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 :

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

Where u_i and v_{ik} are the ith eigenspace and the kth value of the ith 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:

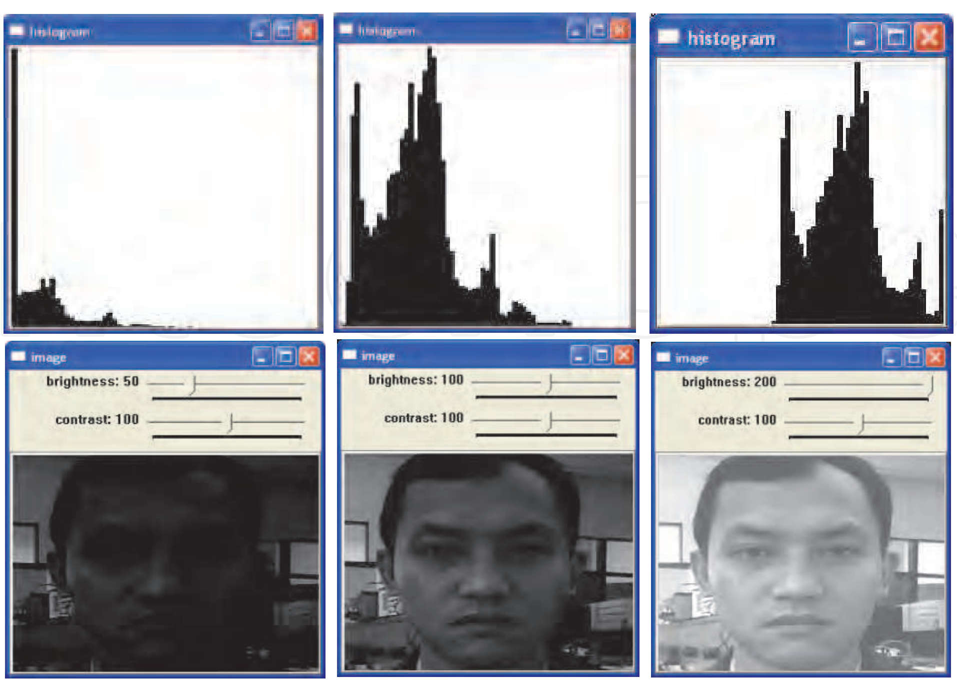


Fig. 2. Effect of varying the illumination for a face.

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:

