Journal of Marine Science and Technology

Volume 22 | Issue 3 Article 6

GLOBAL LOCALIZATION USING DEAD-RECKONING AND KINECT SENSORS FOR ROBOTS WITH OMNIDIRECTIONAL MECANUM WHEELS

Ching-Chih Tsai
Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan, R.O.C., cctsai@nchu.edu.tw

Feng-Chun Tai
Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan, R.O.C.

Yu-Cheng Wang
Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan, R.O.C.

GLOBAL LOCALIZATION USING DEAD-RECKONING AND KINECT SENSORS FOR ROBOTS WITH OMNIDIRECTIONAL MECANUM WHEELS

Acknowledgements

The authors gratefully acknowledge financial support from National Science Council, the Republic of China, under contract NSC 100-2218-E-005-033-.

Journal of Marine Science and Technology, Vol. 22, No. 3, pp. 321-330 (2014)
DOI: 10.6119/JMST-013-0909-4

GLOBAL LOCALIZATION USING DEAD-RECKONING AND KINECT SENSORS FOR ROBOTS WITH OMNIDIRECTIONAL MECANUM WHEELS

Ching-Chih Tsai, Feng-Chun Tai, and Yu-Cheng Wang

Key words: dead-reckoning, extended Kalman filter, global localization, KINECT sensor, low-cost.

ABSTRACT

This paper presents a low-cost global pose localization method using least-squares method and extended Kalman filter (EKF) for an anthropomorphous dual-arm mobile robot (ADAMR) driven by omnidirectional Mecanum wheels in indoor environments. This method is developed by fusing measurements from the KINECT sensor and four encoders mounted on the omnidirectional Mecanum wheels. The KINECT sensor is used to recognize landmarks in the working environments of the robot, and then obtain the azimuth angles and distances between these landmarks and the robot. Based on these measured information, a static global pose initialization algorithm using least-squares method is applied to estimate both unknown start-up position and orientation of the robot. Once the initial pose has been roughly determined, an EKF approach is proposed to fuse the odometric measurements from the four encoders and the azimuth angles and distances from the KINECT sensor to corresponding landmarks. Simulations are conducted to show performance of the proposed method. Experimental results are used to illustrate that the proposed method provides accurate static estimation of both unknown initial poses of the robot.

I. INTRODUCTION

The self-localization capability of an autonomous mobile robot is very important for path tracking control and planned navigation in any given environment [1, 8, 12]. The global localization capability of an autonomous mobile robot is critical for its navigation missions at its prespecified environments [4-6, 11]. Global pose initialization and pose tracking are the two main problems in the global localization technology. The global pose initialization problem determines the robot's initial pose with respect to some reference frame, using no priori or a priori map or landmark information, while the pose tracking problem is concerned with, given initial robot pose from the user, how to continuously maintain the robot's pose using internal and external sensors. Pose initialization is considered crucial for planned or autonomous navigation of the robot because it is relevant not only at start-up, but also during operation for recovery in case of pose localization failures [9].

The dead-reckoning method, which is based on the encoded or odometric information from the wheels, has extensively been utilized to calculate the current location of an autonomous mobile robot. However, this method suffers from the accumulation errors that are caused by wheel slippage or by mechanical tolerances and surface roughness. Hence, the robot may fail to keep track of its true location over long distances. The odometric errors can be reduced by fusing landmark observations using external sensors. Many sensors, such as sonar sensors [2], laser range finders [2], and cameras [7], have been used to detect landmarks and obtained the required measurements. Ultrasonic sensors have widely been applied to develop the useful and economical external sensing systems for localization of autonomous mobile robots [13]. However, sonar sensors and laser range finders introduce several risks: the outbound energy may affect the very characteristics that the sensor is attempting to measure. Furthermore, an active sensor may suffer from interference between its signal and signals beyond its control. Vision is our most powerful sense, providing us with an enormous amount of information about the environment and enables rich, intelligent interaction in dynamic environments. It is therefore not surprising that a great deal of effort has been devoted to providing machines with sensors that mimic the capabilities of the human vision system [10]. Recently, KINECT has been widely utilized for robot localization due to its low cost [3].

Paper submitted 06/17/12; revised 12/10/12; accepted 09/09/13. Author for correspondence: Ching-Chih Tsai (e-mail: cctsai@nchu.edu.tw).
Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan, R.O.C.

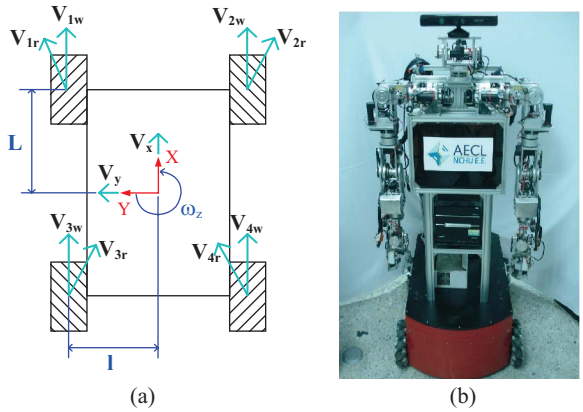

Fig. 1. Anthropomorphous Two-Armed Robots with Omnidirectional Mecanum Wheels. (a) Its simplified geometry with respect to the moving frame; (b) a laboratory-built example of ADAMR with Omnidirectional Mecanum Wheels.

Extended Kalman filter has been proposed for the pose estimation of autonomous mobile robots to obtain better accuracy of pose initialization and pose racking using KINECT [14]. However, the proposed method in [14] has not been experimentally verified yet.

The objectives of this paper are to apply the least-square method and the EKF approach to a low-cost localization method for an ADAMR by fusing dead-reckoning and KINECT measurements, and to demonstrate how the proposed method achieves the better performance of the robot localization. This technique is expected to be useful in improving accuracy and robustness of the ADAMR's position and orientation estimates, and to provide a low-cost global localization method for wheeled mobile robots.

The remainder of this paper is outlined as follows. Section II briefly introduces the kinematics model in the world frame and dead-reckoning model of the system. Section III presents the landmark detection algorithm of the KINECT sensor. This section will utilize Hough transform method to extract the landmarks from the image obtained by KINECT, thus achieving landmark detection. Section IV shows a least-squares pose initialization algorithm to achieve global localization. Section V briefly introduces the extended Kalman filtering method. Section VI introduces the localization using natural landmarks. Computer simulations and experimental results are respectively reported in Sections VII and VIII to show the merit of the proposed methods for global pose initialization and pose tracking of the ADAMR. Section IX concludes this paper.

