,\Phi_M]$ and $\Phi_i = \Gamma_i - \psi$. The matrix C is a very large one and its eigenvectors and eigenvalues are obtained from those of $A^T \cdot A$. Thus, $A \cdot A^T$ and $A^T \cdot A$ have the same eigenvalues and their eigenvectors are related as follows: $u_i = Av_i$. One keeps only $M^{'}$ eigenvectors, corresponding to the largest eigenvalues. Each of these eigenvectors represents an eigenimage (eigenface). Each face image is projected onto each of these eigenfaces, its feature vector, containing $M^{'}$ coefficients, being obtained. Any new input face image is identified by computing the Euclidean distance between its feature vector and each feature training vector. Next, some verification procedures may be necessary to determine if the input image represents a face at all or if it represents a registered person.
We have developed a derived version of this Eigenface approach. Thus, we propose a continuous mathematical model for face feature extraction in the first subsection (Barbu, 2007). Then we discretize this model, the discretized version being presented in the second subsection (Barbu, 2007). A supervised face classification system, that produces the identification results, is described in the third subsection. Facial verification is the last recognition process. In the fourth section we present a threshold-based verification approach. Some experiments performed by using our facial recognition method are presented in the fifth section.

2.1 A continuous model for face feature extraction

We develop a continuous differential model for facial feature extraction. Thus, our approach replaces the 2D face image Ω by a differentiable function $u = u(x, y)$ and the covariance matrix by a linear symmetric operator on the space $L^2(\Omega)$ involving the image vector u and its gradient ∇u. We determine a finite number of eigenfunctions, the identification process being developed on this finite dimensional space (Barbu, 2007).
Therefore, the continuous facial image become $u : \Omega \to R$ and we denote by $L^2(\Omega)$ the space of all L^2 - integrable functions u with the norm $|u|_2 = \left(\int_\Omega u^2(x, y)dxdy\right)^{1/2}$, and by $H^1(\Omega)$ the Sobolev space of all functions $u \in L^2(\Omega)$ with the distributional derivatives $D_x u = \frac{\partial u}{\partial x}$ and $D_y u = \frac{\partial u}{\partial y}$ respectively (Barbu V., 1998). Also, we denote by

$\nabla u(x,y) = (D_x u(x,y), D_y u(x,y))$ the gradient of $u = u(x,y)$. The Euclidean L^2 - norm $\int_\Omega u^2(x,y)dxdy$ is replaced in this new approach by the $H^1(\Omega)$ - energetic norm $\int_\Omega |u|^2 + |\nabla u|^2 dxdy$, that is sharper and more appropriate to describe the feature variations of the human faces. This technique is inspired by the diffusion models for image denoising, representing our main contribution (Barbu, 2007). Thus, the norm of $H^1(\Omega)$ is computed as:

$$|u|_{H^1(\Omega)} = \int_\Omega (u^2(x,y) + (D_x u(x,y))^2 + (D_y u(x,y))^2)dxdy = \int_\Omega (u^2(x,y) + |\nabla u(x,y)|^2)dxdy \quad (1)$$

We are given an initial set of facial images, $\{u_1,...,u_M\} \subset (H^1(\Omega))^M$, representing the training set. Therefore, its average value is computed as:

$$\mu(x,y) = \frac{1}{M}\sum_{i=1}^{M} u_i(x,y), \ \ x,y \in \Omega \quad (2)$$

Next, we compute the differences:

$$\Phi_i(x,y) = u_i(x,y) - \mu(x,y), \ \ i = 1,...,M \quad (3)$$

Then, we set

$$W_i = \nabla u_i = \{D_x u_i, D_y u_i\}, \ \ i = 1,...,M \quad (4)$$

and consider the covariance operator $Q \in L((L^2(\Omega))^3, (L^2(\Omega))^3)$ associated with the vectorial process $\{\Phi_i, W_i\}_{i=1}^M$. If $h = \{h_1, h_2, h_3\} \in L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)$, then we have:

$$\begin{aligned}(Qh)(x,y) = \Bigg\{ &\sum_{i=1}^{M} \Phi_i(x,y)\int_\Omega \Phi_i(\xi)h_1(\xi)d\xi + \sum_{i=1}^{M} W_i^1(x,y)\int_\Omega W_i^1(\xi)h_2(\xi)d\xi + \\ &+ \sum_{i=1}^{M} W_i^2(x,y)\int_\Omega W_i^2(\xi)h_3(\xi)d\xi, \Bigg\}, \ \forall h_k \in L^2(\Omega), k = 1,2,3\end{aligned} \quad (5)$$

where $W_i = \{W_i^1, W_i^2\}, W_i^1(x,y) = D_x U_i(x,y), W_i^2(x,y) = D_y U_i(x,y), \ i = 1,...,M$.

