Proceedings of the 2014 IEEE
International Conference on Robotics and Biomimetics
December 5-10, 2014, Bali, Indonesia

Localization of Leader-Follower Formations Using Kinect and RTK-GPS

Ran Li, and Yunhua Li, *Senior Member, IEEE*

***Abstract*—The localization is an important task for intelligent mobile robot systems and the robot cooperation may bring benefits for the task in a multi-robot team. A localization method for leader-follower formations is proposed in this paper. In the proposed method, the leader robot adopts an Extended Kalman filter to fuses data from multiple sensors. Kinect with depth camera is a rich source of information for the relative localization of the follower mobile robot. In the paper, we focus our attention on Kinect sensor and discuss how to determine the relative localization of the follower robot using the frame information from Kinect sensor. The experiment results have demonstrated the effectiveness of the proposed method.**

I. INTRODUCTION

ROBOTIC systems are the interest research field in many applications [1]. A multi-robot team includes formation control tasks where a team of robots attempt to maintain a given formation, hierarchical control tasks where a leader robot controls another follower robot to move as required, and safety maneuvering tasks where awareness of neighboring team members is used for path planning purposes [2]. In a multi-robot team cooperative localization is an important research topic in mobile robotics [3] [4].

For localization, there are generally two methods: the relative positioning and the absolute positioning. The former calculates the position and heading angle using odometry or inertial sensors. The latter determines the position and direction using an external distance measuring system [6].

Relative localization use the initial starting point information to calculates the relative position. Odometry is the most widely used method for relative positioning. However, as the operating time and moving distance increasing, errors in the position and heading angle constantly increase [7]. Currently, many researchers have worked to decrease these errors by modeling and filter design. These studies were aimed to extend the period of navigation, without the help of external absolute position information. RTK-GPS, Real-Time Kinematic GPS, can give position up to centimeter accuracy. However, it can't work well for cases of less satellite condition, worse whether and larger radio signal noises [1].

A variety of relative measurement sensors which are applicable for ground are reported in literatures. In some field, researchers focused on the method using a vision device [8][9]. Some authors used a single perspective camera or binocular cameras [10] [11]. The main disadvantage of these approaches is the image geometric distortion, discretization error and corner point extraction in the process. Depth cameras are revolutionizing robot perception as a replacement for sensors such as stereo vision systems [12]. The Microsoft Kinect sensor provides a good quality of the depth information of the environment at an affordable price.

In this paper, the system contained two mobile robots as shown in Fig.1. The follower is equipped with a depth cameras (right) and the leader one is equipped with a RTK-GPS (5Hz) receiver. The RTK-GPS can achieve up to 2cm accuracy in a horizontal plane. Cameras are calibrated and images are rectified. GPS and image data are stored under the different computers communicate through wireless.

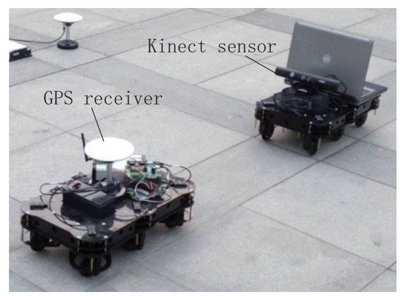


Fig. 1. Two robots, one is equipped with Kinect sensors (right) and the other is equipped with a GPS(left).

While the leader robot marches along the specified path, the follower robot controller controls the follower robot to move for a certain position. In this paper, we restrict our attention to use relative measurement for cooperative localization in a multi-robot team.

Relative localization within such a team allows the follower to determine the relative formation by utilizing the relative measurements. Kinect sensor consists of an infrared laser projector combined with a monochrome CMOS camera, and a second RGB video camera. It can return information in the frame of reference of the leader robot.

The rest of this paper is structured as follows. The next Section II builds related work on kinematics model of the single mobile robot. Section III contains the methods for location of the leader based on Extended Kalman filter. In Section IV we restrict our attention to Kinect sensors which return information in the frame of reference of the leader robot. Section V contains an experimental study. Finally, Section VI presents our conclusions and a brief discussion of future work.

Manuscript received August 12, 2014. This work was supported by National Science Foundation of China (Grant No.51075016).