II. SYSTEM AND DEAD-RECKONING MODEL

1. Kinematics Model in the World Frame

Fig. 1 shows the simplified geometry of the ADAMR with Omnidirectional Mecanum wheels. The inverse kinematics of its base in the moving frame is described by:

$$\mathbf{V}_{\mathrm{w}} = \mathbf{J}_o \cdot \mathbf{V}_{\mathbf{o}} \tag{1}$$

where $\mathbf{V}_{\mathbf{w}} = [V_{1\mathrm{w}} \quad V_{2\mathrm{w}} \quad V_{3\mathrm{w}} \quad V_{4\mathrm{w}}]^T \in R^{4\times 1}$ is the velocity vector in Cartesian coordinates; $\mathbf{V}_{\mathbf{0}} = [v_x \quad v_y \quad \omega_z]^T \in R^{3\times 1}$ is the wheel velocity vector corresponding to the angular velocity. Moreover,

$$\mathbf{J}_{\mathbf{0}} = \begin{bmatrix} 1 & -1 & -(l+L) \\ 1 & 1 & +(l+L) \\ 1 & 1 & -(l+L) \\ 1 & -1 & +(l+L) \end{bmatrix} \in R^{4\times 3} \tag{2}$$

is a transformation matrix; the two parameters l and L are depicted in Fig. 1. To formulate inverse kinematic models of the ADAMR in a global frame, we define that the vector $\begin{bmatrix} x_{\mathrm{w}} & y_{\mathrm{w}} & \theta \end{bmatrix}^T$ denotes the position and orientation of the vehicle in the world frame, and let the moving and world frames have the rotational matrix expressed by

$$\begin{bmatrix} \dot{x}_{\mathrm{w}} \\ \dot{y}_{\mathrm{w}} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ \omega_z \end{bmatrix} \tag{3}$$

Thus, the inverse kinematics model of the MWOR is then expressed by

$$\mathbf{V}_{\mathrm{w}} = \mathbf{J}(\theta) \begin{bmatrix} \dot{x}_{\mathrm{w}} \\ \dot{y}_{\mathrm{w}} \\ \dot{\theta} \end{bmatrix} \tag{4}$$

where $\mathbf{J}(\theta) = \begin{bmatrix} \sqrt{2}\sin(\theta_1) & -\sqrt{2}\cos(\theta_1) & -(l+L) \\ \sqrt{2}\cos(\theta_1) & \sqrt{2}\sin(\theta_1) & (l+L) \\ \sqrt{2}\cos(\theta_1) & \sqrt{2}\sin(\theta_1) & -(l+L) \\ \sqrt{2}\sin(\theta_1) & -\sqrt{2}\cos(\theta_1) & (l+L) \end{bmatrix}$

Since the transformation $\mathbf{J}(\theta)$ exists its pseudo inverse matrix, $\mathbf{J}^{\dagger}(\theta)$ such that $\mathbf{J}(\theta)^{+}\mathbf{J}(\theta) = \mathbf{I}_3$, the forward kinematics of the ADAMR in the world frame becomes

$$\begin{bmatrix} \dot{x}_{\mathrm{w}} \\ \dot{y}_{\mathrm{w}} \\ \dot{\theta} \end{bmatrix} = \mathbf{J}^{\dagger}(\theta)\mathbf{V}_{\mathrm{w}} \tag{5}$$

where $\theta_1 = \theta + \pi/4$

$$\mathbf{J}^{\dagger}(\theta) = \frac{1}{4}\begin{bmatrix} \sqrt{2}\sin(\theta_1) & \sqrt{2}\cos(\theta_1) & \sqrt{2}\cos(\theta_1) & \sqrt{2}\sin(\theta_1) \\ -\sqrt{2}\cos(\theta_1) & \sqrt{2}\sin(\theta_1) & \sqrt{2}\sin(\theta_1) & -\sqrt{2}\cos(\theta_1) \\ -\dfrac{1}{L+l} & \dfrac{1}{L+l} & -\dfrac{1}{L+l} & \dfrac{1}{L+l} \end{bmatrix}$$

2. Dead-Reckoning

The purpose of the dead-reckoning of the ADAMR is, given a correct initial pose, to continuously keep trace of its correct poses with respect to the world frame. Dead-reckoning is the real-time calculation of the robot's position from wheel encoder measurements. The dead-reckoning problem of the ADAMR can be easily solved by using the numerical approach, the velocity information from the encoders mounted on the driving wheels and the forward kinematical model in (5). To find the continuous poses of the robot from (5), many existing numerical approaches, such as the Euler's formula, the Runge-Kutta method and so on, can be employed according to the required numerical accuracy and the step size. One of the simplest dead-reckoning methods is based on the second-order Runge-Kutta formula which approximates the pose differentiation of the robot by the following equation.

$$\begin{bmatrix} x_w(k) \\ y_w(k) \\ \theta(k) \end{bmatrix} = \begin{bmatrix} x_w(k-1) \\ y_w(k-1) \\ \theta(k-1) \end{bmatrix} + T\frac{(\mathbf{J}^+(\theta(k-1))\mathbf{V}_w(k-1) + \mathbf{J}^+(\theta(k))\mathbf{V}_w(k))}{2} \tag{6}$$

where T is the sampling period and sufficiently small; $x(k)$, $y(k)$ and $\theta(k)$ denote respectively the present measurements of x(t), y(t) and $\theta(t)$ at the k^{th} sampling instant.

Notice that Eq. (6) gives the recursive formula to approximately obtain the robot's position and orientation at the next sampling instant. Worthy of mention is that the accuracy of the dead-reckoning method increases as the sampling period decreases. However, if the sampling period T is fixed and not significantly small, then the accuracy of the dead-reckoning method can be improved by using the fourth-order or higher-order Runge-Kutta numerical Method. Thus, this kind of dead-reckoning method can often be used as a useful pose tracking approach for the robot over short traveling distances. This is because, no matter what kind of the numerical method is used, the proposed dead-reckoning method always suffers from unavoidable accumulation errors caused by slippage, surface roughness, surface friction and even the mechanical structure of the ADAMR.

III. KINECT SENSOR

KINECT is a motion sensing input device used by Microsoft for video gaming console. This sensor offers a natural interface using physical gestures and spoken commands for users to control the game without touching the controller. This sensor is equipped with an RGB camera and a pair of depth sensors. The KINECT sensor has the spatial (x/y) resolution of 3 mm at 2 m distance from sensor, while its effective operation range is about 0.8 m-3.5 m with the resolution of 1 cm at 2 m distance from the sensor. In this paper, a KINECT sensor is mounted on the robot as shown in Fig. 1(b), and the height of the KINECT is arbitrarily set.

