The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

High Performance Vision Tracking System for Mobile Robot Using Sensor Data Fusion with Kalman Filter

Jaehong Park, Wonsang Hwang, Hyun-il Kwon, Jong-hyeon Kim, Chang-hun Lee, M. Latif Anjum, Kwang-soo Kim, and Dong-il "Dan" Cho, *Member*, *IEEE*

***Abstract*— This paper introduces a high performance vision tracking system for mobile robot using sensor data fusion. For mobile robots, it is difficult to collect continuous vision information due to robot's motion. To solve this problem, the proposed vision tracking system estimates the robot's position relative to a target and rotates the camera towards the target. This concept is derived from the human eye reflex mechanism, known as the Vestibulo-Ocular Reflex (VOR), for compensating the head motion. This concept for tracking the target results in much higher performance levels, when compared with the conventional method that rotates the camera using only vision information. The proposed system do not require heavy computing loads to process image data and can track the target continuously even during vision occlusion. The robot motion information is estimated using data from accelerometer, gyroscope, and encoders. This multi-sensor data fusion is achieved using Kalman filter. The proposed vision tracking system is implemented on a two-wheeled robot. The experimental results show that the proposed system achieves excellent tracking and recognition performance in various motion scenarios, including scenarios where camera is temporarily blocked from the target.**

I. Introduction

Recently, robots are deeply involved in various fields, such as military, medical services, and entertainment. Advanced sensors and efficient control technologies are required for robots to perform these tasks. Among the various sensors applied to robots, vision sensors are especially crucial to recognize and acquire information about surroundings. Thus, high performance vision tracking systems are required for mobile robot, which can be adapted for various dynamic circumstances. However, as the demands for the high performance vision sensors have increased, complicated image processing algorithms are needed, increasing the cost of the system.

One approach for developing the vision tracking system is to use image processed visual information [1-3]. Cai and Nillius presented visual tracking for multiple targets using particle filter and Bayesian network [1, 2]. Jung developed detection system for moving target using single camera on a mobile robot [3]. These systems require robust filtering algorithm to track regardless of the background clutter and camera motion. Also, these systems do not function well in the absence of visual information. In the practical case, when object is hidden behind the obstacle or illumination is changed, visual information of the object is not generated sufficiently and stably. On the other hand, some researchers used inertial sensors in order to compute motion information of the robot [4-6]. Kaushik showed compensating for perturbations by using 3-axis inertial sensors, and Kanade developed a novel inertial-aided Kanade-Lucas-Tomasi (KLT) feature tracking method using gyroscope data [4, 5]. In most cases, angular motion compensations are performed using inertial sensors. For compensating the translational motion, the tachometer data is used for supplementing visual information [6]. Recently, the Vestibulo-Ocular Reflex (VOR) based vision stabilization systems show fast and accurate stabilization performances [7]. However, the goal of the presented system is to stabilize vision against small rotational movement of the robot rather than to track a certain target. Cho presented the VOR based target tracking system using accelerometer information [8]. His paper computes the translational motion of the robot by using accelerometer information. From this information, a vision sensor which is mounted on the robot rotates towards the selected target, periodically compensating errors from visual information.

In this paper, we propose high performance vision tracking system for two-wheeled robot using multi-sensor data fusion technique. The sensor fusion techniques for active vision system are already presented [9, 10]. Applications of these fusion methods mainly involved the camera attitude estimation, motion estimation, navigation, and 3 dimensional structure reconstructions. On the other hands, the main purpose of the proposed system is to keep the line of sight fixed on the target even when the robot has large and continuous transient motion. The proposed system is able to

Manuscript received March 10, 2010. (Write the date on which you submitted your paper for review.) This work was supported partly by the R&D program of the Korea Ministry of Knowledge and Economy (MKE) and the Korea Evaluation Institute of Industrial Technology (KEIT). [2008-F-037-01, Development of HRI Solutions and Core Chipsets for u-Robot]

Jaehong Park, Wonsang Hwang, Hyun-il Kwon, Jong-hyeon Kim, and Chang-hun Lee are with Automation and Systems Research Institute (ASRI), Inter-university Semiconductor Research Center (ISRC), and School of Electrical Engineering and Computer Science, Seoul National University, Kwanak-ku, Seoul, 151-742, Korea (e-mail: jaehong@snu.ac.kr, wshwang@snu.ac.kr, hikwon@snu.ac.kr, jhkim84@snu.ac.kr, chlee84@snu.ac.kr, latif@snu.ac.kr)

