Image Based Solution to Occlusion Problem for Multiple Robots Navigation

TAJ MOHAMMAD KHAN* AND MUHAMAD AHMED CHOUDHRY**

RECEIVED ON 10.06.2009 ACCEPTED ON 18.11.2009

ABSTRACT

In machine vision, occlusions problem is always a challenging issue in image based mapping and navigation tasks. This paper presents a multiple view vision based algorithm for the development of occlusion-free map of the indoor environment. The map is assumed to be utilized by the mobile robots within the workspace. It has wide range of applications, including mobile robot path planning and navigation, access control in restricted areas, and surveillance systems. We used wall mounted fixed camera system. After intensity adjustment and background subtraction of the synchronously captured images, the image registration was performed. We applied our algorithm on the registered images to resolve the occlusion problem. This technique works well even in the existence of total occlusion for a longer period.

Key Word: **Camera Modelling, Fusion of Data from Multiple Vision Sensors, Detection, Occlusion Detection, Mapping, Navigation.**

1. INTRODUCTION

The need of accurately mapping the workspace can not be tolerated for autonomous mobile robots, working in a factory or workshop environment. The quality and accuracy of the path and trajectory planning of mobile robots greatly depend on the quality and accuracy of the map provided. For mobile robot safe navigation and path planning the detection of dynamic obstacles is an important issue. Cameras are convenient sensors for such tasks [1]. The presence of multiple objects in workspace makes the occlusion problem sever. An independent vision system including camera network, a centralized computing facility for map building and wireless broadcast capabilities is the need for fast and accurate mapping of the dynamic indoor workspace of robots.

Occlusion free, cluttered environment mapping and objects localization is one of the key issues in robots visual navigation and surveillance systems. Typically, during occlusion, all or portion of each object is invisible. While tracking, objects are usually occluded by other objects and tracking system may fail. It severely affects the map quality in the presence of multiple dynamic robots/objects, especially when their shape, direction and speed are close. This is one of highly dynamic and intractable problem for multi-robot localization and safe navigation. To minimize ambiguities due to occlusion, better techniques need be developed to cope with the uncertainty in the map building process especially for mobile robot.

A great number of techniques have been proposed for the detection of occlusion problem. Wang and Liu [2] used blob model, motion model and color model in unified way to track people in an outdoor environment. Felman and Weinshall [3-4] used spatio-temporal intensities for extraction of motion boundaries; however, they assumed

*Ph.D. Scholar, and **Professor,
Department of Electical Engineering, University of Engineering and Technology, Taxila.

that the 2D gradients in layers are distributed isotropically. Nicolescu and Medioni assume piecewise smooth motion and identify the boundaries along motion discontinuities [4]. Silva and Victor used image contour closest to the occlusion point [5]. Lou, et. al., [6] used model based approach with the assumption that 3D wireframe models of vehicles have already been established [6]. Do, et. al. [7] used memory template with the assumption of partial occlusion [7]. Nakamura and Matsuura proposed multiple cameras with camera masking, [8] Nakamura, et. al. [8]. Other researcher proposed various techniques including robust statistics, Hager, et. al. [9], Black, et. al. [10], Mean-shift algorithm, Comaniciu, et. al. [11]. These methods dealt the tracking results at the region level and are devised for motion detection and cannot satisfy the requirements for tracking in the cluttered background where multiple moving objects exist. Mata et al used a genetic algorithm to develop a vision system for recognizing 2D landmarks i.e. quadrangle Shapes on walls (doors, windows and posters) to guide the robot along a corridor, Mata, et. al. [12]. Li and Yang proposed a vision system that used a genetic algorithm for detecting numerical signs in an outdoor environment, Li, et. al. [13]. However, these systems have not addressed the problem of recognizing occluded landmarks. Yan et al proposed a histogram and intensity mixed template which is effective for the detection of reappeared objects, Yan, et. al. [14] and is not effective while the occlusion persist. Zitnick and Kanade use uniqueness and continuity assumptions which are based on match convergence, Zitnick, et. al. [15] are useful for depth ordering in disparity space and for partial occlusion detection.

Active contour-based methods are more simple, effective and fast. They work well only in partial occlusion and limit the tracking precision to the contour level. Feature-based tracking methods utilizing local features and dependence graphs can handle partial occlusion. However, due to non linear perspective distortion their performance is poor in case of unrelated structures.