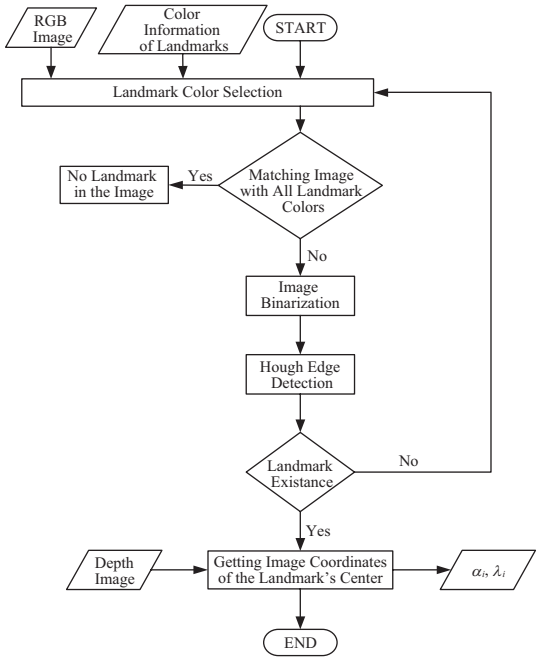

Fig. 2. Flow diagram of KINECT sensing.

1. KINECT Calibration

This subsection describes a calibration method converting the KINECT disparity into the corresponding distances. Acquiring the depth image of KINECT is based on the light coding technology [14]. The coded light is captured by the IR camera in order to produce the KINECT disparity matrix. The relationship between the disparity and actual depth value for a normal stereo system is given by,

$$z = \frac{b \times f}{d} \tag{7}$$

where z is the depth, b is the horizontal baseline between the cameras, f is the focal length of the cameras in pixels, and d is the disparity. However, a zero KINECT disparity does not correspond to infinite distance, so the KINECT disparity is related to a normalized disparity by the subsequent relation

$$d = \frac{1}{8}(d_{off} - d_{kinect}) \tag{8}$$

where d_{off} is an offset value for a given KINECT device, and d_{kinect} is the KINECT disparity which provides 2048 levels of sensitivity in VGA resolution with 11-bit depth. The factor 1/8 is used due to the fact that d_{kinect} is in 1/8 pixel units. However this calibration results may slightly differ from one sensor to another. Therefore each sensor should be calibrated separately before using it for depth estimation.

2. Landmark Detection Using KINECT

Fig. 2 depicts the detailed flowchart of the proposed

Table 1. The degree value of H with respect to its color.

Degrees	0	60	120	240	300
Color	Red	Yellow	Green	Blue	Magenta

landmark detection algorithm for the circular artificial landmarks. In the algorithm, the HSI color space is adopted because its representation of colors is similar to how the human eyes sense. The HSI model represents every color with three components: hue (H), saturation (S), and intensity (I). The H component utilizes an angle between [0,360] degrees, as shown in Table 1. The S and I components utilize a real number between [0,1] separately. The value $I = 0$ means black and 1 means white. The value of S represents how much the color is mixed with white color.

In the measurement update step, the ADAMR detects landmarks around it to estimate its pose. In this section, circles with different colors are used as landmarks. The RGB images and their depth values are obtained by the KINECT sensor. Hough transform filters are utilized to detect landmarks in the RGB image. The landmarks are distinguished from each other by HSI color model. The azimuth angle α_i with respect to the heading direction of ADAMR and the distance λ_i to the i^{th} landmark $L_i(x_{L_i}, y_{L_i}, z_{L_i})$ at time k can be calculated using the KINECT measurements. The value of the measurement function can be determined by

$$Z_i = \begin{bmatrix} \alpha_i \\ \lambda_i \end{bmatrix} = \begin{bmatrix} \tan^{-1}\dfrac{y_{L_i} - \hat{y}_{k^-}}{x_{L_i} - \hat{x}_{k^-}} - \hat{\theta}_{k^-} \\ \sqrt{(x_{L_i} - \hat{x}_{k^-})^2 + (y_{L_i} - \hat{y}_{k^-})^2 + (z_{L_i} - \hat{z}_{k^-})^2} \end{bmatrix} \quad (9)$$

IV. POSE INITIALIZATION

This section presents a hardware configuration and a static pose initialization approach for localization of the robot using KINECT measurements. Suppose that m landmarks are installed at the known positions denoted by

$L_1 = (x_1, y_1, z_1)^T$, $L_2 = (x_2, y_2, z_2)^T$, …, $L_m(x_m, y_m, z_m)^T$, respectively and the robot pose is denoted by $(x, y, z, \theta)^T$. The KINECT data from the landmarks to the robot can be converted into their corresponding distances from the landmarks to the robot. Let $r_1, r_2, \cdots, r_m$ denote the distances from the landmark positions $(x_1, y_1, z_1)^T$, $(x_2, y_2, z_2)^T$, …, $(x_m, y_m, z_m)^T$ to the robot, respectively. Then one obtains the following set of equations

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 = r_1^2 \\ (x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 = r_2^2 \\ \vdots \\ (x_m - x)^2 + (y_m - y)^2 + (z_m - z)^2 = r_m^2 \end{cases} \quad (10)$$

which give the following matrix equation for the robot position (x, y, z).

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{B} \quad (11)$$

where

$$A = \begin{bmatrix} 2(x_2 - x_1) & 2(y_2 - y_1) & 2(z_2 - z_1) \\ 2(x_3 - x_2) & 2(y_3 - y_2) & 2(z_3 - z_2) \\ \vdots & \vdots & \vdots \\ 2(x_1 - x_p) & 2(y_1 - y_p) & 2(z_1 - z_p) \end{bmatrix},$$

$$B = \begin{bmatrix} r_1^2 - r_2^2 - (x_1^2 - x_2^2) - (y_1^2 - y_2^2) - (z_1^2 - z_2^2) \\ r_2^2 - r_3^2 - (x_2^2 - x_3^2) - (y_2^2 - y_3^2) - (z_2^2 - z_3^2) \\ \vdots \\ r_p^2 - r_1^2 - (x_p^2 - x_1^2) - (y_p^2 - y_1^2) - (z_p^2 - z_1^2) \end{bmatrix}$$

In (11), the least-squares method can be used to solve for the robot position $(x, y, z)^T$, i.e.

$$\mathbf{X} = (\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^{\mathrm{T}} \cdot \mathbf{B} = \begin{bmatrix} x & y & z \end{bmatrix}^T \quad (12)$$

if the matrix **A** has a full rank. Note that if the matrix **A** is not of full rank, then the pseudo inverse method will be adopted to solve for the robot position in (11). Once the robot position $(x, y, z)^T$ has been calculated via (12), the robot heading is then found by

$$\theta = \frac{1}{m}\sum_{i=1}^{m} \tan^{-1}\frac{y_i - y}{x_i - x} - \alpha_i \quad (13)$$

As a result, the initial static pose (position and orientation) of the robot can be uniquely determined using (12) and (13).