Equivalently we may view Q as covariance operator of the process $\Phi = \{\Phi_i\}_{i=1}^M$ in the space $H^1(\Omega)$ endowed with norm $|\bullet|_{H^1}$. Indeed, for $h_1 = z, h_2 = D_x z, h_3 = D_y z$, the equation (5) becomes:

$$(Qh)(x,y) = \{\Phi \cdot (\Phi, z)_{L^2(\Omega)}, \nabla\Phi \cdot (\nabla\Phi, \nabla z)_{L^2(\Omega)}\} = \{\Phi, \nabla\Phi\} < \Phi, z >_{H^1(\Omega)}, \ \forall z \in H^1(\Omega) \quad (6)$$

This means Q is just the covariance operator of the $H^1(\Omega)$ - variable Φ. We may consider Φ to be a random variable. The operator Q is self-adjoint in $(L^2(\Omega))^3$ and has an orthonormal complet system of eigenfunctions $\{\varphi_j\}$, i.e., $Q\varphi_j = \lambda_j \varphi_j$, $\lambda_j > 0$ (Barbu V., 1998).

Moreover, $\varphi_j \in (H^1(\Omega))^3$, $\forall j$. Let us associate with Q the $[3M \times 3M]$ matrix $\tilde{Q} = A^T A$, where $A : R^{3M} \to (L^2(\Omega))^3$ is given by $AY = \left[\sum_{i=1}^{M} \Phi_i y_1^i, \sum_{i=1}^{M} W_i^1 y_2^i, \sum_{i=1}^{M} W_i^1 y_3^i \right]$, $Y = \{y_1^i, y_2^i, y_3^i\}_{i=1}^M$, and $A^T : (L^2(\Omega))^3 \to R^{3M}$, that represents the adjoint operator, is given by

$$A^T h = \left[\int_\Omega \Phi_1(\xi) h_1 d\xi, ..., \int_\Omega \Phi_M h_1 d\xi, \int_\Omega W_1^1(\xi) h_2 d\xi, ..., \int_\Omega W_M^1(\xi) h_2 d\xi, \int_\Omega W_1^2(\xi) h_3 d\xi, ..., \int_\Omega W_M^2(\xi) h_3 d\xi \right]$$

where $h = (h_1, h_2, h_3) \in (L^2(\Omega))^3$. One results:

$$A^T A = \left\| \begin{matrix} \int_\Omega \Phi_i \Phi_j d\xi & 0 & 0 \\ 0 & \int_\Omega \Phi_i \Phi_j d\xi & 0 \\ 0 & 0 & \int_\Omega \Phi_i \Phi_j d\xi \end{matrix} \right\|_{i,j=1}^{M} \tag{7}$$

We consider $\{\lambda_j\}_{j=1}^{3M}$ and $\{\psi_j\}_{j=1}^{3M} \subset R^{3M}$ a linear independent system of eigenvectors for $\tilde{Q}$, therefore $\tilde{Q}\psi_j = \lambda_j \psi_j$, $j = 1, ..., 3M$. One can see that the sequence $\{\varphi_j\}_{j=1}^{3M}$, defined by $\varphi_j = \psi_j \Phi_j$, $for\ 1 \le j \le M$, $\varphi_j = \psi_j W_j^1$, $for\ M+1 \le j \le 2M$, and $\varphi_j = \psi_j W_j^2$, for $2M+1 \le j \le 3M$, represent the eigenfunctions of operator Q, i.e., $Q\varphi_j = \lambda_j \varphi_j$, $j = 1, ..., 3M$. The eigenfunctions of the covariance operator Q maximizes the variance of projected samples:

$$\varphi_j = arg\{max < Qh, h >_{(L^2(\Omega))^3} ; |h|_{(L^2(\Omega))^3} = 1\} \tag{8}$$

In this way the eigenfunctions $\{\varphi_j\}_{j=1}^{3M}$ capture the essential features of images from the initial training set. From $\{\varphi_j\}$ we keep a smaller number of eigenfaces $\{\varphi_j\}_{j=1}^{3M'}$, with $M' < M$ corresponding to largest eigenvalues λ_j and consider the space

$$X = lin\{\varphi_j\}_{j=1}^{3M'} \subset (L^2(\Omega))^3 \tag{9}$$

Assuming that systems $\{\varphi_j\}$ is normalized (orthonormal in $(L^2(\Omega))^3$) one can project any initial image $\{\Phi_i, W_i^1, W_i^2\} \in (L^2(\Omega))^3$ on X by formula

$$\Psi_i(x,y) = \sum_{j=1}^{3M'} \varphi_j(x,y) < \varphi_j, T_i >_{(L^2(\Omega))^3},\ i = 1, ..., 3M \tag{10}$$

where $T_i = \{\Phi_i, W_i^1, W_i^2\}$, $i = 1, ..., 3M$ and $< \cdot, \cdot >_{(L^2(\Omega))^3}$ is the scalar product in $L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)$. We denote the weights $w_i^j =< \varphi_j, T_i >_{(L^2(\Omega))^3}$, $i = 1, ..., 3M$, and the resulted feature vector will be

$$V(u_i) = (w_i^1,...,w_i^{3M'}),\ \ i = 1,...,3M \tag{11}$$

2.2 Discretization of the feature extraction approach

We discretize the continuous model described in the last section. Let us assume that $\Omega = [0,L_1] \times [0,L_2]$ and let us set $x_i = i\varepsilon,\ i = 1,...N_2$ and $y_j = j\varepsilon,\ j = 1,...N_1$, $\varepsilon > 0$ (Barbu, 2007).