In this paper we used two cameras installed in the environment from where the whole work space is observable. It is an extension to our previous work, Khan, et. al. [16] in which we have used single camera for mobile robot localization in indoor environment. The use of multiple cameras for indoor cluttered environment application provides the most reliable and practical solution for mapping uncertainty. Tracking with a single camera easily generates ambiguities due to occlusion and depth. Multiple camera-based visual surveillance systems can be extremely helpful because the surveillance area can be expanded and multiple view information can overcome occlusion. However, multi-cameras application needs the proper selection of positions for camera installation, camera calibration, object identification, image registration, and data fusion and finally the map updates to be used by the active agents of the environment.

2. MULTIPLE ROBOT NAVIGATION APPROACH

Fig. 1 illustrates the overall architecture of our localization system, in which two identical cameras were installed in work space on such positions from where whole of the environment is observable. Both the cameras are connected with CPU. The camera system (left and right) captured images of the scene synchronously. These images are stored on CPU, which performed image processing, object detection, identification and localization. The CPU also performs the image registration and occlusion detection. Finally it updates the environment map and broadcast it within the environment. The mobile robots of the environment are assumed to be capable to communicate with CPU. On receiving the updated map with positional and occlusion information of objects/robots, the mobile robot can perform their path and trajectory planning execution more efficiently. The goal of the proposed technique is to robustly detect mobile robots/objects in 2D plane and provide the 2D map of the environment.

For object identification and occlusion detection using left and right camera images is extremely helpful because multiple view information can resolve the occlusion problem as well as increases the confidence for object identification. The ambiguities generated by one camera can be easily removed by other one. However, increasing the number of camera increases computational time and algorithm complexity.

Calibration is an important consideration when using multiple cameras. For camera calibration we used offline photogrammetric technique which is performed by observing a calibration object whose geometry is known with very good precision. We use a 2D square grid shown in Fig. 2 for this purpose. The square corner coordinates are manually marked from checkerboard's image at different orientations. With the known sizes of the square, MatLab toolbox calculates the camera parameters for each view.

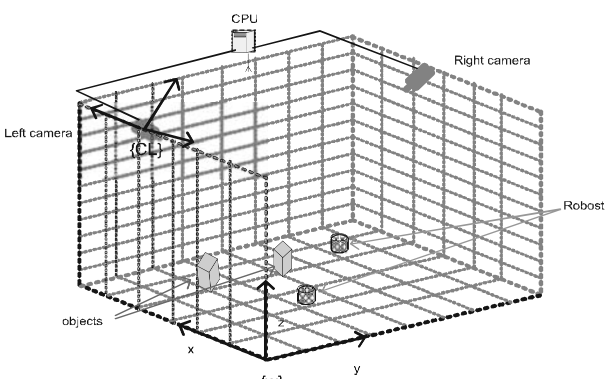

FIG. 1. THE VISION SYSTEM WITH WALL MOUNTED CAMERA NETWORK, CPU, AND DYNAMIC OBJECTS

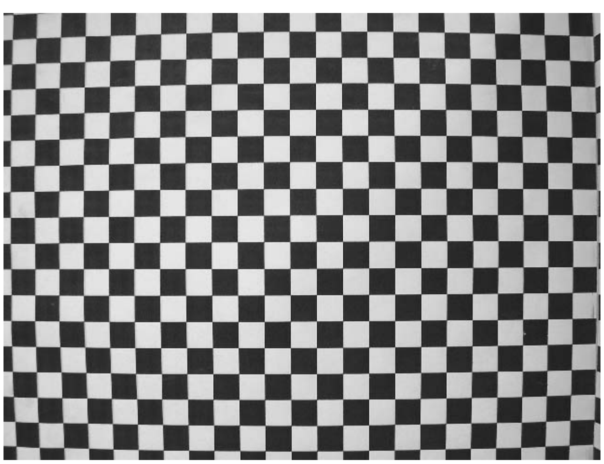

FIG. 2. AN IMAGE OF THE CHECKER BOARD FOR OFFLINE CAMERA CALIBRATION

The overall flow of the activities for the detection of occluded object is illustrated in Fig. 3. While the occlusion detection and mapping system is in operation, the left and right cameras capture synchronously the environment images. These images are subtracted from their respective background images. The objects in the individual images are then identified and localized. In next step image registration is performed for alignment and finally the occlusion detection and map updates.