Kwangsoo Kim is with Department of Control and Instrumentation Engineering, Hanbat National University, Yuseong-gu, Deajeon, 305-719, Korea (e-mail: kskim@hanbat.ac.kr)

Dong-il "Dan" Cho is with Automation and Systems Research Institute (ASRI), Inter-university Semiconductor Research Center (ISRC), and School of Electrical Engineering and Computer Science, Seoul National University, Kwanak-ku, Seoul, 151-742, Korea (corresponding author; phone: 82-2-880-6488; fax: 82-2-877-9304; e-mail: dicho@snu.ac.kr).

compute the random motion of the robot by using inertial sensors, encoders, and vision sensor. The key concepts of the proposed system are based on the human eye reflex mechanism, known as the VOR. By adapting human eye reflex mechanism, vision tracking system is able to reduce the dependency on visual information and track the selected target during locomotion. Using the robot motion information, the proposed system achieved excellent tracking performance in the absence of visual information.

This paper is organized as follows. Section II describes the concept and overall system organizations of proposed vision tracking system. Section III shows the coordinate systems which are used for explaining the robot motion and target location for vision tracking system. Section IV gives brief explanation of sensor data fusion mechanism. Section V and Section VI show the implementation and experimental results of the proposed vision tracking system. Finally, we concluded this paper in Section VII with some relevant discussion.

II. Concept and System Organization

In order to improve performance, the human eye reflex mechanism is applied to the proposed vision tracking system. In the human eye reflexes, during locomotion the head oscillates passively on the trunk and the images of the visual surroundings tend to smear across the retina. Image motion is minimized and visual recognition is facilitated by the initiation of eye movements in the opposite direction of head movement. These compensatory eye movements are generated by the VOR and the Opto-Kinetic Reflexes (OKR). Both reflexes are therefore activated simultaneously and generate eye movements in the same direction but differ in their dynamic properties [11].

The block diagram of the proposed vision tracking system is shown in Fig. 1. When robot starts to move according to robot trajectory input, robot motion information are generated by inertial sensors and encoders of robot wheels. This information is transferred to the controller for estimation of robot location and heading angle relative to the target. This information corresponds to the VOR mechanism. Target position information in image plane, which is produced in image processing block, is translated to controller. This information is related to OKR mechanism. The target position information in image plane is the angle difference between the center of the image plane and center of the target. In the image processing block, extracting and matching features of the target are performed using initial and present image. Then target is detected by matching result, and the angle difference between the center of the target and the image plane can be estimated by using a pin-hole camera model [12, 13]. For the purpose of target detection for estimating angle difference, the Speeded Up Robust Features (SURF) image processing method is used [14]. To estimate the expected camera rotation angle, the multi-sensor data fusion is achieved using Kalman filter and filter fusion algorithm in the controller block. The output of controller is the fed to the plant, which contains the

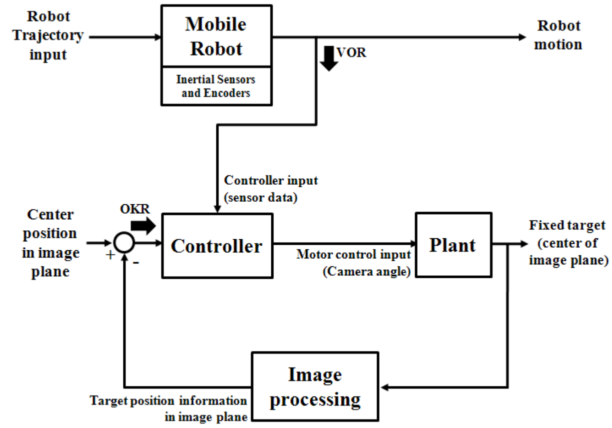


Fig. 1. Block diagram of the proposed vision tracking system

camera and motor. Finally, motor controls gaze of the camera, and target is fixed in the center of the image plane during locomotion.