Therefore, we obtain M matrices of size $N_1 \times N_2$, which represent the discrete images. We denote their corresponding $N_1 \cdot N_2 \times 1$ image vectors as $I_1,...,I_M$. Now, we have $\tilde{Q} = A^T \cdot A$, where

$$A = \begin{Vmatrix} \Phi_1,...,\Phi_M & 0 & 0 \\ 0 & W_1^1,...,W_M^1 & 0 \\ 0 & 0 & W_1^2,...,W_M^2 \end{Vmatrix} \tag{12}$$

and

$$\begin{cases} \Phi_k = \left\| \Phi_k(x_i, y_j) \right\|_{i,j=1}^{N_2,N_1} \\ W_k^1 = \left\| \Phi_k(x_{i+1}, y_j) - \Phi_k(x_i, y_j) \right\|_{i,j=1}^{N_2,N_1} \\ W_k^2 = \left\| \Phi_k(x_i, y_{j+1}) - \Phi_k(x_i, y_j) \right\|_{i,j=1}^{N_2,N_1} \end{cases} \tag{13}$$

The obtained matrix has a $3M \times 3M$ dimension. Thus, we determine the eigenvectors ψ_i of the matrix $\tilde{Q}$. Then, the eigenvectors of the discretized covariance operator Q are computed as following:

$$\tilde{\varphi}_i = A \cdot \psi_i,\ \ i = 1,...,M \tag{14}$$

We keep only $M' < M$ eigenimages corresponding to the largest eigenvalues and consider the space $X = linspan\{\tilde{\varphi}_i\}_{i=1}^{3M'}$. Then, the projection of $[\Phi_i, W_i^1, W_i^2]$ on X is given by the discrete version of Ψ_i, that is computed as:

$$P_X([\Phi_i, W_i^1, W_i^2]) = \sum_{j=1}^{3M'} w_i^j \cdot \varphi_j,\ \ i = 1,...,3M \tag{15}$$

where $w_i^j = \tilde{\varphi}_j^T \cdot [\Phi_i, W_i^1, W_i^2]^T$. So, for each facial image I_i a corresponding training feature vector is extracted as the sequence of all these weights:

$$V(I_i) = [w_i^1,...,w_i^{3M'}]^T,\ \ i = 1,...,3M \tag{16}$$

Therefore, the feature training set of the face recognition system is obtained as $\{V(I_1),...,V(I_M)\}$, each feature vector being given by (16). The Euclidean metric could be used to measure the distance between these vectors.

2.3 Facial feature vector classification and verification

The next step of face recognition is the classification task (Duda et. al, 2000). We provide a supervised classification technique for facial authentication. Thus, we consider an unknown input image I to be recognized using the face training set $\{I_1,...,I_M\}$. The feature training set of our classifier is $\{V(I_1),...,V(I_M)\}$ and its metric is the Euclidean distance for vectors. We normalize the input image, first. So, we get $\Phi = I - \Psi$, where $\Psi = \frac{1}{M}\sum_{i=1}^{M} I_i$. The vectors W^1 and W^2 are computed from Φ using formula (13). Then it is projected on the eigenspace, using the formula $P(\Phi) = \sum_{i=1}^{M'} w^i \varphi_i$, where $w^i = \tilde{\varphi}_i^T \cdot [\Phi, W^1, W^2]^T$. Obviously, its feature vector is computed as $V(I) = [w^1,...,w^{M'}]^T$ (Barbu, 2007).

A threshold-based facial test could be performed to determine if the given image represents a real face or not. Thus, if $\|P(\Phi) - \Phi\| \le T$, where T is a properly chosen threshold value, then I represents a face, otherwise it is a non-facial image. We consider K registered (authorized) persons whose faces are represented in the training set $\{I_1,...,I_M\}$. We redenote this set as $\{I_1^1,...,I_1^{n(1)},...,I_i^1,...,I_i^{n(i)},...,I_K^1,...,I_K^{n(K)}\}$, where $K < M$ and $\{I_i^1,...,I_i^{n(i)}\}$ represents the training subset provided by the i^{th} authorized person, $n(i)$ being the number of its registered faces.