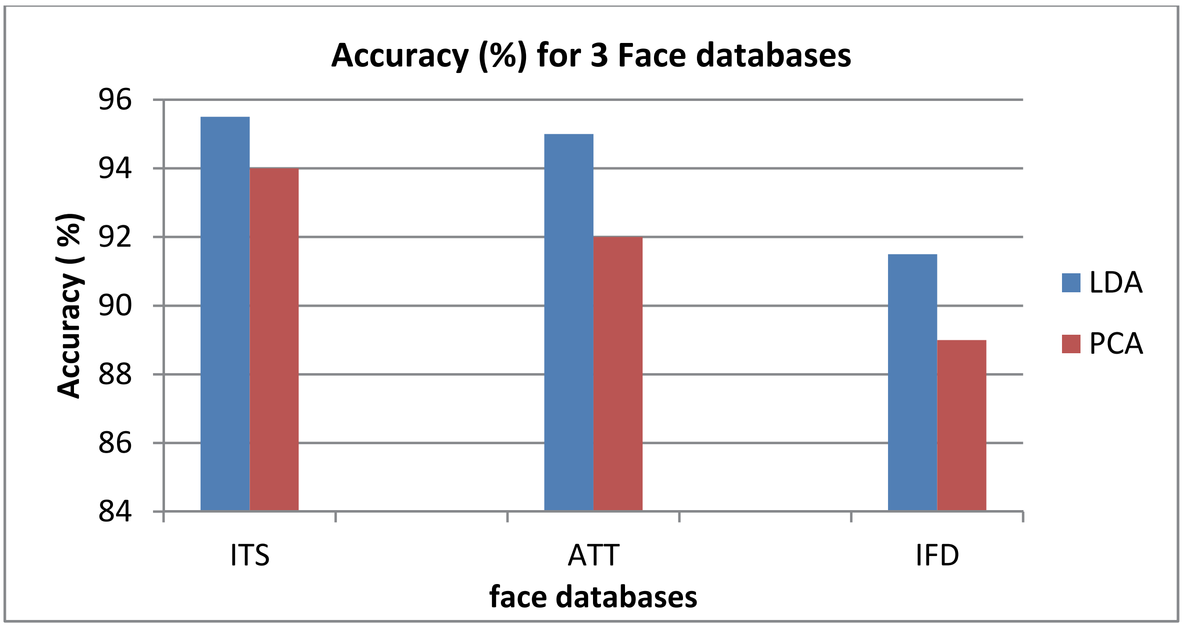


Fig. 4. Success rate comparison of face recognition between 3 faces databases, each using 10 sets face. It shown clearly that ITS database have highest success rate than ATT and Indian face database when the illumination of testing images is varied. The success rate using PCA in our proposed method and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

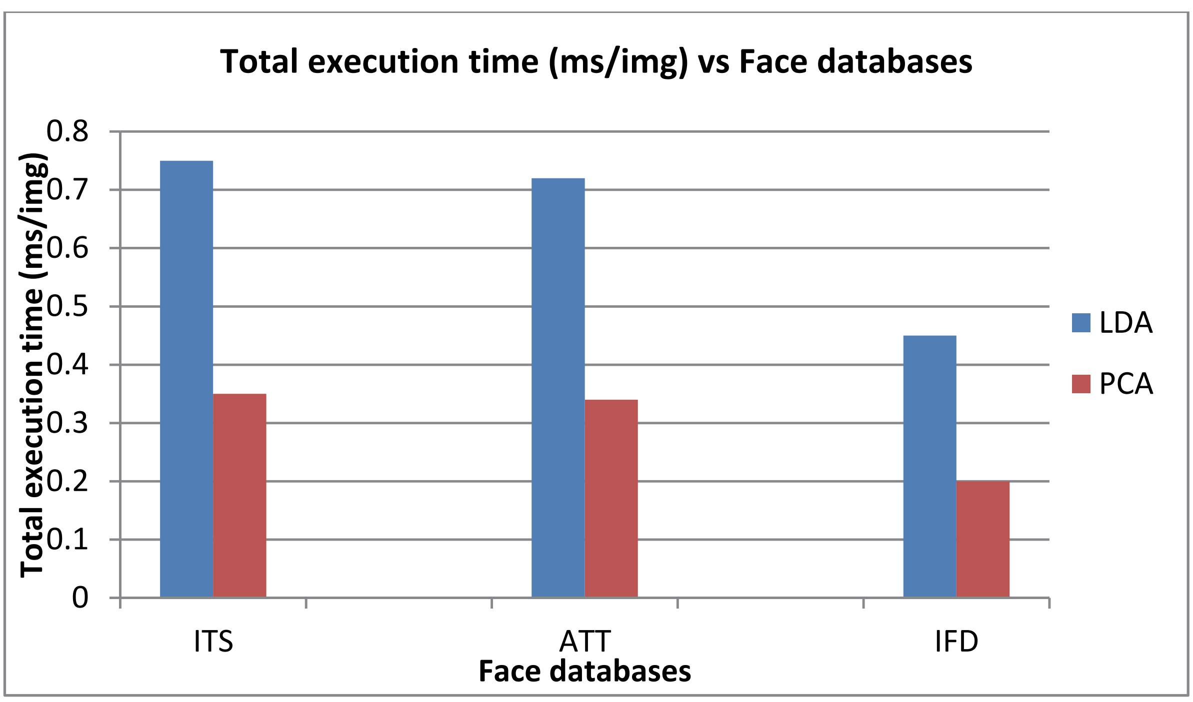


Fig. 5. For total execution time, we can see the Indian face database (IFD) is shortest because the size of each image is lowest then ITS and ATT.

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.