3. OBJECTS IDENTIFICATION

A number of techniques are available for object identification in the image, i.e. SAD (Sum of Absolute Gray Value Difference), SSD (Sum of Squired Gray Value Difference), NCC (Normalized Cross Correlation), the mean SED (Squired Edge Distance) etc. Due to simplicity and the availability of constant illumination in the indoor environment we selected the SSD approach for locating the objects/robots on image plane presented, Sulaiman, et. al. [17] .

$$S(r,c) = s\left[t(u,v),\ f(r+u,c+v);(u,v) \in T\right] \qquad (1)$$

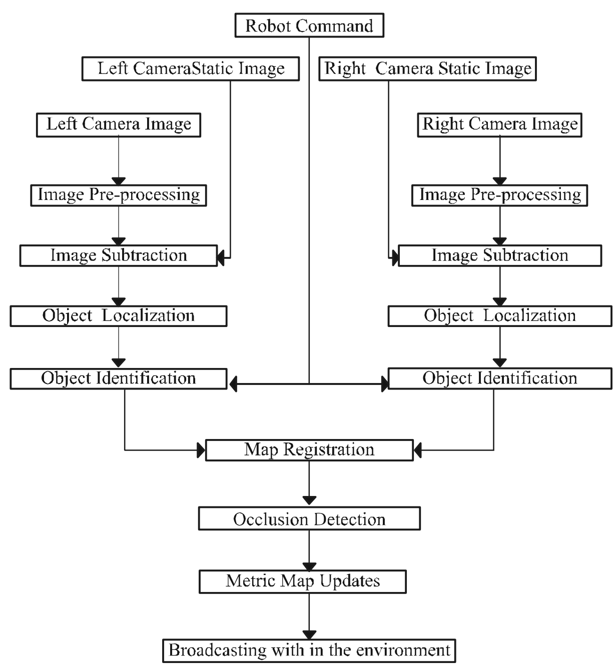

FIG. 3. OBJECT IDENTIFICATION, LOCALIZATION AND OCCLUSION DETECTION SYSTEM

Here $t(u,v)$ is the template and T is its corresponding region of interest. For matching purpose the template was moved over all the locations in image and the similarity measure was obtained.

Gray values in the template and gray values in the shifted region of interest at the current position in the image $f(r+u,c+v)$ are the inputs for similarity measure . It calculates a scalar value that indicates the similarity based on the gray values within the region of interest. Image regions are typically compared using SSD (Sum of Square Differences).

$$ssd(r,c) = \frac{1}{n} \sum_{(u,v) \in T} \left[t(u,v) - f(r+u,c+v) \right]^2 \tag{2}$$

Where n is the number of points in the template region of interest. A similarity measure 0 indicates that image and template are identical. While greater value of similarity measure indicates that image and template are more dissimilar. Further-more the similarity image ssd(r,c) was thresholded which gives a region that contains several adjacent pixels. To obtain a unique location for the template, local minima of the similarity image was selected within each connected component of the thresholded region.

The above ssd similarity measure is sensitive for illumination variation. Luckily due to availability of constant illumination in the indoor environment, it gives reliable results.

4. OBJECTS LOCALIZATION

For calculating the position and size of the objects, we determine the central pixels of the object using centroid method in Equation (3). and determine the contour points using Equations (4-5) as described Lin, et. al. [18]. We use d=f(θ) as the polar representation of the contour graphs. The total number of pixels 'n' that an object occupied is calculated. Fig. 4 is an image of our lab environment, Fig. 5 is the background image, and Fig. 6 is the image after background subtraction showing the noise and floor reflections effects which are removed by using smoothing filters. Fig. 7 shows the contour graph along with the centroid of the object.