A minimum average distance classifier, representing an extension of the classical variant of minimum distance classifier (Duda et. al, 2000), is used for feature vector classification. So, for each registered user, one calculates the average distance between its feature training subset and the feature vector of the input face. The input image is associated to the user corresponding to the minimum distance value. That user is the k^{th} registered person, where

$$k = \arg\min_i \frac{\sum_{j=1}^{n(i)} d(V(I), V(I_i^j))}{n(i)} \tag{17}$$

where d is the Euclidean metric. Thus, each input face I is identified as a registered person by this formula. However, the face identification procedure has to be completed by a facial verification step, to determine if that face really represents the associated person. We propose a threshold-based verification approach.

First, we provide a novel threshold detection solution, considering the overall maximum distance between any two feature vectors belonging to the same training feature sub-set as a threshold value (Barbu, 2007). Therefore, the input face I may represent the k^{th} registered user if the following condition is fulfilled:

$$\min_i \frac{\sum_{j=1}^{n(i)} d(V(I), V(I_i^j))}{n(i)} \le T \tag{18}$$

where k is computed by (17) and the threshold is given by formula:

$$T = \max_{i} \max_{j \neq t} d(V(I_i^j), V(I_i^t)) \tag{19}$$

If the condition (18) is not satisfied, then the input face is rejected by our recognition system and labeled as *unregistered*.

2.4 Experiments

The proposed Eigenface-based recognition system has been tested on numerous face datasets. We have performed many experiments and achieved satisfactory facial recognition results. Our technique produces a high recognition rate, of approximately 90%. We have used "Yale Face Database B", containing thousands of 192×168 face images corresponding to many persons, in our research (Georghiades et. al, 2001). In the next figures, there is represented such a face recognition example. In Fig. 1 one can see a set of 10 input faces to be recognized.

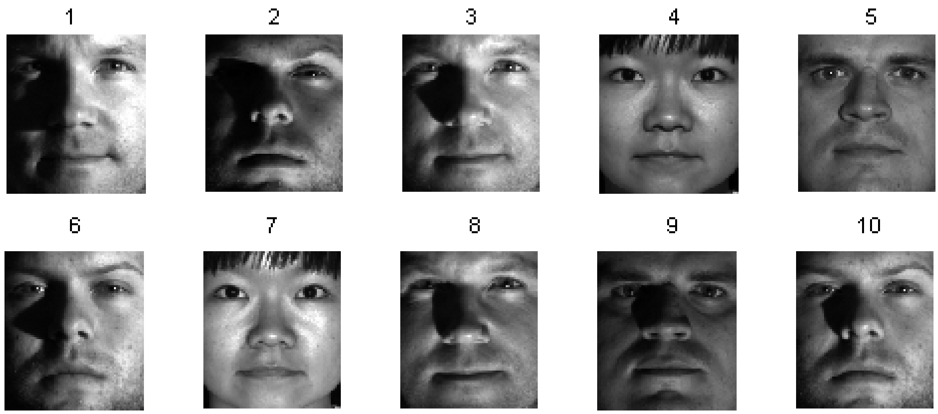


Fig. 1. Input face images

This example uses a training set containing 30 facial images belonging to 3 persons. The face set is illustrated in Fig. 2, where each registered individual has 10 photos positioned on two consecutive rows.
Therefore, one computes 90 eigenfaces for this training set but only the most important 27 ($M' = 9$) of them are necessary. The most significant eigenfaces are represented in Fig. 3. Thus, the feature training set, containing face feature vectors, is obtained on the basis of these eigenimages.
The mean distances between these feature vectors are computed and the identification process provided by (17) is applied. Faces 2, 6 and 10 are identified as belonging to the first person, faces 4 and 7 are identified to the second person. Also, the faces 5 and 9 are associated to the first registered individual but their distance values, 5.795 and 5.101, are greater than the threshold value provided by (19), computed as $T = 2.568$, so the verification procedure labels them as unregistered.