V. POSE TRACKING

This section is dedicated to elucidating the EKF method used for the pose estimation of the ADAMR. Although researchers have verified that particle filter (PF) get more accurate results than the extended Kalman filter (EKF) does [3], but when the method be extended into 3D space, PF will consume too much calculation resource so that it will cause control delay. In order to avoid that, we used the EKF instead the PF. The initial pose estimate $\hat{x}_0$ is taken as the initialized pose in the previous section. The extrapolation steps of the filter are executed using the dead-reckoning information. The state updating step will be preceded only when a landmark is detected. Whenever a landmark is not detected, the predicted state is taken as the state estimate for the next iteration of the filter, i.e. $\hat{x}_{k^+} = \hat{x}_{k+1^-}$.

In this paper, the process and measurement noises of the

ADAMR system are assumed to be stationary Gaussian white noise with zero mean. As a result, the noise covariance matrices, Q_k, and R_k are considered diagonal [7].

$$w_k \equiv N(0, Q_k),\ Q_k = \begin{bmatrix} \sigma_{x_k}^2 & 0 & 0 \\ 0 & \sigma_{y_k}^2 & 0 \\ 0 & 0 & \sigma_{\theta_k}^2 \end{bmatrix} \tag{14}$$

$$v_k \equiv N(0, R_k),\ R_k = \begin{bmatrix} \sigma_{\alpha_k}^2 & 0 \\ 0 & \sigma_{\lambda_k}^2 \end{bmatrix} \tag{15}$$

The process noise variances of (x, y) coordinates and orientation are represented by $\sigma_{x_k}^2$, $\sigma_{y_k}^2$, and $\sigma_{\theta_k}^2$ respectively, while $\sigma_{\alpha_k}^2$ and $\sigma_{\lambda_k}^2$ are the measurement noise variances.

Consider the following nonlinear discrete-time system model and measurement model of the ADAMR:

$$X_k = f(X_{k-1}, k-1) + w_{k-1} \tag{16}$$

$$Z_k = h(X_k, k) + v_k \tag{17}$$

where $f(\cdot)$ and $h(\cdot)$ are nonlinear functions of the state $X_k = [x(k)\ y(k)\ z(k)\ \theta(k)]^T$, and twice differentiable. Note that both nonlinear functions, $f(\cdot)$ and $h(\cdot)$, can be found in (6) and (9). Thus, the proposed EKF whose prediction and estimation equations can be easily stated as follows:

(i) One-step prediction

$$\hat{x}_{k^-} = f_{k-1}(\hat{x}_{k-1^+}, k-1) \tag{18}$$

$$\hat{z}_k = h_k(\hat{x}_{k^-}, k) \tag{19}$$

$$P_{k^-} = \Phi_{k-1}^{[1]} P_{k-1^+} \Phi_{k-1}^{[1]\,T} + Q_{k-1} \tag{20}$$

(ii) Estimation (measurement update)

$$\hat{x}_{k^+} = \hat{x}_{k^-} + K_k(z_k - \hat{z}_k) \tag{21}$$

$$P_{k^+} = P_{k^-} - P_{k^-} H_k^{[1]T} (H_k^{[1]} P_{k^-} H_k^{[1]T} + R_k^{-1})^{-1} H_k^{[1]} P_{k^-} \tag{22}$$

$$K_k = P_{k^+} H_k^{[1]T} R_k^{-1} \tag{23}$$

where

$$\Phi_{k-1}^{[1]} = \left.\frac{\partial \phi_{k-1}}{\partial x}\right|_{x=\hat{x}_{k-1^-}},\ H_k^{[1]} = \left.\frac{\partial h_k}{\partial x}\right|_{x=\hat{x}_{k^-}}.$$

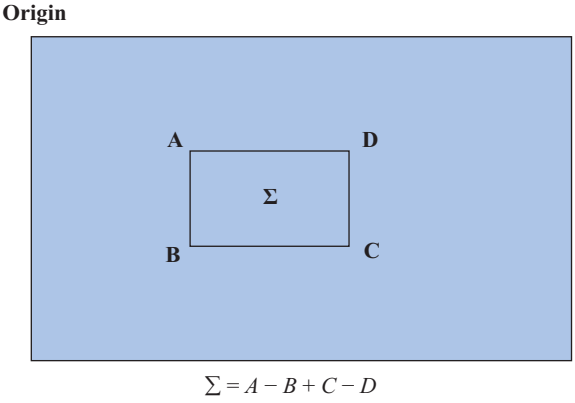

Fig. 3. The advantage of integral images.

VI. LOCALIZATION AND MAPPING USING NATURAL LANDMARKS

This section is aimed to develop a KINECT-based EKF approach using SURF for global localization and mapping of the ADAMR navigating in partially known indoor environments. Since SURF algorithm has been widely used to extract and match features of objects of interests, it can be applied to detect features of natural landmarks in the indoor environments.

1. Feature Detection of Natural Landmarks

An essential part of establishing a map of the environment with image features is the robust matching. This section is aimed to briefly describe the procedure of image feature detection and matching with SURF algorithm.

SURF is a novel scale and rotation invariant detector and descriptor and is utilized to extract and match features between two successive images. The detector is based on the Hessian matrix because of its good performance in scale invariant. More precisely, SURF replaces the time-consuming computation of determinant of the Hessian matrix with box filters and integral images.

The integral images play a very important role in using box type of filters. It is defined as follows,

$$I_\Sigma\left((x,y)^T\right) = \sum_{i=0}^{x}\sum_{j=0}^{y} I(i,j) \tag{24}$$

denotes the sum of the pixels within a rectangular region of the image, and the rectangular region is formed by origin and radius. Therefore, it takes only three additions to compute the sum of pixels inside the region of any size, as shown in Fig. 3.

The original Hessian matrix is defined as follows

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \tag{25}$$

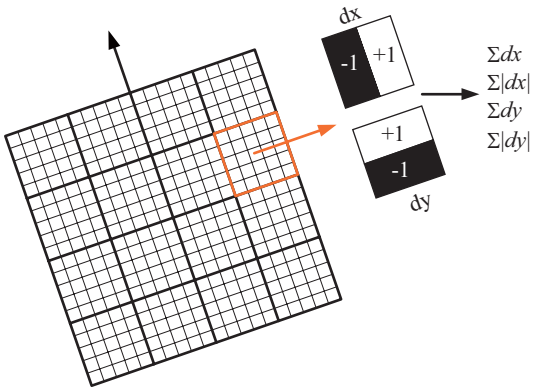

Fig. 4. Descriptor of the feature based on Haar.

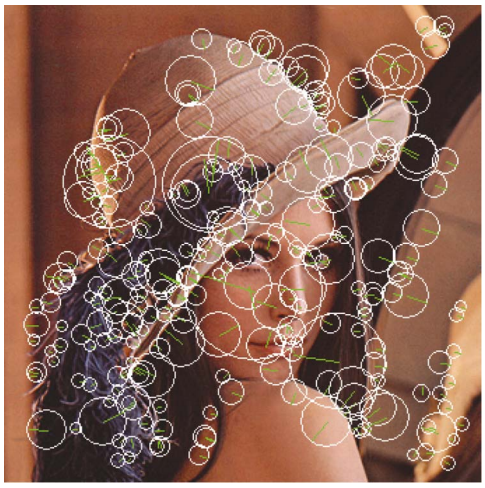

Fig. 5. Experiment of detected interest points.