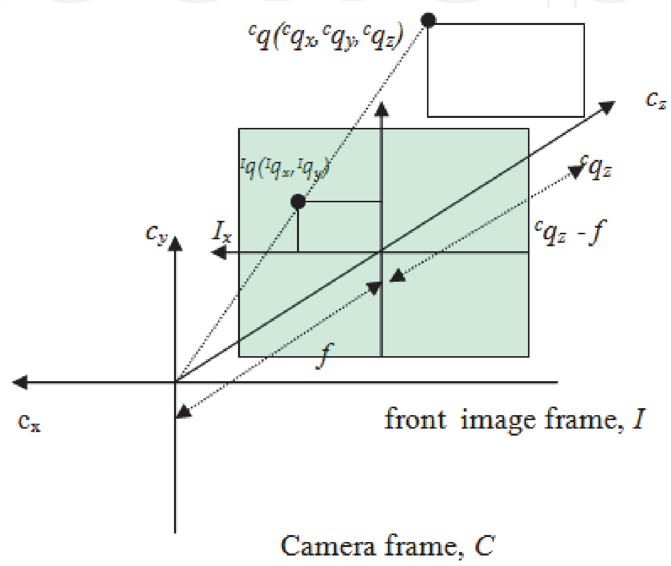


Fig. 6. Projection of point on front image frame.

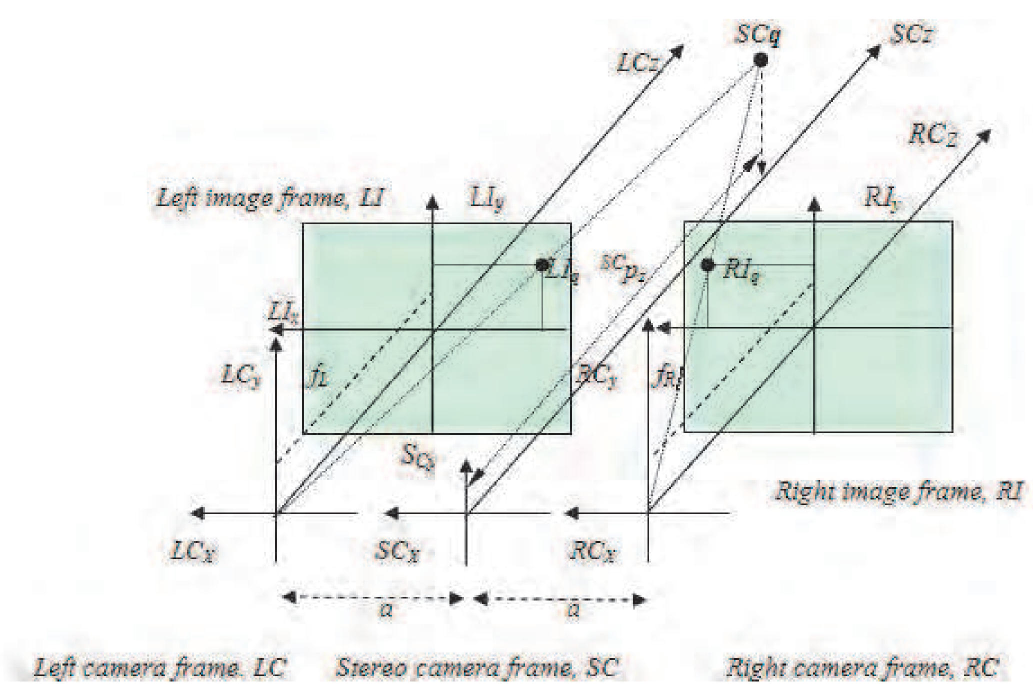


Fig. 7. Stereo Imaging model

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:

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

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 :