Fig. 2. Facial training set

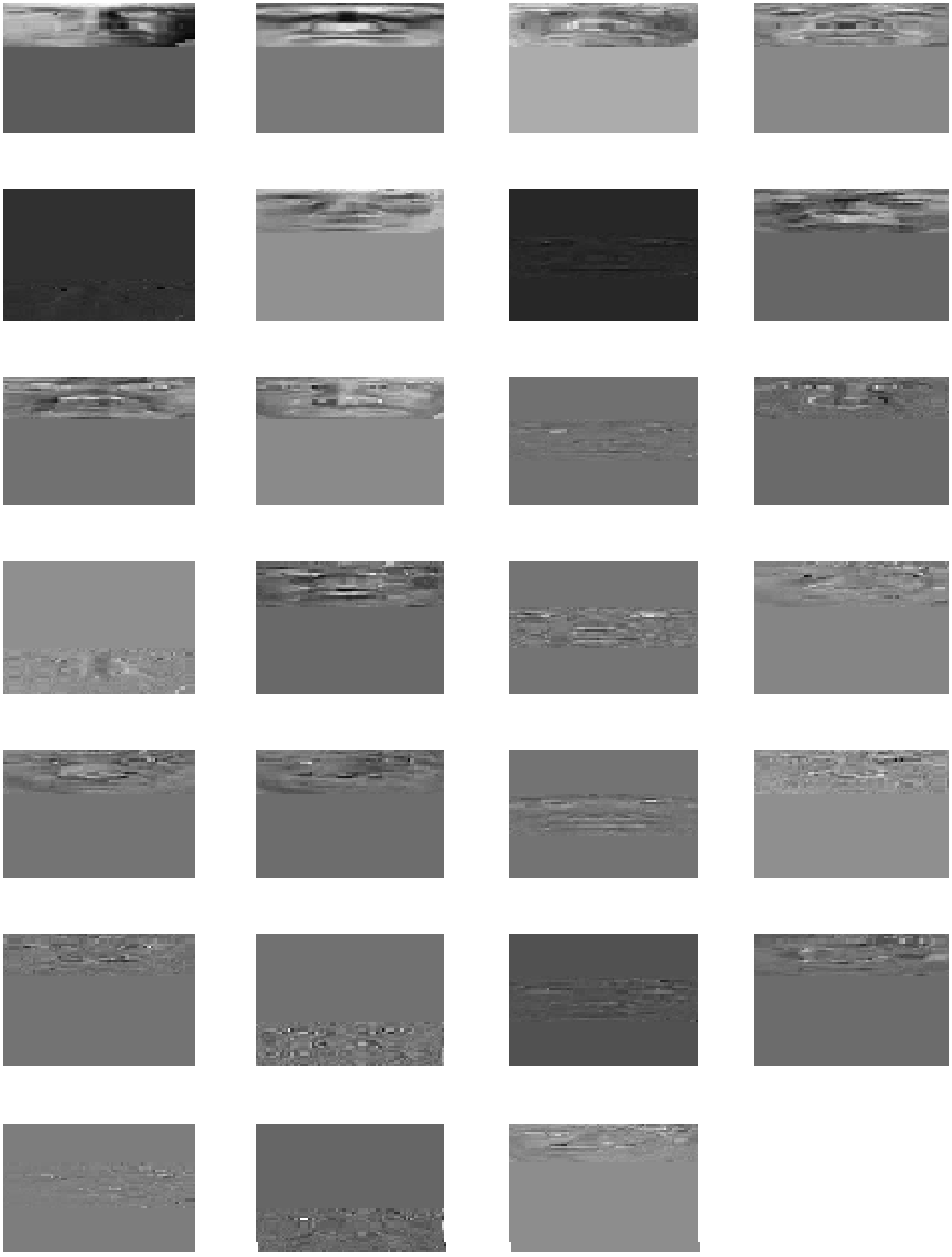


Fig. 3. The main eigenfaces

3. Face recognition technique using two-dimensional Gabor filtering

The second face authentication approach is based on two-dimensional Gabor filtering. As one knows, a great amount of research papers have been published in literature for Gabor filter-based image processing (Movellan, 2008). Besides face recognition, Gabor filters are successfully used in many other image processing and analysis domains, such as: image smoothing, image coding, texture analysis, shape analysis, edge detection, fingerprint and iris recognition.

We use the Gabor filters in the feature extraction process, that is described in the next subsection. Then, we use a feature vector classification approach that is similar to the supervised method proposed in the previous section. A threshold-based face verification is also described in the second subsection. Some facial recognition experiments are presented in the last subsection.

3.1 A Gabor filter based facial feature extraction

We intend to obtain some feature vectors which provide proper characterizations of the visual content of face images. For this reason we use the two-dimensional Gabor filtering as a feature extraction tool (Barbu, 2010).

The Gabor filter represents a band-pass linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. Thus, a bidimensional Gabor filter constitutes a complex sinusoidal plane of particular frequency and orientation modulated by a Gaussian envelope. It achieves optimal resolutions in both spatial and frequency domains (Movellan, 2008, Barbu, 2010). Our approach designs 2D odd-symmetric Gabor filters for face image recognition, having the following form:

$$G_{\theta_k,f_i,\sigma_x,\sigma_y}(x,y) = exp\left(-\left[\frac{x_{\theta_k}^2}{\sigma_x^2} + \frac{y_{\theta_k}^2}{\sigma_y^2}\right]\right) \cdot cos\left(2\pi f_i x_{\theta_k} + \varphi\right) \tag{20}$$

where $x_{\theta_k} = x\cos\theta_k + y\sin\theta_k$, $y_{\theta_k} = y\cos\theta_k - x\sin\theta_k$, f_i provides the central frequency of the sinusoidal plane wave at an angle θ_k with the x – axis, σ_x and σ_y represent the standard deviations of the Gaussian envelope along the axes x and y.