III. Coordinate System

The two-wheeled robot has in-plane motion with 3 degrees of freedom (DOF), which contain linear motions in x/y axis and rotation motions in z-axis. Considering these characteristics, we applied Cartesian coordinate system to express robot motion and location as shown in Fig. 2 [15]. There are two coordinate systems: global coordinate (centered at O) and intermediate coordinate system (centered at O_k) as shown in Fig 2. In this coordinate system, r_k, vector from robot to target, can be expressed as

$$\underline{r}_k = \underline{r}_o - \underline{p}_k \tag{1}$$

where, $\underline{r}_0$ is vector from origin to target, $T(X_T, Y_T)$ is the target position in global coordinate, and $\underline{p}_k$ is robot location vector in kth step. In the kth step, according to present coordinate system, $\underline{r}_k$ can be expressed as

$$\underline{r}_k = \begin{bmatrix} \cos\theta_k & \sin\theta_k \\ -\sin\theta_k & \cos\theta_k \end{bmatrix} \cdot (\underline{r}_o - \underline{p}_k) \tag{2}$$

where θ_k is robot heading angle with respect to initial coordinate system. Using equation (1) and (2), the expected rotation angle of camera $\angle r_k$ is calculated by

$$\angle r_k = \tan^{-1}\left(\frac{r_{y,k}}{r_{x,k}}\right) = \tan^{-1}\left(\frac{-(X - p_{x,k}) \cdot \sin\theta_k + (Y - p_{y,k})\cos\theta_k}{(X - p_{x,k})\cos\theta_k + (Y - p_{y,k})\sin\theta_k}\right) \tag{3}$$

IV. Fusion Mechanism of Sensor Data

As mentioned in Section II, the multi-sensor data fusion for estimating rotation angle of camera is performed in controller block. The simple block diagram of controller is shown in Fig. 3. In this block, Kalman filter is mainly used as the sensor data fusion, and considering characteristics of each sensor the slip detector is added. After the robot location and heading

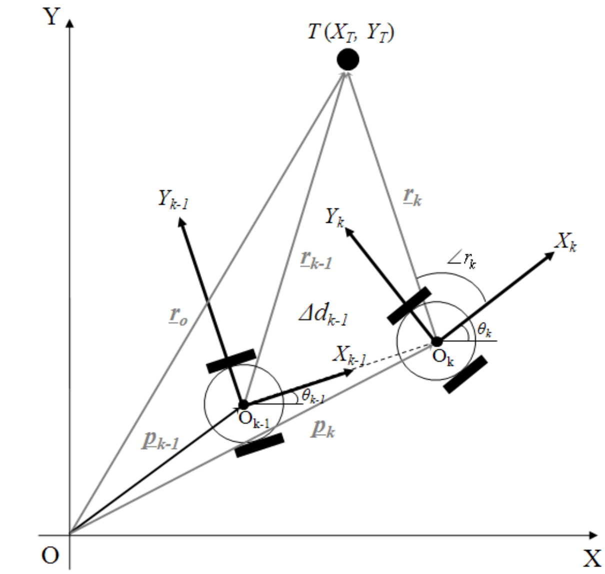


Fig. 2. Cartesian coordinate system for vision tracking system

angle are estimated by the extended Kalman filter and indirect Kalman filter, one of them is selected in filter output fusion block based on slip detector output.

A. Extended Kalman Filter Block

The extended Kalman filter block estimate location and heading angle of robot using accelerometer data, gyroscope data, and vision data [16]. Considering the nonlinear property of inertial sensor data system, the extended Kalman filter is used. In this block, encoder data are not used. Thus this estimation result is reliable on slippery or rugged roads with wheel slip [17]. However, errors of the accelerometer, such as systematic errors and random errors, cause an unbounded growth in the error of integrated measurements [18]. In order to compensate the accelerometer errors, the vision data, which is the angle difference between the center of the target and the image plane is used for measurement update of the extended Kalman filter. Considering above two characteristics, the output of an extended Kalman filter is utilized only for slip condition. The state model for the designed extended Kalman filter is as follows

$$\begin{aligned}
&\underline{x}(k) = \left[p_x(k) \quad p_y(k) \quad v(k) \quad \theta(k) \right]^T \\
&\underline{u}(k) = \left[0 \quad 0 \quad a(k) \quad \Omega(k) \right]^T \\
&p_x(k) = p_x(k-1) + v(k-1)\cdot \Delta t \cos(\theta(k-1)) + w_1^{EKF}(k) \\
&p_y(k) = p_y(k-1) + v(k-1)\cdot \Delta t \sin(\theta(k-1)) + w_2^{EKF}(k) \\
&v_R(k) = v(k-1) + a(k-1)\cdot \Delta t + w_3^{EKF}(k) \\
&\theta_R(k) = \theta(k-1) + \Omega(k-1)\cdot \Delta t + w_4^{EKF}(k)
\end{aligned} \tag{4}$$

where $\underline{x}(k)$ is state vector, $\underline{u}(k)$ is input, p_x is robot x-coordinate, p_y is robot y-coordinate, v is robot velocity, θ_R is robot heading angle, a is measured acceleration from accelerometer data, Ω is measured angular velocity from gyroscope data, and w_n^{EKF} is process noise.