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

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

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

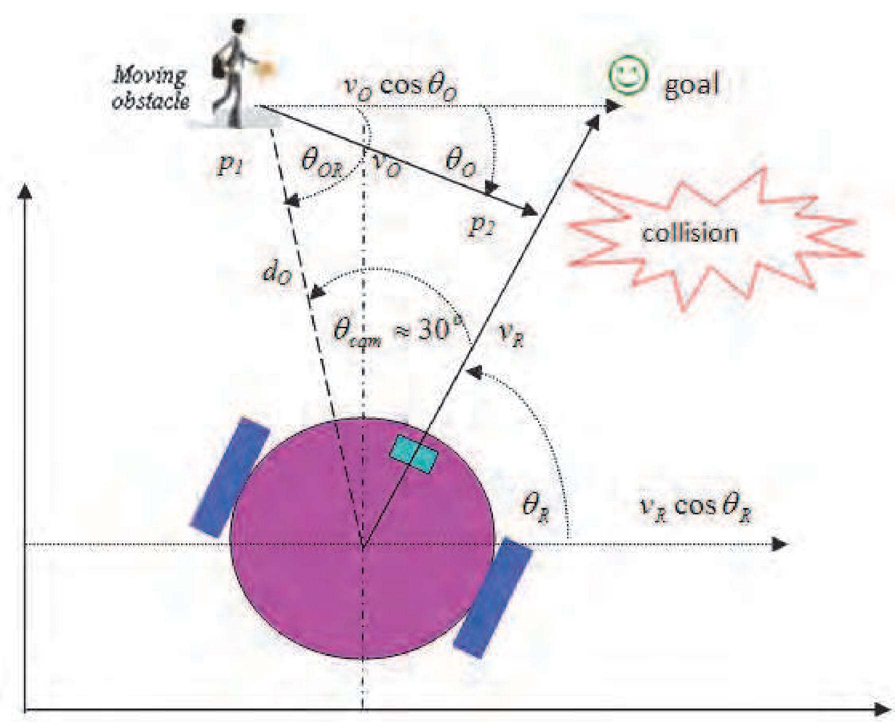


Fig. 9. Proposed Cartesian model of mobile robot with moving obstacle

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.

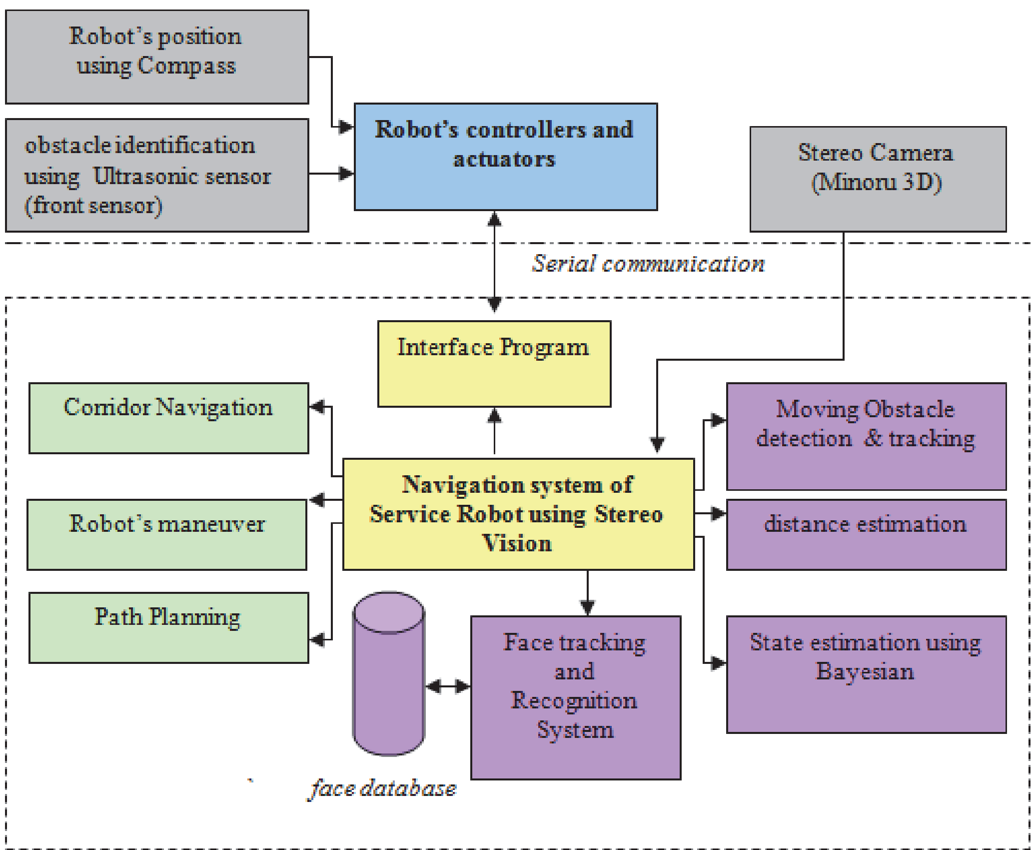


Fig. 10. General architecture of service robot called Srikandi III. Hardware and software parts are separated by the dashed line. All arrows indicate dataflow.