Then, we set the phase $\varphi = \pi / 2$ and compute each orientation as $\theta_k = \frac{k\pi}{n}$, where $k = \{1,...,n\}$. The 2D filters $G_{\theta_k,f,\sigma_x,\sigma_y}$ computed by (20) represent a group of wavelets which optimally captures both local orientation and frequency information from a digital image (Barbu, 2010).

Each facial image has to be filtered by applying $G_{\theta_k,f,\sigma_x,\sigma_y}$ at various orientations, frequencies and standard deviations. A proper design of Gabor filters for face authentication requires an appropriated selection of those parameters. Therefore, we consider some proper variance values, a set of radial frequencies and a sequence of orientations, respectively.

The settings for the filter parameters are: $\sigma_x = 2$, $\sigma_y = 1$, $f_i \in \{0.75,\ 1.5\}$ and n = 5, which means $\theta_k \in \left\{\frac{\pi}{5}, \frac{2\pi}{5}, \frac{3\pi}{5}, \frac{4\pi}{5}, \pi\right\}$. One results the filter bank $\left\{G_{\theta_k,f_i,2,1}\right\}_{f_i \in \{0.75,1.5\}, k \in [1,5]}$, that is composed of 10 channels. The current face image is convolved with each 2D Gabor filter from this set. The resulted Gabor responses are then concatenated into a three-dimensional feature vector. So, if I represents a $[X \times Y]$ face image, then the feature extraction is modeled as:

$$V(I)[x,y,z] = V_{\theta(z),f(z),\sigma_x,\sigma_y}(I)[x,y] \tag{21}$$

where $x \in [1,X]$, $y \in [1,Y]$ and

$$\theta(z) = \begin{cases} \theta_z, & z \in [1,n] \\ \theta_{z-n}, & z \in [n+1,2n] \end{cases}, \quad f(z) = \begin{cases} f_1, & z \in [1,n] \\ f_2, & z \in [n+1,2n] \end{cases} \tag{22}$$

and

$$V_{\theta(z),f(z),\sigma_x,\sigma_y}(I)[x,y] = I(x,y) \otimes G_{\theta(z),f(z),\sigma_x,\sigma_y}(x,y) \tag{23}$$

A fast 2D convolution is performed using Fast Fourier Transform, so, the face feature vector is computed as $V_{\theta(z),f(z),\sigma_x,\sigma_y}(I) = FFT^{-1}[FFT(I) \cdot FFT(G_{\theta(z),f(z),\sigma_x,\sigma_y})]$. For each face I one obtains a 3D face feature vector $V(I)$, that has a $[X \times Y \times 2n]$ dimension.
This feature vector proves to be a satisfactory content descriptor of the input face. In Fig. 4 one depicts a facial image and its 10 Gabor representations, representing the components of its feature vector.

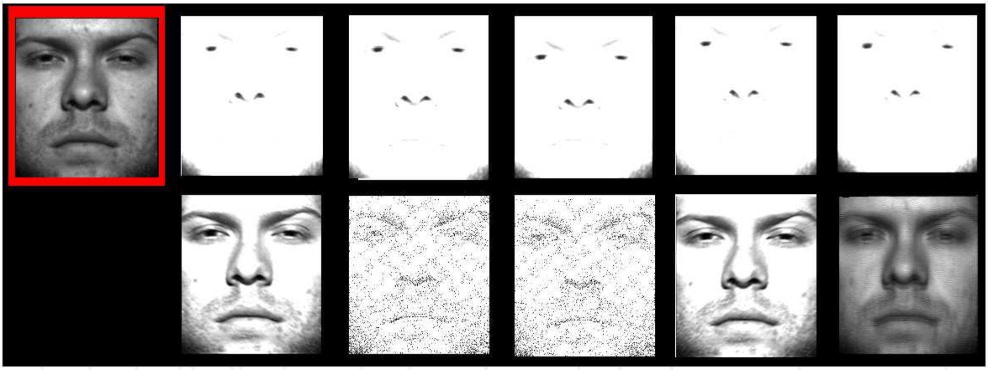


Fig. 4. Feature vector components of a face image

The distance between these feature vectors must be computed using a proper metric. Since the size of each vector depends on the dimension of the featured image, a resizing procedure has to be performed on the face images, first. Then, we compute the distance between these 3D feature vectors using a *squared Euclidean metric*, which is characterized by formula:

$$d(V(I),V(J)) = \sum_{x=1}^{X}\sum_{y=1}^{Y}\sum_{z=1}^{2n} \left| V(I)[x,y,z)] - V(J)[x,y,z)] \right|^2 \tag{24}$$

where I and J are two face images, resized to the same $[X \times Y]$ dimension.