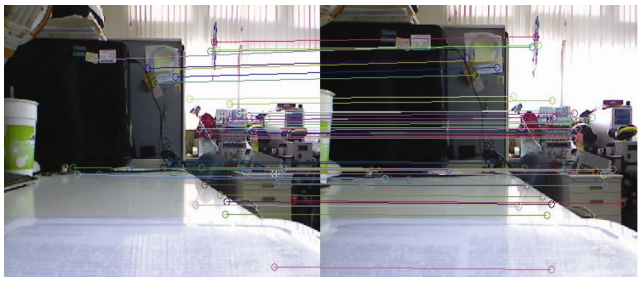

Fig. 6. Experiment of feature matching in two sequential still images.

where $L_{ij}(x,\sigma)$ express the convolution of image and Gaussian second order partial derivative $\frac{\partial^2 g(\sigma)}{\partial i \partial j}$. In order to reduce the computation, the approximate determinant of the Hessian matrix is used and defined as follows

$$\det(H_{approx}) = D_{xx}D_{yy} - (wD_{xy})^2 \tag{26}$$

where D_{ij} is the approximation for the L_{ij}, the second order Gaussian partial derivative in x-direction, y-direction and xy-direction, and w is weight of the filter responses used to balance the expression for the Hessian's determinant.

Due to the use of box filters and integral images, the scale space is analyzed by up-scaling the filter size rather than iteratively reducing the image size as SIFT algorithm. This makes better computational efficiency.

Subsequently the descriptor vectors of the detected feature are calculated. Each one is calculated by the feature itself and surrounding region centered around it. To build the descriptor, the size of feature descriptor region is set to 20×20 pixels and divided into 4×4 sub-regions. For each square, the wavelet responses are computed from 5×5 samples. Consequently the descriptor vector of a descriptive region has total $4 \times 4 \times 4 = 64$ elements, as shown in Fig. 4. For each field, we collect the sums dx, $|dx|$; dy, and $|dy|$, computed relatively to the orientation of the grid Fig. 5 shows an experiment of the detected interest points.

In the matching stage, the trace of the Hessian matrix for the underlying interest point is included. Fig. 6 shows an experiment of matching result of the fifty detected interest points.

2. EKF-Based Global Localization and Mapping Algorithm

EKF is utilized to address the global localization and mapping problem for global localization and mapping problem of the ADAMR. This section is devoted to describing the EKF algorithm and proposing a procedure of the global localization and mapping.

In doing so, the robot's dead-reckoning method and the least-square pose initialization algorithm are the same to those proposed in sec 2.2 and sec. 5. Once an EKF-based global localization and mapping is implemented, we have to estimate the pose of the robot X_w and the features in the environment L. The state vector is composed by the pose of the robot and the positions of the m matching features from the environment, as shown by

$$x = \begin{bmatrix} x_w \\ y_w \\ z_w \\ \theta \end{bmatrix} \text{ and } L = \begin{bmatrix} l_1 & l_2 & \cdots & l_m \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & & x_m \\ y_1 & y_2 & \cdots & y_m \\ z_1 & z_2 & & z_m \end{bmatrix}$$

In the measurement update step, KINECT detects and matches features around it with SURF to estimate the pose of the mobile base. Let α_i be the azimuth angle with respect to the heading direction of ADAMR, ϕ_i be the elevation angle with respect to the horizontal plane, and λ_i be the distance from the robot to the i^{th} landmark $l_i(x_i, y_i, z_i)$ at time k. Note that these three measurements can be calculated using the KINECT measurements.

Thus, the measurement vector function is formulated as below;

$$Z=\begin{bmatrix}\lambda_1\\ \alpha_1\\ \phi_1\\ \vdots\\ \lambda_m\\ \alpha_m\\ \phi_m\end{bmatrix}=\begin{bmatrix}\sqrt{(x_1-\hat{x}_w)^2+(y_1-\hat{y}_w)^2+(z_1-\hat{z}_w)^2}\\ \tan^{-1}\dfrac{y_1-\hat{y}_w}{x_1-\hat{x}_w}-\hat{\theta}_w\\ \tan^{-1}\dfrac{z_1-\hat{z}_w}{\sqrt{(x_1-\hat{x}_w)^2+(y_1-\hat{y}_w)^2}}\\ \vdots\\ \sqrt{(x_m-\hat{x}_w)^2+(y_m-\hat{y}_w)^2+(z_m-\hat{z}_w)^2}\\ \tan^{-1}\dfrac{y_m-\hat{y}_w}{x_m-\hat{x}_w}-\hat{\theta}_w\\ \tan^{-1}\dfrac{z_m-\hat{z}_w}{\sqrt{(x_m-\hat{x}_w)^2+(y_m-\hat{y}_w)^2}}\end{bmatrix} \quad (27)$$

With (25), the proposed EKF-based global localization and mapping can be easily stated as follows:

Step 1:

(i) Utilize the least-squares pose initialization algorithm to find the global position and orientation of the ADAMR, i.e. (x_{w0}, y_{w0}, z_{w0}, θ_0).

(ii) Utilize SURF to detect features L_{SURF_0} and calculate the 3D-location L_{depth} of the points with respect to all pixels within the depth image with the initial pose of the mobile base.

(iii) Obtain the data L_0 composed by the repeated points simultaneously appeared in L_{SURF_0} and L_{depth}.

$$L_0=\begin{bmatrix}l_1 & l_2 & \cdots & l_m & \cdots & l_n\end{bmatrix} \quad (28)$$

Step 2:

(i) One-step Prediction

$$\hat{x}_{k^-}=f_{k-1}(\hat{x}_{k-1^+},k-1) \quad (29)$$

$$P_{k^-}=F_{k-1}^{[1]}P_{k-1^+}F_{k-1}^{[1]T}+Q_{k-1} \quad (30)$$

where $f(\cdot)$ represents the dead-reckoning method, i.e.

$$\begin{bmatrix}x_w(k)\\ y_w(k)\\ \theta(k)\end{bmatrix}=\begin{bmatrix}x_w(k-1)\\ y_w(k-1)\\ \theta(k-1)\end{bmatrix}+T\frac{J^{\dagger}(\theta(k-1))V_w(k-1)+J^{\dagger}(\theta(k))V_w(k)}{2}$$

and $F_{k-1}^{[1]}$ means the Jacobian matrix with respect to *f*.

$$F_{k-1}^{[1]}=\left.\frac{\partial f_{k-1}}{\partial x}\right|_{x=\hat{x}_{k-1^-}}$$

(ii) Use SURF to detect features L_{SURF_1} and match with L_{SURF_0}. Obtain the data L_1 composed by the location of the *m* matched features from L_0.