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

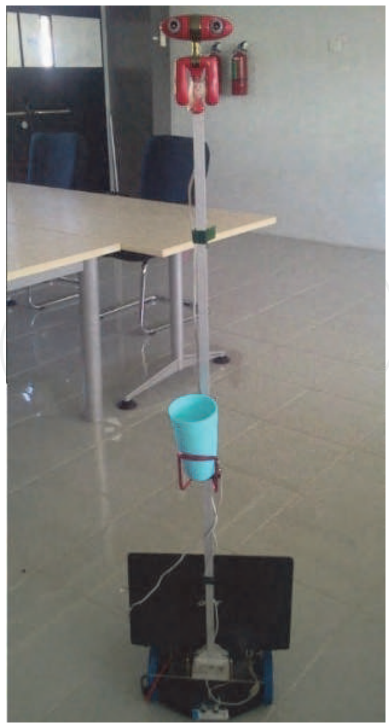


Fig. 11. Prototype of Service robot Srikandi III using stereo camera

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:

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

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.

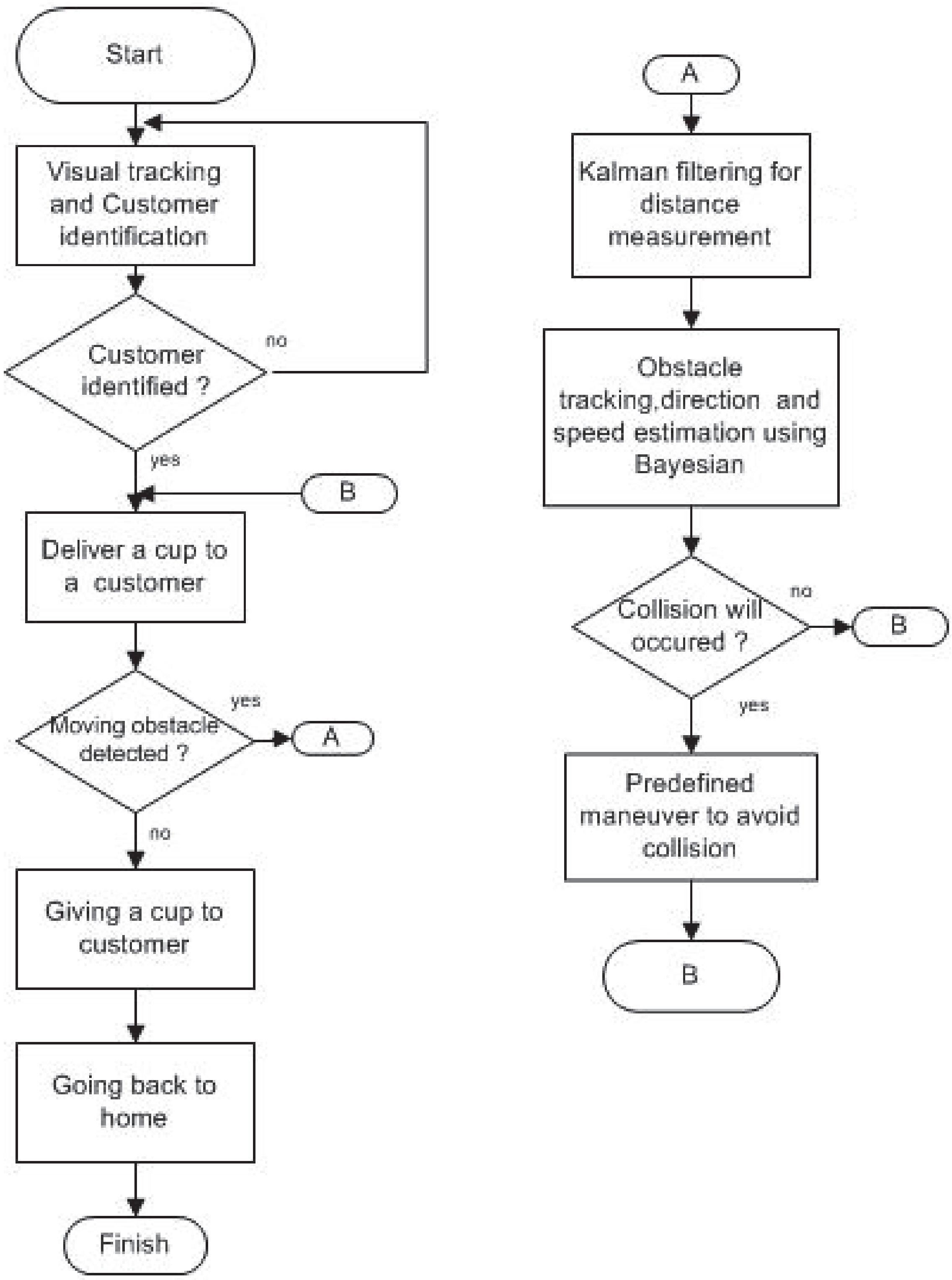


Fig. 12. Flow chart of Navigation System from start to goal position for service robot.

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