The measurement model with measurement noise, η^{EKF}, is shown below,

$$\begin{aligned}
z(k) &= \angle r(k) \\
&= \tan^{-1}\left(\frac{(X_T - p_x(k))\cdot \sin\theta_R(k) + (Y_T - p_y(k))\cdot \cos\theta_R(k)}{(X_T - p_x(k))\cdot \cos\theta_R(k) - (Y_T - p_y(k))\cdot \sin\theta_R(k)} \right) + \eta^{EKF}
\end{aligned} \tag{5}$$

B. Indirect Kalman Filter Block

The indirect Kalman filter block estimates the location and heading angle of robot using gyroscope data, encoder data, and vision data [19]. Because encoder data is not reliable on slippery surface, this estimation result is used for non-slip condition. The error state model of the designed indirect Kalman filter is as follows

$$\begin{bmatrix} \delta x(k+1) \\ \delta y(k+1) \\ \delta\varphi(k+1) \\ \delta s_R(k+1) \\ \delta s_L(k+1) \\ \delta d(k+1) \\ \delta\phi(k+1) \\ \delta g_s(k+1) \\ \delta b_g(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cos\varphi(k)\cdot\frac{R+L}{2} & \sin\varphi(k)\cdot\frac{L}{2} & \sin\varphi(k)\cdot\frac{R}{2} & 0 & 0 & 0 & 0 \\ 0 & 1 & -\sin\varphi(k)\cdot\frac{R+L}{2} & \cos\varphi(k)\cdot\frac{L}{2} & \cos\varphi(k)\cdot\frac{L}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{R}{d} & -\frac{L}{d} & \frac{L-R}{d^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \Omega & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta x(k) \\ \delta y(k) \\ \delta\varphi(k) \\ \delta s_R(k) \\ \delta s_L(k) \\ \delta d(k) \\ \delta\phi(k) \\ \delta g_s(k) \\ \delta b_g(k) \end{bmatrix} + \begin{bmatrix} \frac{\sin\varphi(k)}{2} & \frac{\sin\varphi(k)}{2} & 0 \\ \frac{\cos\varphi(k)}{2} & \frac{\cos\varphi(k)}{2} & 0 \\ \frac{1}{d} & -\frac{1}{d} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_R \\ \varepsilon_L \\ \varepsilon_g \end{bmatrix} + \underline{w}^{IKF} \tag{6}$$

where, x and y are the location information of robot which are computed by encoder data, R and L are moving distances which are computed by right and left wheel encoder data, ε_R and ε_L are noise of R and L, s_R and s_L are radius error of right wheel and left wheel, d is distance error between right and left wheel, g_s is scale factor of gyroscope, b_g and ε_g are bias and random noise of gyroscope, φ is computed robot heading angle from encoder data, Φ is computed robot heading angle from gyroscope data, and $\underline{w}^{IKF}$ is process noise vector.

In the absence of vision data, measurement update is performed by calculating the difference between heading angles based on encoder data and gyroscope data. Otherwise, measurement update is performed using vision data and calculated heading angle. The equations for measurement update are as follows.

$$\psi_c(k) = \hat{\varphi}(k) - \hat{\phi}(k) = \delta\varphi(k) - \delta\phi(k) + \eta_1^{IKF} \tag{7}$$

$$\begin{aligned}
&\underline{\psi} = H_k \cdot \underline{x} + \underline{\eta}_2^{IKF} \\
&\underline{\psi} = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}, \quad \psi_1 = \psi_2 \\
&\psi_1 = \tan^{-1}\left(\frac{Y_T - (y + \delta y)}{(x + \delta x) - X_T} \right) - (\varphi + \delta\varphi) + \eta_{2-1}^{IKF} \\
&\psi_2 = \tan^{-1}\left(\frac{Y_T - (y + \delta y)}{(x + \delta x) - X_T} \right) - (\phi + \delta\phi) + \eta_{2-2}^{IKF}
\end{aligned} \tag{8}$$