(iii) Use L_1 to compute

$$\hat{z}_k=h_k(\hat{x}_{k^-},k) \quad (31)$$

where *h*(·)denotes equation (27), i.e.

$$Z=\begin{bmatrix}\lambda_1\\ \alpha_1\\ \phi_1\\ \vdots\\ \lambda_m\\ \alpha_m\\ \phi_m\end{bmatrix}=\begin{bmatrix}\sqrt{(x_1-\hat{x}_w)^2+(y_1-\hat{y}_w)^2+(z_1-\hat{z}_w)^2}\\ \tan^{-1}\dfrac{y_1-\hat{y}_w}{x_1-\hat{x}_w}-\hat{\theta}_w\\ \tan^{-1}\dfrac{z_1-\hat{z}_w}{\sqrt{(x_1-\hat{x}_w)^2+(y_1-\hat{y}_w)^2}}\\ \vdots\\ \sqrt{(x_m-\hat{x}_w)^2+(y_m-\hat{y}_w)^2+(z_m-\hat{z}_w)^2}\\ \tan^{-1}\dfrac{y_m-\hat{y}_w}{x_m-\hat{x}_w}-\hat{\theta}_w\\ \tan^{-1}\dfrac{z_m-\hat{z}_w}{\sqrt{(x_m-\hat{x}_w)^2+(y_m-\hat{y}_w)^2}}\end{bmatrix}$$

Step 3:

(i) Estimation (Measurement Update)

$$\hat{x}_{k^+}=\hat{x}_{k^-}+K_k(z_k-\hat{z}_k) \quad (32)$$

$$P_{k^+}=P_{k^-}-P_{k^-}H_k^{[1]T}(H_k^{[1]}P_{k^-}H_k^{[1]T}+R_k^{-1})^{-1}H_k^{[1]}P_{k^-} \quad (33)$$

$$K_k=P_{k^+}H_k^{[1]T}R_k^{-1} \quad (34)$$

Where $H_k^{[1]}$ means the Jacobian matrix with respect to *h*.

$$H_k^{[1]}=\left.\frac{\partial h_k}{\partial x}\right|_{x=\hat{x}_{k^-}}.$$

(ii) Calculate and update the 3D-position of the points L_{depth} with respect to all pixels within the depth image with the update pose of the mobile base.

(iii) Update the data L_0 given by the repeat points appeared in L_{SURF_1} and L_{depth}.

(IV) Replace L_{SURF_0} to L_{SURF_1}.

Step 4:

Repeat step 2 and step 3.

Fig. 7 shows the flowchart of the Step 1. And Fig. 8 shows the iteration of Step 2 and Step 3.

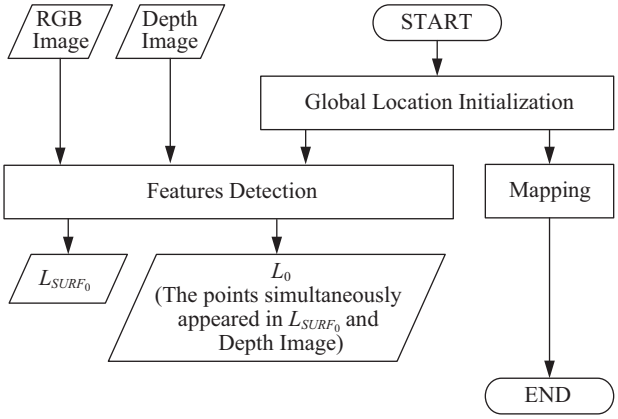

Fig. 7. Flowchart of the proposed global localization and mapping algorithm at the start-up phase.

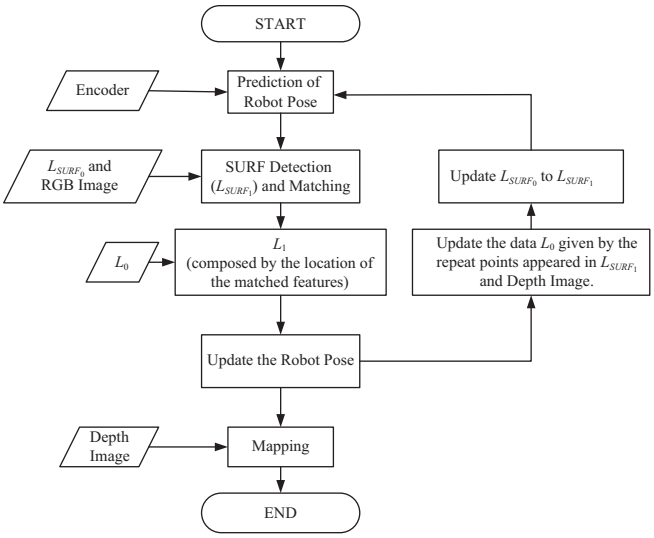

Fig. 8. Deployment of the four landmarks in the second experiment.

VII. SIMULATION AND DISCUSSION

In this section, one simulation result is performed to examine the performance of the proposed EKF-based global localization algorithms of ADAMR using artificial landmarks, four encoders and KINECT sensor. This simulation is done using Matlab/Simulink codes. In the simulation, the pose localization algorithm is only verified when the ADAMR moves in the flat environments.

1. The Proposed Pose Tracking Method

To do the simulation of the proposed pose initialization and tracking, the second simulation sets up the size of environment as 4 m × 3 m, installs twelve circle landmarks whose positions and color are known. All the landmarks are fixed at the same height. In the beginning of pose initialization, four landmarks are given in front of the ADAMR. Then, a rounded-corner trajectory is set in Fig. 9, where there is a clearance of 0.5 m against the wall. After these steps, each distance between two

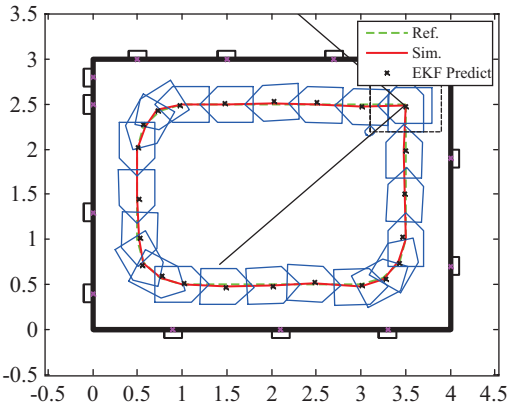

Fig. 9. Simulation results of the proposed pose initialization and posture tracking method.