FIG. 4. AN IMAGE OF ROBOT IN OUR LAB

FIG. 5. BACK GROUND IMAGE OF ROBOT ENVIRONMENT

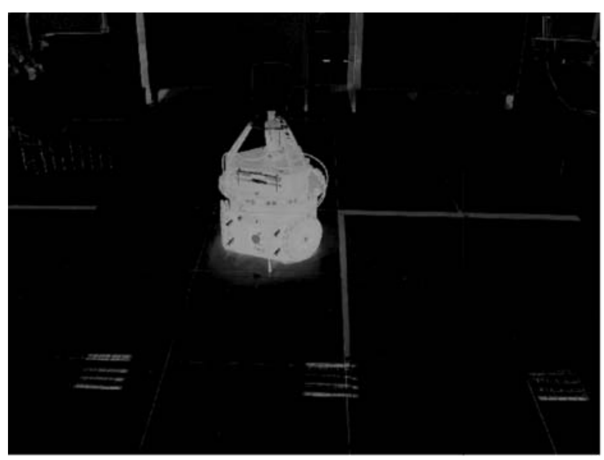

FIG. 6. AFTER BACKGROUND SUBTRACTION

$$x_c = \sum_{i=1}^{n} x_i / n \tag{3}$$

$$y_c = \sum_{i=1}^{n} y_i / n \tag{4}$$

$$\theta = \tan^{-1}\left\{\frac{(y - y_c)}{(x - x_c)}\right\} \tag{5}$$

For object actual size determination and mapping purpose, we divide the workspace of the robot in grid cells and assigned an experimentally predetermined scaling factor to each cell. A lookup table is used for these scaling factors. Addressing the Look-Up-able using the centroid of the object $p(x_i,y_j)$ as an index, we get the scaling factor. For the determination of actual size of an object we multiplied the total number of pixels of an object with the scaling factor.

Based on color and size information of the objects we assign a parametric vector $M_t\{p(x_i,y_j),V,C\}$ to each object, containing the centroid, velocity and color information of the object. These dynamic vectors are created and destroyed based on the appearance and disappearance of the objects in the workspace of the robot. These parametric vectors are used by the robot of the environment in their path planning and obstacle avoidance algorithms.

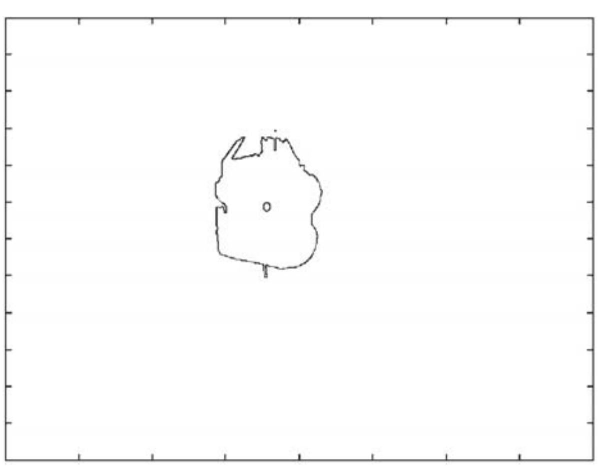

FIG. 7. CONTOUR OF THE ROBOT POSITION

5. OCCLUSION DETECTION

For mobile robot to navigate within its workspace, needs to have the complete knowledge of its environment. Occlusion usually produces the following major types of problems for mobile robots.

The reduction of available workspace for path planning and navigation (due to falsified occupancy).

Due to lack of knowledge of the objects, the probability of collision increases while navigating.

The tracking and surveillance of moving objects.

As shown in Fig. 8 the left and right camera arrangement provides the capability to detect the occluded objects on 2D plane. In Fig. 8(a) the left side camera image taken with an oblique angle, a solid triangle occluded completely a smaller object (toward its right) due to the occlusion effect of the triangle height. Similarly in the image taken from the right side camera, Fig. 8(b) shows the solid triangle as well as the smaller object. While utilizing image from a single

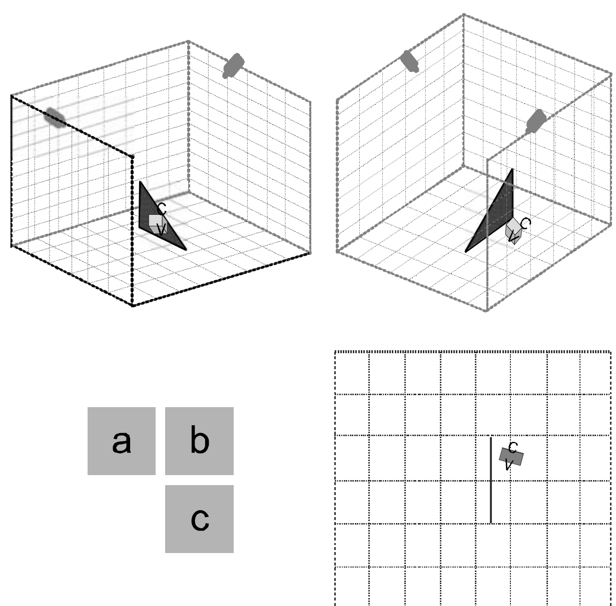

FIG. 8. OCCLUSION DETECTION (A) IN LEFT CAMERA IMAGE THE OCCLUDED ABJECT BEHIND THE SOLID TRIANGLE, IN RIGHT CAMERA IMAGE (B) THE OBJECT IS NOT OCCLUDED, AND (C) AFTER RESOLVING THE OCCLUSION (TOP VIEW)

camera (left), the occlusion problem can not be resolved effectively.