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

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

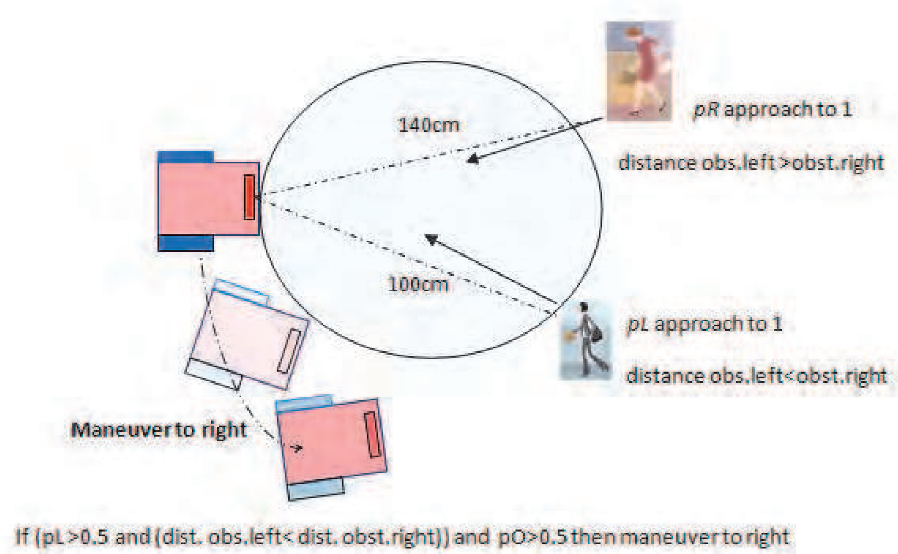
We define a loss function *L(a, θ)* which gives a measure of the loss incurred in taking action *a* when the state is *θ*. The robot should chooses an action a from the set **A** of possible actions based on the observation **z** of the current state of the path *θ*. This gives the posterior distribution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