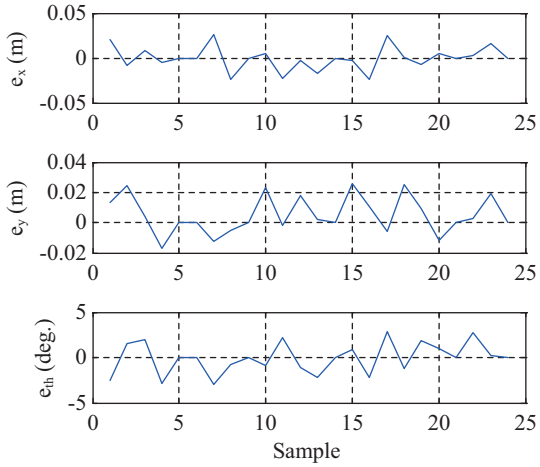

Fig. 10. The errors of pose tracking in simulation.

localized points is 0.5 m. Fig. 9 and Fig. 10 depict the simulation results and the pose tracking errors. The simulation results reveal that the proposed pose initialization and continuous pose tracking methods are capable of giving accurate pose estimates for position and orientation of the ADAMR.

2. The Proposed Pose Tracking Method

To do the simulations of the proposed global localization and mapping, one sets up the size of environment as 3 m × 3 m. In the beginning of pose initialization, four movable circle landmarks whose positions and color are known are given in front of the robot. Then, a circular trajectory with the radius of 1 m is set as the motion trajectory of the robot, as shown in Fig. 11. The velocities of the four wheels of the mobile base are set as follows;

$$\begin{bmatrix} V_{1w} \\ V_{2w} \\ V_{3w} \\ V_{4w} \end{bmatrix} = \begin{bmatrix} 0.1047 \\ 0.4189 \\ 0.1047 \\ 0.4189 \end{bmatrix} \quad m/\text{sec} \tag{35}$$

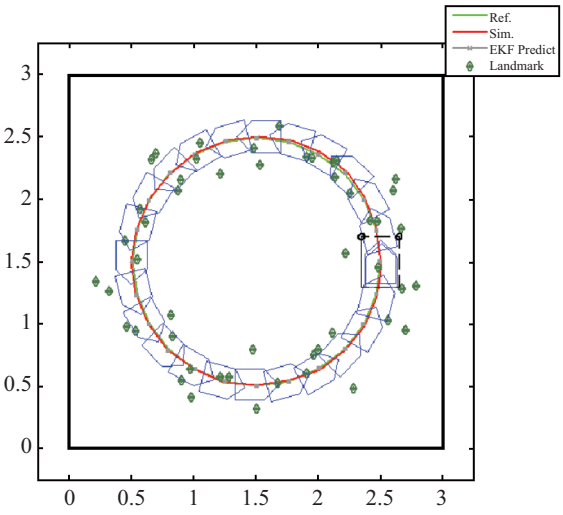

Fig. 11. Simulation results of the global simulation and mapping method.

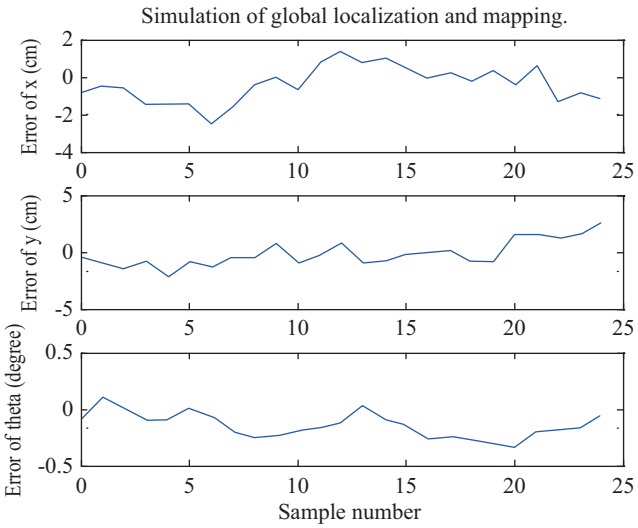

Fig. 12. The errors of pose tracking in simulation.

From the simulation results in Fig. 12, the average of the errors equals $[-0.38\text{ cm}\ \ -0.14\text{ cm}\ \ -0.1412°]^T$. As can be seen in Fig. 11, those landmarks are recognized and their calculated positions are almost the same to their true ones. Through the simulations, the proposed localization and mapping algorithm has been shown to find accurate locations of unknown natural landmarks and give accurate pose estimates for position and orientation of the ADAMR.

VIII. EXPERIMENT RESULTS AND DISCUSSION

This experiment is devoted to exploring the merit of the proposed pose initialization using the lease-squares method, and the second experiment focuses on the effectiveness of the proposed EKF-based tracking algorithms of the ADAMR. Worthy of mention is that the pose initialization and tracking algorithm are tested using only four circle landmarks with distinct colors.

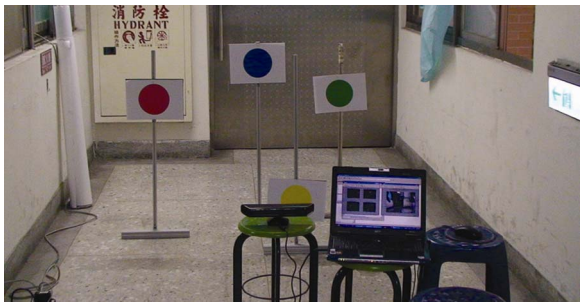

Fig. 13. Deployment of the four landmarks in the second experiment.

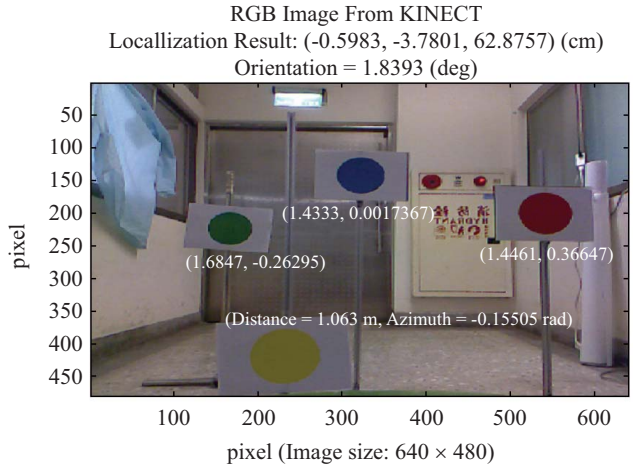

Fig. 14. Computed global position and orientation of the KINECT in the world frame using the four distances and azimuths of the four landmarks.