The left and right images are subtracted from their respective background images. After object detection and localization the image registration is performed for alignment. For occlusion detection we applied the following simplified algorithm.

if pixel Sl(r,c)>Th & pixel Sr(r,c)>Th then.....So(r,c)=1

if pixel Sl(r,c)<Th & pixel Sr(r,c)>Th then.....So(r,c)=0

if pixel Sl(r,c)<Th & pixel Sr(r,c)>Th the.....So(r,c)=0

if pixel Sl(r,c)<Th & pixel Sr(r,c)<Th the.....So(r,c)=0

Here Sl(r,c), Sr(r,c) and So(r,c) are transformed left, right and the output images, Th is low-threshold. This operation identified the occluded object in left and right camera images as shown in Fig. 8(c). The images are binarized with gray thresh Matlab function. The threshold was adjusted in such a way to cover most of the objects of the environment. Finally the map of the environment is updated and broadcasted within the environment.

6. EXPERIMENTAL SETUP

The rectangular environment contained a ball and an oval shape smaller object. In the image (640x480) taken from left side camera with oblique angle, as shown in Fig. 9, the smaller object is occluded (covered) by the lager spherical object. If a map is built using single image the smaller object cannot be identified.

In the right camera image (640x480) shown in Fig. 10 the smaller object is clearly visible in front of larger spherical object. For development of map it is necessary to locate all objects of the environment. Therefore fist of all we subtracted the left and right images from back-ground image, then applied projective transformation to both the images.

In next step the left image is rotated for alignment with the right image. Finally the proposed algorithm is applied. The resultant image shows both the objects as shown in Fig.11. The developed algorithm can be successfully applied for more than two objects of the environment.

After the application of our proposed algorithm the 2D circular shape of the spherical object is transformed to oval shape. The reason behind this conversion is that due to the oblique viewing angle from opposite sides, more area under the object is in the view of both the cameras. As our method converts the uncovered area in any image as unoccupied area in the resultant image as is visible in Fig. 11.

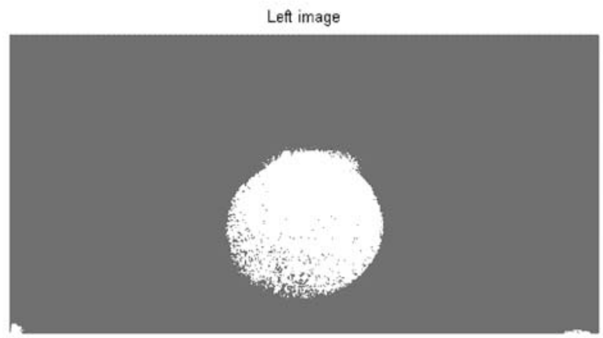

FIG. 9. LEFT CAMERA IMAGE (TRANSFORMED AND ROTATED), THE BIGGER OBJECT OCCLUDED THE SMALLER OBJECT

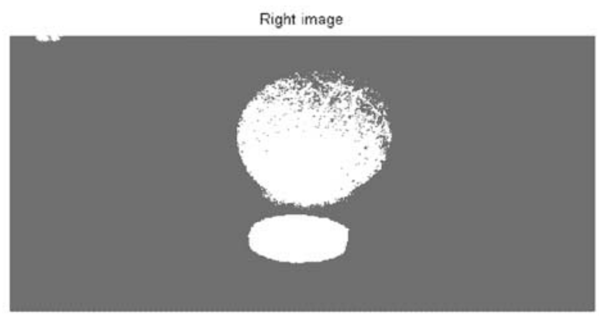

FIG. 10. RIGHT CAMERA IMAGE (TRANSFORMED), THE SMALLER OBJECT IS VISIBLE

FIG. 11. AFTER IMAGE REGISTRATION AND OCCLUSION RESOLVING (COMBINING LEFT AND RIGHT CAMERA IMAGE) ACCORDING TO OUR ALGORITHM

6.1 Application

A Reliable, Easy and Fast Solution to the Occlusion Problem: The occlusion detection is a important parameter for obstacle avoidance algorithms. Our technique reduces the occlusion problem to the limits where it does not matter for path planning in indoor mobile robot applications.

A Real Time Solution to Problem of Multi-Robot Localization and Cooperation: As centralized Computational facility is used for overall localization process. This wave off a lot of processing burden from the robots and helps them in real time localization. It is extremely efficient for multiple robot cooperation.