3.2 Face feature vector classification and verification

Feature vector classification represents the second step of the face identification process. We provide a similar supervised classification approach for these Gabor filter-based 3D feature vectors. Some well-known supervised classifiers (Duda et. al, 2000), such as minimum distance classifier or the *K-Nearest Neighbour* (*K-NN*) classifier, could be also used in this case.

The training set of this classifier is created first. Therefore, one considers N authorized system users. Each of them provides a set of faces of its own, named templates, which are included in the training set. Thus, the training face set can be modeled as $\left\{\left\{F_j^i\right\}_{j=1,...,n(i)}\right\}_{i=1,...,N}$, where F_j^i represents the j^{th} template face of the i^{th} user and $n(i)$ is the number of training faces of the i^{th} user. The classification algorithm produces N face classes, each one corresponding to a registered person. Next, one computes the training vector set as $\left\{\left\{V(F_j^i)\right\}_{j=1,...,n(i)}\right\}_{i=1,...,N}$.

If the input faces to be recognized are $\{I_1,...,I_K\}$, the classification procedure inserts each of them in the class of the *closest* registered person, representing the user corresponding to the minimum *average distance*. The minimum average distance classification process is described by a formula that is similarly to (17):

$$Class(j) = arg \min_{i\in[1,N]} \frac{\sum_{t=1}^{n(i)} d(V(I_j),V(F_t^i))}{n(i)}, \forall j \in [1,K] \tag{25}$$

where the obtained $Class(j) \in [1,N]$ represents the index of the face class where I_j is introduced. Let $C_1,...,C_N$ be the resulted classes, representing the facial identification result.

Next, a verification procedure is performed, to complete the face recognition task. We use a similar automatic threshold-based verification approach (Barbu, 2010). Therefore, one computes a proper threshold value as the overall maximum distance between any two training face feature vectors corresponding to the same registered user, that is:

$$T = \max_{i\le N}\left(\max_{j\neq k\in[1,n(i)]} d\left(V(F_j^i),V(F_k^i)\right)\right) \tag{26}$$

If the average distance corresponding to an image from a class is greater than threshold T, then that image has to be rejected from the face class. The verification process is represented formally as follows:

$$\forall i \in [1,N], \forall I \in C_i : \frac{\sum_{j=1}^{n(i)} d(V(I),V(F_j^i))}{n(i)} > T \Rightarrow C_i = C_i - \{I\} \tag{27}$$

The rejected images, that could represent non-facial images or faces of unregistered users, are included in a new class C_{N+1} , labeled as *Unauthorized*.

3.3 Experiments and method comparisons

We have performed many facial recognition experiments, using the proposed Gabor filter based approach. Our recognition system has been tested on various large face image datasets and good results have been obtained.
A high face recognition rate, of approximately 90%, has been reached by our recognition system in the experiments involving hundreds frontal images. We have obtained high values (almost 1) for the performance parameters, *Precision* and *Recall*, and for the combined measure F_1 . That means our approach produce a small number of false positives and false negatives (missed hits).
We have used the same database as in the previous case, *Yale Face Database B*, containing thousands of 192×168 faces at different illumination conditions, representing various persons, for our authentication tests (Georghiades et. al, 2001). The obtained results prove the effectiveness of the proposed human face authentication approach. We have obtained lower recognition rates for images representing rotated or non-frontal faces, and higher authentication rates for frontal images.
We have performed some comparisons between the two proposed facial recognition techniques. Also, we have compared them with other face authentication methods. Thus, we have compared the performance of our approaches with the performances of Eigenface-based systems, which are the most popular in the facial recognition area.
Therefore, we have tested the Eigenface algorithm of Turk & Pentland, our Eigenface technique and the Gabor filter based method on the same facial dataset. One have computed the statistical parameters *Precision* and *Recall*, using the number of correctly recognized faces and the number of correctly rejected faces, for these approaches, the values registered in Table 1 being obtained.

	Eigenface (T&P)	Eigenface (TB)	Gabor filter based
Precision	0.95	0.85	0.88
Recall	0.94	0.85	0.90

Table 1. Performance parameter comparison

As one can see in this table, the three face recognition techniques produce comparable performance results. The original Eigenface technique performs slightly better than our two methods.

4. Conclusions

We have proposed two automatic supervised facial recognition approaches in this chapter. As one can observe, the two face authentication techniques performs the same sequence of recognition related processes: feature extraction, feature vector classification and verification of the face identity.
While the two recognition methods differ substantially in the feature extraction stage, they use quite similar classification and verification techniques. In both cases, the feature vector classification process is supervised and based on a facial training set. We propose a

minimum average distance classifier that produces proper face identification. Also, both methods use a threshold-based verification technique. We have provided a threshold value detection approach for face verification.