Ran Li and Yunhua Li are with school of Automation Science and Electrical Engineering, Beihang University, Beijing, China. (e-mail: liran91000@163.com; email: yhli@buaa.edu.cn).

978-1-4799-7397-2/14/$31.00 © 2014 IEEE

II. Robot Model

A. Coordinated System

Fig. 2 shows the global coordinate system ($O_W X_W Y_W$). For the whole system, it is convenient to define body coordinate systems $O_m X_m Y_m$ .It is attached to the mobile robot with origin in its center of mass (COM).

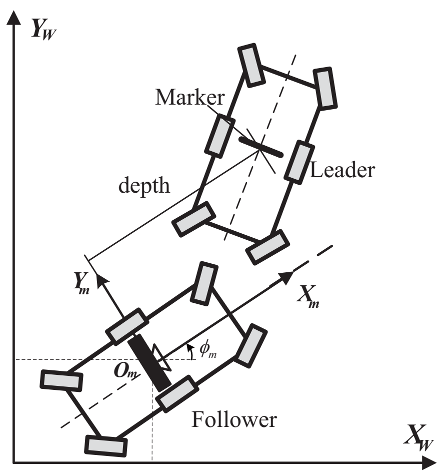


Fig. 2. Coordinate estimate via robot tracker: observation of the reference robot by the follower robot

B. Kinematics Model of Mobile Robot

The leader robot and the follower robot share a same kinematics model.

Consider a robot moving on two dimensional plane with global coordinate system ($O_W X_W Y_W$) as depicted in Figure 3. (x, y) is the coordinates of its center of mass in the global coordinate system. ϕ_L is an orientation the local coordinate system with respect to the global coordinate system. Let $\boldsymbol{u}_L = [v_{Lm} \ w_L]^T \in R^2$ be an input, and v_{Lm} is the speed at the centre along X_m axis, ω_L the angular velocity of the robot. From Fig.3, the steering radius R_{Lm} , the front-left wheel steering angle θ_{Lfl} and the front-right wheel steering angle θ_{Lfr} can be derived as follows

$$R_{Lm} = v_{Lm} / \omega_L \ . \tag{1}$$

$$\tan\theta_{Lfl} = L^{-1}(R_{Lm} - 0.5\mathrm{B}), \cot\theta_{Lfr} - \cot\theta_{Lfl} = B / L \tag{2}$$

where L is the distance between each end and the midpoint, and B is the distance between left side wheel and right side wheel. Similarly, the angle and speed of other wheels can be also obtained.

As seen from (1), to adopt $[v_{Lm} \ \ \theta_{Lfl}]^T$ as the control input and to adopt $[v_{Lm} \ \ \omega_L]^T$ are equivalent.

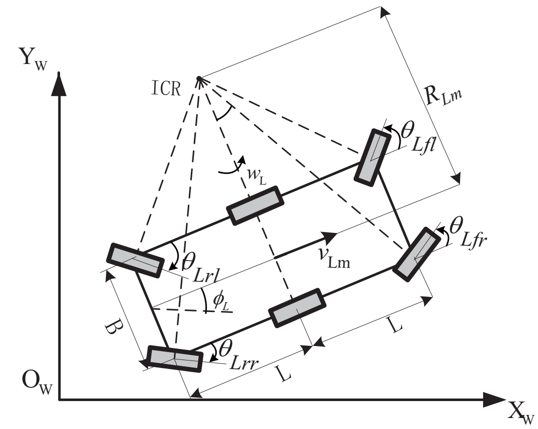


Fig. 3. Local coordinate system

From Fig.3, it is easy to derive kinematic equation of motion of the wheel robot as follow

$$\begin{cases} \dot{x} = v_{Lm} \cos\phi_L \\ \dot{y} = v_{Lm} \sin\phi_L \\ \dot{\phi}_L = \omega_L \end{cases} \tag{3}$$

III. Positioning of Leader Mobile Robot

A. State equation

Rotor encoder can convert the angular rate of a rotor into a digital signal. The vehicle is equipped with encoders on the wheels. Estimation on velocity v and angular velocity w is provided at each sampling period ΔT .