7. CONCLUSION

We can summarize our contribution as follows:

We have developed a technique that performs intensity based comparison of corresponding pixels of the left and right camera images and produce an output image that is free of occlusion. We tested our technique with two camera systems for occlusion detection for indoor mobile robot localization and mapping. Our technique works even when a large object completely occlude smaller one. In future the proposed occlusion detection technique can be extended to more than two cameras for indoor as well as outdoor environment.

ACKNOWLEDGEMENT

Authors would like to thank the Higher Education Commission of Pakistan, for providing the financial support and facilities for our research work.

REFERENCE

[1] Gecks, T., and Henrich, D., "Multi-Camera Collision Detection Allowing for Object Occlusions", 4th German Conference on Robotics (Robotik) München, Germany May 15-17, 2006.

[2] Wang, X., and Liu, J.L., "Tracking Multiple People Under Occlusion and Across Cameras Using Probabilistic Models", Journal Zhejiang University Science-A, Volume 10, No. 7, pp. 985-996, 2009.

[3] Feldman, D., and Weinshall, D., "Motion Segmentation and Depth Ordering Using an Occlusion Detector", IEEE Transaction on Pattern Analysis and Machine Intelligence, Volume 30, No. 7, pp. 1171-1185, July, 2008.

[4] Nicolescu, M., and Medioni, G., "A Voting-Based Computational Framework for Visual Motion Analysis and Interpretation", IEEE Transaction. Pattern Analysis and Machine Intelligence, Volume 27, No. 5, pp. 739-752, May, 2005.

[5] Silva, C., and Victor, J.S., "Motion from Occlusions", Robotics and Autonomous Systems, Volume 35, pp. 153-162, 2001.

[6] Lou, J., Tan, T., Weiming, H., Yang H., and Maybank, S.J., "3-D Model-Based Vehicle Tracking", IEEE Transactions on Image Processing, Volume 14, No. 10, pp. 1561-1569, October, 2005.

[7] Do, Q.V., Lozo, P., and Jain, L.C., "A Vision System for Partially Occluded Landmark Recognition", LNAI 3809, pp. 1246-1252, 2005.

[8] Nakamura, Y., Matsuura, T., Satoh, K., and Ohta, Y., "Occlusion Detectable Stereo-Occlusion Patterns in Camera Matrix", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 371-378, San Francisco, June 18-20,1996.

[9] Hager, G.D., and Belhumeur, P.N., "Efficient Region Tracking with Parametric Models of Geometry and Illumination", IEEE Transactions on Pattern Analysis Machine Intelligence, Volume 20, pp. 1025-1039, October, 1998.

[10] Black, M.J., and Jepson, A.D., "Eigentracking; Robust Matching and Tracking of Articulated Objects Using a View-Based Representation", Proceedings of European Conference on Computer Vision, Volume 1, pp. 329-342, 1996.

[11] Comaniciu, D., Ramesh, V., and Meer, P., "Kernel Based Object Tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 25, No. 5, May, 2003.

[12] Mata, M., Armingol, J.M., Escalera, A.de.la., and Salichs, M.A., "A Visual Landmark Recognition System for Topological Navigation of Mobile Robots", Proceedings of IEEE International Conference on Robotics and Automation, pp.1124-1129, 2001.

[13] Li, H., and Yang, S.X., "A Behavior-Based Mobile Robot with a Visual Landmark-Recognition System", IEEE/ ASME Transactions on Mechatronics, Volume 8, pp. 390-400, 2003.

[14] Yan, Z., Bo, H., and Zhang, J., "Occlusion Detection and Tracking Method Based on Bayesian Decision Theory", LNCS 4319, pp. 474-482, 2006.

[15] Zitnick, C.L., and Kanade, T., "A Cooperative Algorithm for Stereo Matching and Occlusion Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, No. 7, pp. 675-684, July, 2000.

[16] Khan, T.M., and Choudhry, M.A., "Simple Algorithm for Robot Localization by Using Mathematical Correlation of Image Pixels with Physical Data", IEEE International Conference on Control, Automation and systems, COEX, Seoul, Korea, October 17-20, 2007.

[17] Sulaiman, S.N., Alias, M.F., Ashidi, N.M., and Rahman, M.F., "An Expert Image Processing System on Template Matching", International Journal of Computer Science and Network Security, Volume 7, No. 7, 2007.

[18] Lin, H.J., Kao, Y.T., Yen, S.H., and Wang, C.J., "A Study of Shape-based Image Retrieval", Proceedings of 24th International Conference on Distributed Computing Systems Workshops, [0-7695-2087-10], 2004.