The initial pose of the KINECT got started at the posture, i.e. $[x_0\ \ y_0\ \ y_0\ \ \theta_0]^T = [0.0\text{ cm}\ \ 0.0\text{ cm}\ \ 58\text{ cm}\ \ 0\text{ deg.}]^T$ and the four landmarks were located at the corresponding four positions $[132.4\text{ cm}\ \ 57.5\text{ cm}\ \ 74.4\text{ cm}]^T$, $[102.3\text{ cm}\ \ -16.1\text{ cm}\ \ 27.8\text{ cm}]^T$, $[163.8\text{ cm}\ \ -47.6\text{ cm}\ \ 73.2\text{ cm}]^T$, and $[142.0\text{ cm}\ \ 0.0\text{ cm}\ \ 92.3\text{ cm}]^T$. Fig. 13 shows the deployment of the four landmarks in the experiment, and the four distances and azimuths of the four landmarks are depicted, thereby obtaining the global position and orientation of the KINECT in the world frame in Fig. 14, In comparison with the true distance and azimuth of the yellow landmark, the errors are $[1.6\text{ cm}\ \ -0.06015\text{ deg.}]^T$. Furthermore, in comparison with the true position and orientation of the KINECT, one can find the errors are $[-0.5983\text{ cm}\ \ -3.7801\text{ cm}\ \ 4.8757\text{ cm}\ \ 1.8393\text{ deg.}]^T$ in Fig. 15. Through the experiment results, the proposed pose initialization method has been shown capable of giving initial pose localization with acceptable accuracy for position and orientation of the ADAMR.

The experiment result reveals that the average of the errors equaled $[-0.75879\text{ cm}\ \ -0.29552\text{ cm}\ \ 0.89673°]^T$ in three axes. The above-mentioned results indicate that the proposed

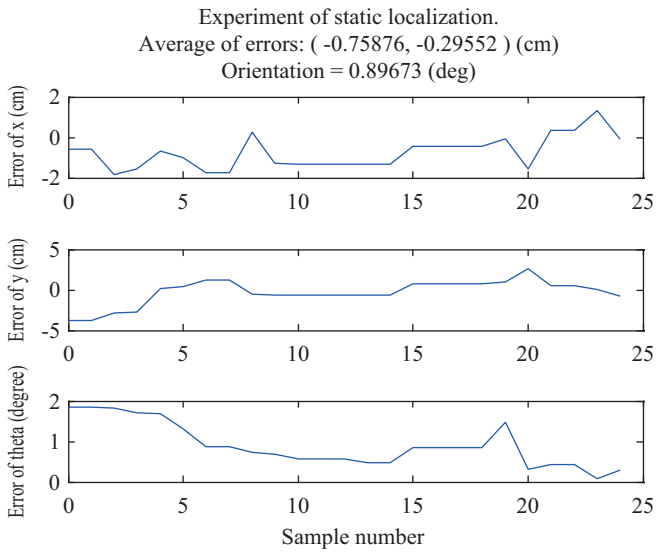

Fig. 15. Errors of pose tracking in experiment.

continuous pose tracking method is capable of giving accurate pose estimates for position and orientation of the ADAMR for the static case.

IX. CONCLUSIONS

This paper has developed a global localization approach for the ADAMR based on the least-square estimate method using KINECT, a KINECT-based detection algorithm of artificial landmarks, and an EKF-based localization algorithm by fusing the information from the odometric encoders and the KINECT sensor in indoor environments. Accumulation errors of the dead-reckoning method can be significantly reduced by fusing the KINECT information. By fusing the KINECT data and odometric measurements, the proposed EKF-based pose estimation method using the artificial landmarks has been shown capable of keeping track of the continuous poses of the robot navigating on a flat terrain at slow speeds. The simulation results for both pose initialization and continuous pose tracking have verified that the proposed methods provide more accurate pose estimates for the mobile ADAMR. The experimental results for static pose localization have confirmed the effectiveness and performance of the proposed global localization method. It is worthwhile to note that the proposed localization system can be installed at several distinct and pre-specified location within the robot's working space, in order to calibrate the robot's poses. Except the static experiment, the authors are conducting experiments on obtaining EKF-based pose estimation of the ADAMR while navigating around its working environment.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from National Science Council, the Republic of China, under contract NSC 100-2218-E-005-033-.

REFERENCES

1. Adachi, N., Fukao, T., and Nakagawa, H., "Adaptive tracking control of a nonholonomic mobile robot," *IEEE Transactions on Robotics Automation*, Vol. 16, No. 5, pp. 609-615 (2000).
2. Dellaert, F., Burgard, W., Fox, D., and Thrun, S., "Monte carlo localization for mobile robots," *Proceeding on IEEE International Conference on Robotics and Automation*, Vol. 2, pp. 1322-1328 (1999).
3. Ganganath, N. and Leung, H., "Mobile robot localization using odometry and kinect sensor," *IEEE International Conference on Emerging Signal Processing Applications (ESPA)*, pp. 91-94 (2012).
4. Graf, B., Hans, M., and Schraft, R. D., "Mobile robot assistants," *IEEE Robotics and Automation Magazine*, Vol. 11, No. 2, pp. 67-77 (2004).
5. Jensen, B., Froidevaux, G., Greppin, X., Lorotte, A., Mayor, L., Meisser, M., Ramel, G., and Siegwart, R., "The interactive autonomous mobile system Robox," *IEEE/RSJ International Conference on Intelligent Robots and System*, Vol. 2, pp. 1221-1227 (2002).
6. Kim, G., Chung, W., Han, S., Kim, K. R., Kim, M., and Shinn, R. H., "The autonomous tour-guide robot Jinny," *IEEE/RSJ International Conference on Intelligent Robots and System*, Vol. 4, pp. 3450-3455 (2004).
7. Kiriy, E. and Buehler, M., "Three-state extended Kalman filter for mobile robot localization," Technical Representative, McGill University, Montreal, Canada (2002).
8. Maksarov, D. and Durrant-Whyte, H., "Mobile vehicle navigation in unknown environments: a multiple hypothesis approach," *Proceeding on Institute Electric Engineering - Control Application Theory*, Vol. 142, No. 4, pp. 385-400 (1995).
9. Maybeck, P. S., *Stochastic Models, Estimation, and Control*, Vol. 1, Academic Press, New York (1979).
10. Siegwart, R. and Nourbakhsh, I. R., *Introduction to Autonomous Mobile Robots*, The MIT Press.
11. Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Hahnel, D., Rosenberg, C., Roy, N., Schulte, J., and Schulte, D., "MINERVA: a second-generation museum tour-guide robot," *IEEE International Conference on Robotics and Automation*, Vol. 3, pp. 1999-2005 (1999).
12. Triggs, B., "Model-based sonar localisation for mobile robots," *Robotics and Autonomous System*, Vol. 12, Nos. 3-4, pp. 173-186 (1994).
13. Tsai, C. C., Lin, H. S., and Hsu, J. C., "Ultrasonic localization and pose tracking of an autonomous mobile robot via fuzzy adaptive extended information filtering," *IEEE Transactions Instrument Measurements*, Vol. 57, No. 9, pp. 2024-2034 (2008).
14. Tsai, C. C. and Wang, Y. C., "Global localization using dead-reckoning and KINECT sensors for anthropomorphous two-armed robots with omnidirectional mecanum wheels," *Proceedings of 2012 National Symposium on Systems Science and Engineering*, Keelung, Taiwan (2012).