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

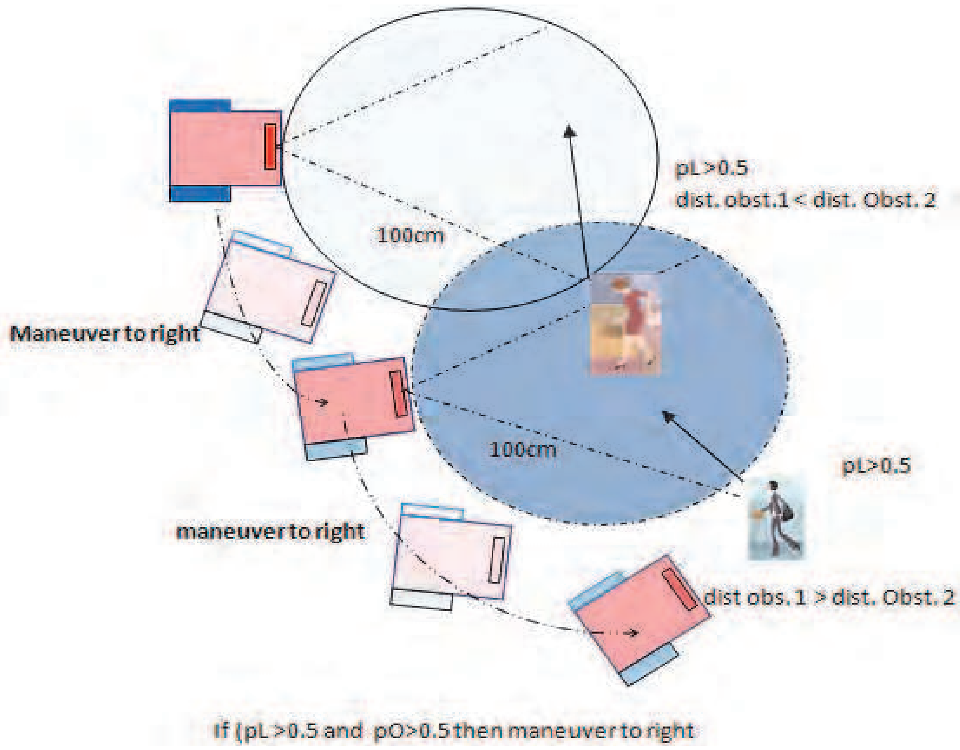
Then, based on the posterior distribution in (17), we can compute the posterior expected loss of an action (Hu, H et al., 1994):

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

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.

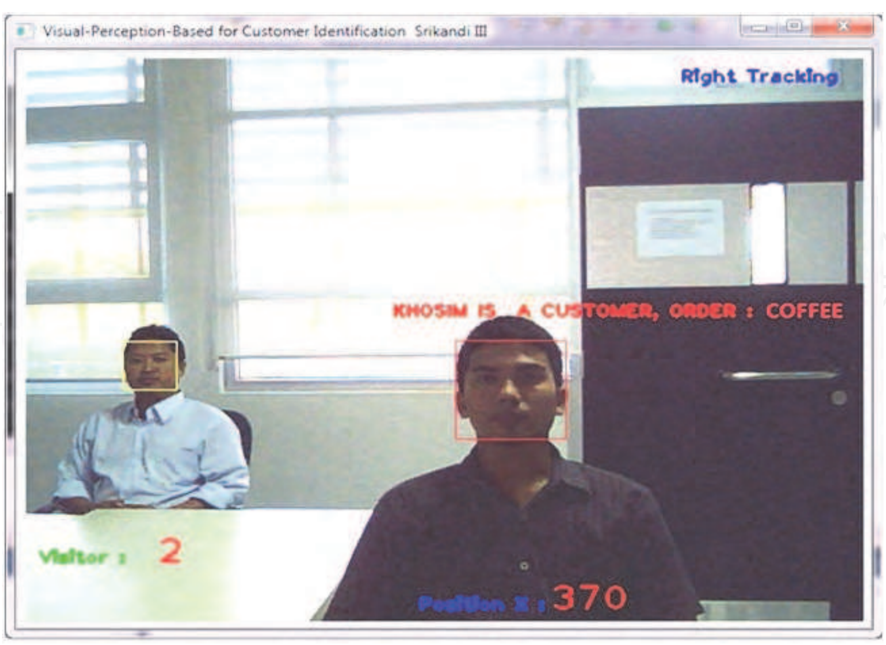


Fig. 14. Result of simulation using improved face recognition system and implemented to a service robot to identify a customer.

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.

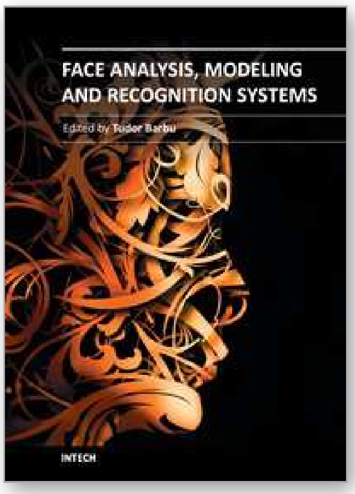
Yang M. (2002) Detecting faces images: A survey, IEEE Transactions on Pattern Analysis and Machine Inteligence, vol. 24 no. 1, pp.34-58

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.