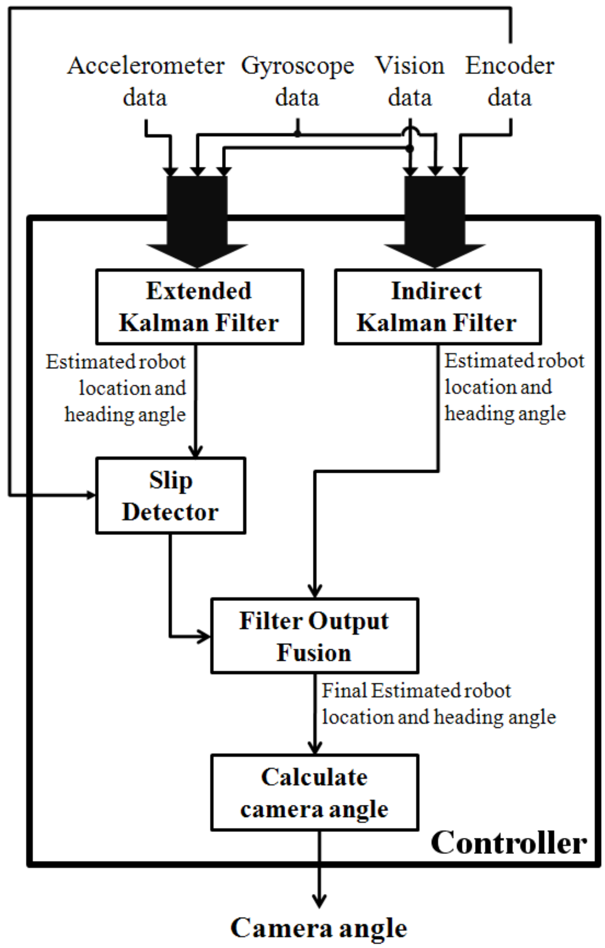


Fig. 3. The block diagram of controller

where ψ_c is angle difference between φ and Φ, ψ is computed robot heading angle error from vision sensor data, and η^{IKF} is measurement noise

Above are the equations for measurement update without vision data (7) and equations for measurement update with vision data (8). Due to the Output Data Rate (ODR) difference, two measurement update equations are used.

C. *Slip Detector Block*

Slip detector decide slip level using comparison between the encoder data and the output of the extended Kalman filter. When difference between the encoder data and the extended Kalman filter output is larger than threshold level, slip detector shows slip occurrence. Threshold level is defined by repetitive experiments for various surface conditions.

D. *Filter Output Fusion Block*

Filter output fusion block fuse the extended Kalman filter and indirect Kalman filter based on the slip detector output. In the slip condition, final robot location and heading angle are defined using estimation results of the extended Kalman filter. In the non-slip condition, final robot location and heading angle are defined using estimation results of the indirect Kalman filter. The output of controller is expected camera rotation angle, which is calculated by final robot location and heading angle information.

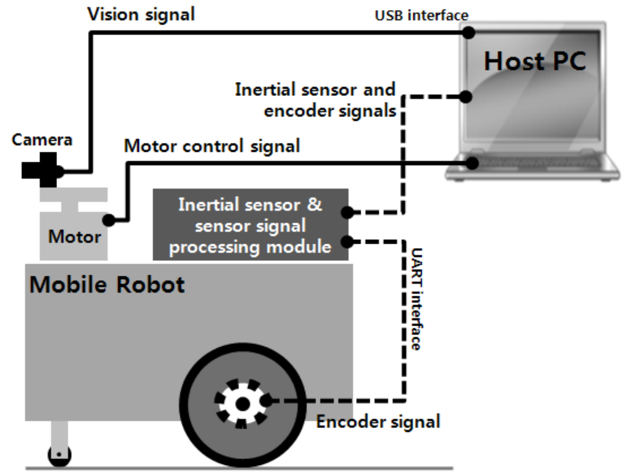
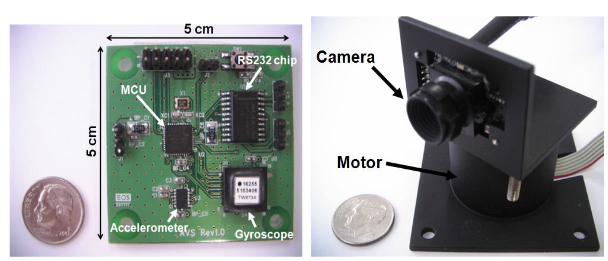


Fig. 4. Schematic diagram of hardware configuration

V. Implementation of Proposed System

The schematic diagram of proposed vision tracking system is shown in Fig. 4. The proposed vision tracking system consists of an accelerometer, a gyroscope, a camera, a dc motor for rotating the camera, and a host computer. Robot motion information, such as acceleration data, angular velocity data, and encoder data, are transferred from sensors to the host computer by using UART interface. The vision signal and motor control signal are transferred by using a USB interface.