The main contributions of this work are brought in the feature extraction stages of the proposed recognition techniques. The most important contribution is the continuous model for the Eigenface-based feature extraction. Then, the discretized version of this model represents another important originality element of this chapter.

The proposed Gabor filtering based feature extraction procedure represents another contribution to face recognition domain. We have performed a proper selection of the Gabor 2D filter parameters and obtained a powerful Gabor filter set which is successfully applied to facial images.

We have compared the two recognition methods, based on the results of our experiments, and found they have quite similar performances. Also, both of them are characterized by high face recognition rates. The facial authentication techniques described here work for cooperative human subjects only. Our future research in the face recognition domain will focus on developing some recognition approaches for non-cooperative individuals.

The recognition techniques of this type can be applied successfully in surveillance systems and law enforcement domains. A facial authentication approach that works for non-cooperative persons could be obtained by combining one of the recognition techniques proposed in this chapter with a face detection method (Barbu, 2011).

5. References

Zhao, W., Chellappa, R., Phillips, P. J. (2003). Face Recognition: A Literature Survey, *ACM Computing Surveys*, Volume 35, Number 4, pp. 399-458.

Turk, M. A. & Pentland, P. A. (1991). Face recognition using eigenfaces, *Proc. Of Computer Vision and Pattern Recognition*, pp. 586-591, IEEE.

Barbu, T. (2007). Eigenimage-based face recognition approach using gradient covariance, *Numerical Functional Analysis and Optimization*, Volume 28, Issue 5 & 6, pp. 591 – 601.

Yin, B., Bai, X., Shi, Q., Sun, Y. (2005). Enhanced Fisherface for Face Recognition, *Journal of Information and Computational Science*, No. 3, pp. 591-595.

Samaria, F., Young S. (1994). *HMM based architecture for face identification*, Image and Computer Vision, volume 12, pp. 537-583.

Wiskott, L., Malsburg, C. (1996). Face Recognition by Dynamic Link Matching, In J. Sirosh, R. Miikkulainen and Y. Choe editors, *Lateral Interactions in the Cortex: Structure and Function*. UTCS Neural Networks Research Group, Austin, TX, ISBN 0-9647060-0-8.

Barbu, T. (2010). Gabor filter-based face recognition technique, *Proceedings of the Romanian Academy*, Series A: Mathematics, Physics, Technical Sciences, Information Science, Volume 11, Number 3, pp. 277 – 283.

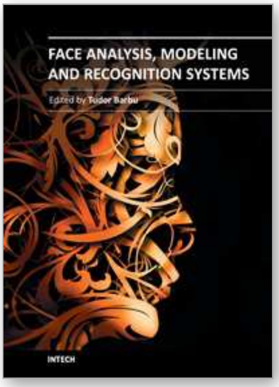
Barbu, V. (1998). *Partial Differential Equations and Boundary Value Problems*, Kluwer Academic Publishers, Dordrecht, Boston, London.

Duda, R. O., Hart, R. P., Stork, D. G. (2000), *Pattern Classification*. John Wiley & Sons.

Georghiades, A. S., Belhumeur, P. N., Kriegman, D. J. (2001). From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, Vol. 23, No. 6, pp. 643-660.

Movellan, J.R. (2008). *Tutorial on Gabor filters*, http://mplab.ucsd.edu/tutorials/gabor.pdf.

Barbu, T. (2011). An Automatic Face Detection System for RGB Images, *International Journal of Computers, Communications & Control*, Vol. 6, No.1, pp. 21-32



Face Analysis, Modeling and Recognition Systems

Edited by Dr. Tudor Barbu

ISBN 978-953-307-738-3
Hard cover, 212 pages
Publisher InTech
Published online 30, September, 2011
Published in print edition September, 2011

The purpose of this book, entitled Face Analysis, Modeling and Recognition Systems is to provide a concise and comprehensive coverage of artificial face recognition domain across four major areas of interest: biometrics, robotics, image databases and cognitive models. Our book aims to provide the reader with current state-of-the-art in these domains. The book is composed of 12 chapters which are grouped in four sections. The chapters in this book describe numerous novel face analysis techniques and approach many unsolved issues. The authors who contributed to this book work as professors and researchers at important institutions across the globe, and are recognized experts in the scientific fields approached here. The topics in this book cover a wide range of issues related to face analysis and here are offered many solutions to open issues. We anticipate that this book will be of special interest to researchers and academics interested in computer vision, biometrics, image processing, pattern recognition and medical diagnosis.

How to reference

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

Tudor Barbu (2011). Two Novel Face Recognition Approaches, Face Analysis, Modeling and Recognition Systems, Dr. Tudor Barbu (Ed.), ISBN: 978-953-307-738-3, InTech, Available from: http://www.intechopen.com/books/face-analysis-modeling-and-recognition-systems/two-novel-face-recognition-approaches

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