In order to estimate the position and orientation of the object, a discrete time state space dynamic model has to be considered as follows

$$\begin{cases} x(k+1) = x(k) + \Delta T(v_{Lm}(k) + \nu(k))\cos\phi_L(k) \\ y(k+1) = y(k) + \Delta T(v_{Lm}(k) + \nu(k))\sin\phi_L(k) \\ \phi_L(k+1) = \phi_L(k) + \Delta T(\omega_L(k) + w(k)) \end{cases} \tag{4}$$

where $\boldsymbol{q} = [x \ y \ \phi]^T \in R^3$ is state vector expressed as the mobile robot pose. To Express (4) in terms nonlinear state space yields:

$$\boldsymbol{q}(k+1) = f(\boldsymbol{q}(k), \boldsymbol{u}(k), \boldsymbol{\nu}(k)) \tag{5}$$

where $f(\bullet)$ represents the state transition function. $\boldsymbol{\nu}(k) = [\nu(k), w(k)]^T$ represents unpredictable process noise, that is assumed to be Gaussian with zero, and covariance $\boldsymbol{Q}(k)$. The process noise covariance $\boldsymbol{Q}(k)$ was modeled on the assumption of two independent sources of error, velocity and angular, $v(k)$ and $w(k)$ are added with corresponding uncertainties. The expression for $\boldsymbol{Q}(k)$ is as

$$Q(k)=\begin{bmatrix}\sigma_v^2 & 0\\ 0 & \sigma_w^2\end{bmatrix} \tag{6}$$

where $\sigma_v{}^2$ and $\sigma_w{}^2$ denotes respectively variances of $v(k)$ and $w(k)$.

B. Measurement equation based on RTK-GPS

RTK-GPS, mean Real time kinematic GPS, is a method that can be able to get cm-level real-time positioning accuracy. The use of the dynamic real-time carrier phase measurement is a major milestone in GPS application.RTK technology is widely used in engineering survey, maritime precision positioning, topographic map, cadastral survey because of high precision, wide application.

The GPS observation from RTK-GPS is used as position measurement. As in the urban environment, GPS suffers from multi-path problems, and the non stationary noise of GPS measurement noise affects the observation model, the linear observation equation of GPS positions is

$$\boldsymbol{Y}(k)=\begin{bmatrix}x^{gps}(k)\\ y^{gps}(k)\end{bmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}x(k)\\ y(k)\\ \theta(k)\end{pmatrix} \tag{7}$$

while the GPS observation ($x^{gps}(k), y^{gps}(k)$) is provided by GPS position measurement

In our system, we choose the observation from RTK-GPS as position measurement. The measurement function $\boldsymbol{z}(k)=h(\boldsymbol{q}(k), \boldsymbol{w}(k))$ is

$$\boldsymbol{z}(k)=[x(k), y(k)]^{\mathrm{T}} \tag{8}$$

where $\boldsymbol{w}(k)$ is the measurement noise (Gaussian with zero mean and variance $r(k)$. Measurement covariance matrix $\boldsymbol{R}(k)$ is a diagonal matrix with the elements $r(k)$ .As GPS measurements are affected by many independent noise sources, the measurement noise of GPS position error can be estimated by

$$\boldsymbol{R}(k)=\begin{bmatrix}\sigma_x{}^2 & 0\\ 0 & \sigma_y{}^2\end{bmatrix} \tag{9}$$

where $\sigma_w{}^2$ and $\sigma_\theta{}^2$ denotes respectively variances of $x(k)$ and $y(k)$. The Jacobian matrix of $h(\bullet)$ with respect to $\boldsymbol{q}$ is

$$\boldsymbol{H}(k)=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}. \tag{10}$$

C. Extended Kalman Filtering

1) Prior estimation of the state

First, we need to use system process model to predict the next state of the system. The state transition function $f(\bullet)$ uses the state vector at the current time instant and the current control input to compute the state vector at the next time step

$$\hat{\boldsymbol{q}}^{-}(k+1)=f(\hat{\boldsymbol{q}}(k), \boldsymbol{u}(k), \boldsymbol{v}(k)) \ . \tag{11}$$