For signal processing of the inertial sensors, sensor module is developed as shown in Fig. 5 (a). The sensor module to detect the rotational and translational motions of robot functions in a similar manner as the semicircular canals and otoliths in human vestibular system. The sensor module contains an accelerometer (KXPS5-3157, Kionix, Inc.), a gyroscope (ADIS16255, Analog Devices, Inc.), a micro controller unit, and RS232 chip [20, 21]. The actuation module with camera is also developed as shown in Fig. 5 (b). Required camera motion for wheeled robot is panning motion for tracking the target. The single camera (SPC 520NC, PHILPS) and DC motor (Series 2619, MicroMo Electronics, Inc.) are implemented for actuation module [22, 23]. The resolution and the frame rate of camera are set to 320×240 pixels and 30 fps for achieving ideal processing of the image. As shown in Fig. 6, the developed systems are implemented



(a) (b)

Fig. 5. Picture of sensor module (a) and actuation module with camera (b) for vision tracking system

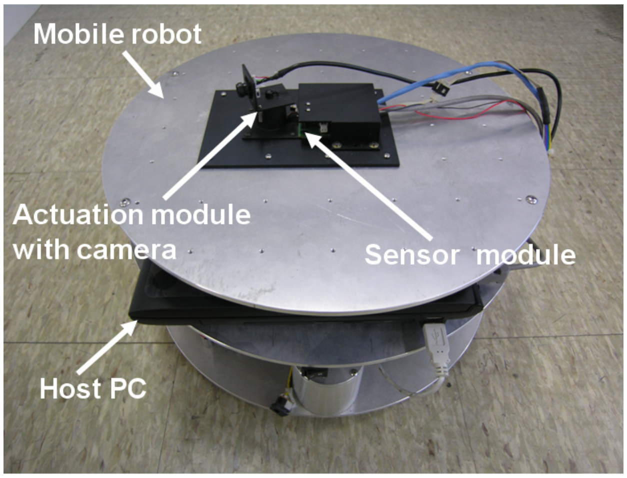


Fig. 6. Developed vision tracking system mounted on mobile robot

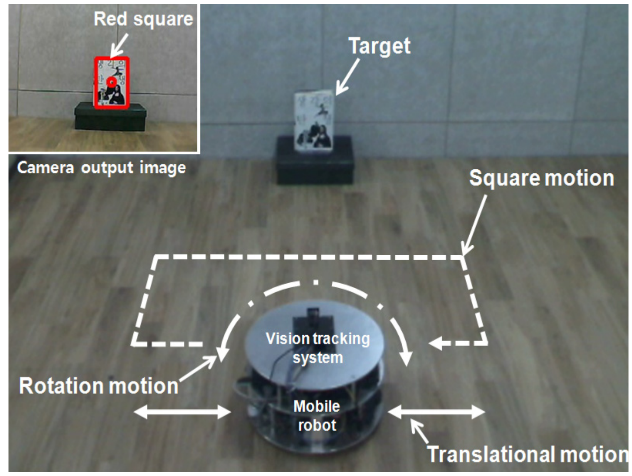


Fig.7. Picture of experimental environment setup with robot trajectories

on the two-wheeled robot (Mobile Robot, Customer & Robot Co., LTD.). The host computer is also mounted on mobile robot platform. Overall software program was designed for controlling the robot and the dc motor using C++, which is based on the Windows environment.

VI. Experimental Results

Experimental setup for proposed vision tracking system is shown in Fig. 7. In the experiment, initial distance between vision system and target is 1.8 m, while driving and rotation velocity of mobile robot is 0.2 m/s and 30 deg/s, respectively. In recognition test, SURF based recognition program is used. In the case of recognition success, red square is formed on the edge of the target as shown in Fig. 7. Tracking success rate and recognition rate of developed vision tracking system are tested for various robot trajectories. Tracking success rate is 100 % in all three cases, while recognition rate is 100 % except in the case of square motion. Due to the blurring effect in image plane, the target cannot be recognized successfully in some output images of square motion. Table 1 shows the summarized experimental results for various robot trajectories. All of the presented results are the average of five experiments.

To verify robustness of proposed system to various circumstances, we performed experiments for two cases. First, tracking success rate is tested when an obstacle temporarily block the camera sight. As shown in Fig. 8 (b) and (c), complete vision information is not produced because of poor recognition results. In this case, purely vision-based tracking system cannot track the target with this deficiency of visual information. However, using motion information of robot, the proposed system fixes camera sight to the target while the target is temporarily blocked by the obstacle. The target is tracked and recognized continuously after the obstacle is removed. In the experimental results, the tracking success rate of proposed vision tracking system is 100 % for the square motion. In the second place, the tracking success rate is tested when outside illumination changes while robot is moving. The target is recognized easily in the normal lighting

TABLE I

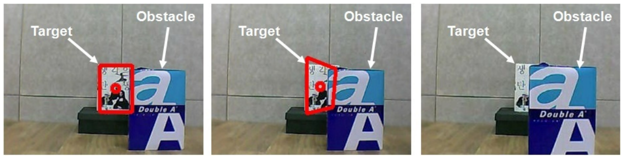
Summary of Experimental Results for Various Robot Trajectories

Robot trajectory	Tracking success rate (# of successful tracking/ # of total image frame)	Recognition rate (# of successful recognition/ # of total image frame)
Translational motion (±0.5 m)	100 % (129/129)	100 % (129/129)
Rotational motion (±90 deg)	100 % (108/108)	100 % (108/108)
Square motion	100 % (170/170)	92 % (156/170)

condition but it does not appear clearly in the image frame while illumination is lowered as shown in Fig. 9 (b) and (c). In these cases, either unstable vision information is produced from recognition results, or vision information is not produced at all. Therefore, purely vision-based tracking system cannot track the target in the absence of normal visual information. However, in spite of the illumination change, the proposed vision tracking system can track the target continuously because of robot motion information. In the experimental results, tracking success rate of the proposed vision tracking system is 100 % while outside illumination is changed occasionally during the square motion of the robot. In above both cases, purely vision-based tracking system is not able to perform the tracking function after the target is blocked by obstacles or illumination is changed. Therefore, the tracking success rates of purely vision-based tracking system are declined remarkably. The experimental results for various circumstances are summarized in Table 2. As previous experimental results, all of the presented results are the average of five experiments.

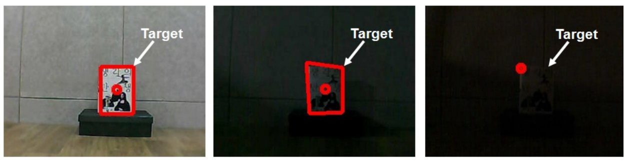
VII. Conclusions

In this paper, high performance vision tracking system using multi-sensor data fusion through Kalman filter is



(a) (b) (c)

Fig. 8. Output images of camera with the obstacle. (a) The obstacle blocks the target (perfect red square is formed), (b) the obstacle partially blocks the target (distorted red square is formed), (c) the obstacle blocks the target mostly (red square is not formed)



(a) (b) (c)

Fig. 9. Output images of camera with illumination change. (a) Outside light on (perfect red square is formed), (b) outside light partially off (distorted red square is formed), (c) outside light off (red square is not formed)

TABLE II

SUMMARY OF EXPERIMENTAL RESULTS FOR VARIOUS CIRCUMSTANCES

Circumstance	Vision tracking system	Tracking success rate (# of successful tracking/ # of total image frame)
Obstacle blocks target	Proposed system	100 % (175/175)
	Purely vision-based tracking system	33 % (57/175)
Outside illumination change	Proposed system	100 % (172/172)
	Purely vision-based tracking system	37 % (63/172)

presented. To improve the performance in vision tracking, the proposed vision tracking system uses not only the visual information but also robot motion information, which is computed using inertial sensors and wheel encoders. This concept is derived from the human eye reflex mechanism for compensating the head motion. In the control scheme, the extended Kalman filter and indirect Kalman filter are used for estimating location and heading angle of the robot. To fuse outputs of the two Kalman filters, a slip detector is implemented to detect the slip occurrence. Based on the slip detector output, final robot motion and expected rotation angle of camera is estimated using one of the two filters. The proposed vision tracking system is mounted on a mobile robot, and the tracking success rates and recognition rates are evaluated for various robot trajectories and circumstances. In the experimental results, the proposed vision tracking system demonstrates excellent tracking and recognition performance in various motion scenarios.

REFERENCES

[1] Y. Cai, N. de Freitas, and J. J. Little, "Robust visual tracking for multiple targets," in *Proc. European Conf. Computer Vision*, vol. 4, pp. 107-118, 2006

[2] P. Nillius, J. Sullivan, and S. Carlsson, "Multi-target tracking - linking identities using baysian network inference," in *Proc. Computer Vision and Pattern Recognition*, pp. 2187 - 2194, 2006.

[3] B. Jung and G. S. Sukhatme, "Detection moving objects using a single camera on a mobile robot in an outdoor environment," in *Proc. Conf. Intelligent Autonomous Systems*, pp. 980-987, 2004.

[4] R. Kaushik, M. Marcinkiewicz, J. Xiao, S. Parsons, and T. Raphan, "Implementation of bio-Inspired Vestibulo-Ocular Reflex in a quadrupedal Robot," in *Proc. 2007 IEEE Int. Conf. Robotics and Automation*, pp 4861-4866, 2007.

[5] M. Hwangbo, J. –S. Kim, and T. Kanade, "Inertial-aided KLT feature tracking for a moving camera," in *Proc.2009 IEEE/RJS Int. Conf. Intelligent Robots and Systems*, pp. 1909-1916, 2009

[6] Z. Jia, A. Balasuriya, and S. Challa, "Sensor fusion based 3d target visual tracking for autonomous vehicles with the out-of sequence measurements solution," in *Proc. 2005 IEEE Int. Conf. Robotics and Automation*, pp. 1841-1846, 2005.

[7] A. Lenz, T. Balakrishnan, A. G. Pipe, and C. Melhuish, "An adaptive gaze stabilization controller inspired by the vestibulo-ocular reflex," *Bioinspiration & Biomimetics*, vol.3, pp. 1-11, 2008.

[8] H. K. Ouh, W. Bahn, J. Park, W. S. Hwang, H. I. Kwon, K. S. Kim, and D. I. Cho, "VOR based target tracking system with an accelerometer for mobile robots," in *Proc. 6th Int. Conf. Ubiquitous Robots and Ambient Intelligence*, pp. 133-137, 2009.

[9] P. Corke , J. Lobo , and J. Dias, "An introduction to inertial and visual sensing," *International Journal of Robotics Research*, vol.26, pp. 519-535, 2007.

[10] P. Gemeiner, P. Einramhof, and M. Vincze, "Simultaneous motion and structure estimation by fusion of inertial and vision data," *International Journal of Robotics Research*, vol.26, pp. 591-605, 2007.

[11] N. Dieringer, "Vestibular system," *Encyclopedia of life science*, 2006

[12] R. Y. Tsai, "A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras lenses," *IEEE Trans. on Robotics and Automation*, vol. 3, pp. 323-344, 1987.

[13] Y. Lui, T. S. Huang, and O. D. Faugeras, "Determination of camera location from 2-D to 3-D line and point correspondences," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp.28-37,1990

[14] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," in *Proc. of the 9th European Conf. Computer Vision*, pp. 404-417, 2006.

[15] A. Alessandri, G. Bartolini, P. Pavanati, E. Punta, and A. Vinci, "An application of the extended Kalman filter for integrated navigation in mobile robotics," in *Proc. of the American Control Conference*, pp. 527-531, 1997.

[16] D. Simon, "*Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*," John Wiley & Sons, 2006

[17] J. S. Lim, T. H. Lee, T. D. Ahn, K. H. Yoo, A. R. Lee, and D. I. Cho, "Sensor fusion for mobile robot navigation using accelerometers and odometer," in *Proc. 13th Int. Conf. Advanced Robotics*, pp. 343-348, 2007.

[18] J. Borenstein, H.R. Everett, L. Feng, and D. Wehe "Mobile robot positioning & sensors and techniques," Invited paper for *Journal of Robotic Systems, Special Issue on Mobile Robots*, vol. 14, no. 4, pp. 231 – 249, 1997

[19] K. C. Park, D. Chung, H. Chung, and J. G. Lee, "Dead reckoning navigation of a mobile robot using an indirect Kalman filter," in *Proc. 1996 IEEE/SICE/RSJ Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, pp.132-138, 1996.

[20] http://www.kionix.com/accelerometers/accelerometer-KXPS5.html

[21] http://www.analog.com/en/other-products/multi-chip/adis16255/products/product.html

[22] http://www.consumer.philips.com/c/webcams-headsets/36739/cat/kr/

[23] http://www.micromo.com/uploadpk/2619SR_IE2-16_DFF.pdf