2) Correction of the estimation of the state

Then, we find a priori estimate $\boldsymbol{P}^{-}(k+1)$ of the error covariance matrix $\boldsymbol{P}^{+}(k)$

$$\boldsymbol{P}^{-}(k+1)=\boldsymbol{A}(k)\boldsymbol{P}^{+}(k)\boldsymbol{A}^{T}(k)+\Delta T^2\boldsymbol{W}(k)\boldsymbol{Q}(k)\boldsymbol{W}^{T}(k) \tag{12}$$

where $\boldsymbol{A}(k)=\frac{\partial f(\boldsymbol{q})}{\partial \boldsymbol{q}}\Big|_{\boldsymbol{q}=\hat{q}^{-}(k+1)}$ is the Jacobian matrix of $f(\bullet)$ with respect to $\boldsymbol{x}$,while $\boldsymbol{W}=\frac{\partial f(\boldsymbol{q})}{\partial \boldsymbol{u}}\Big|_{\boldsymbol{q}=\hat{q}^{-}(k+1)}$ is the Jacobian matrix of $f(\bullet)$ with respect to control input $\boldsymbol{u}$. In the real application, we generally choose $\boldsymbol{P}^{+}(0)=\boldsymbol{WQW}^{T}$.

Then we can get the posteriori state estimate in the correction step using

$$\hat{\boldsymbol{q}}(k+1)=\hat{\boldsymbol{q}}^{-}(k+1)+\boldsymbol{K}(k)(\boldsymbol{Y}(k)-h(\hat{\boldsymbol{q}}^{-}(k+1)) \tag{13}$$

where the Kalman gain $\boldsymbol{K}(k)$ is calculated by

$$\boldsymbol{K}(k)=\boldsymbol{P}^{-}(k+1)\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{HP}^{-}(k+1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}(k)^{\mathrm{T}})^{-1} \ . \tag{14}$$

3) Motion Update

Finally, we compute the posteriori estimate of the error covariance matrix

$$\boldsymbol{P}^{+}(k+1)=(\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{H})\boldsymbol{P}^{-}(k+1), \tag{15}$$

and $\boldsymbol{P}^{+}(k+1)$ is chosen as the error covariance matrix at the next time step.

IV. Positioning of Follower Mobile Robot

In this part, we introduce the principle of Kinect-based position measurement, and then introduce how to track the fixed point in the leader by using the Camshift tracking algorithm.

The experimental results show that the method can tracking the target object real-time.

A. Principle of Kinect-based Position Measurement

According to paper [14] and [15], the sensor is capable of operation in dynamic environments. The Kinect sensor does not contain large systematic errors when compared with a laser scanning data, the density of points also decreases with increasing distance to the sensor. Its low cost, high frame rate and absolute depth accuracy over a useful range make it suitable for use on robotic platforms.

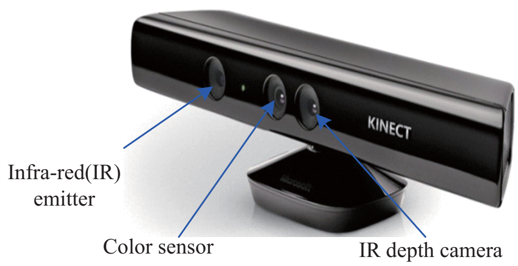


Fig. 4. Kinect sensor

Fig. 4 is the Microsoft Kinect. It is a low cost peripheral human machine for use as a game controller with the Xbox 360 game system. The Prime sense chipset in the Kinect uses a form of structured light, a proprietary Light Coding technique to compute depth. The Kinect sensor consists of an infrared laser projector combined with a monochrome CMOS camera, and a second RGB video camera. Both cameras provide 640 x 480 pixel images at 30 Hz [13].

The infra-red camera is a depth camera, and it can capture images of the object with different color in the image block represent the distance.

Unlike the RGB video camera, the resulting data from the infra-red camera is a two-dimensional matrix contains the depth information. The depth information is depth image frames in depth pixel. Each depth pixel consists of 2 bytes (16 bits, high 13 bits represent the distance between the camera plane and a certain point; and low 3 bits represents the index number. The distance in mm can be obtained by extracting the high 13 bit depth values in depth pixels.

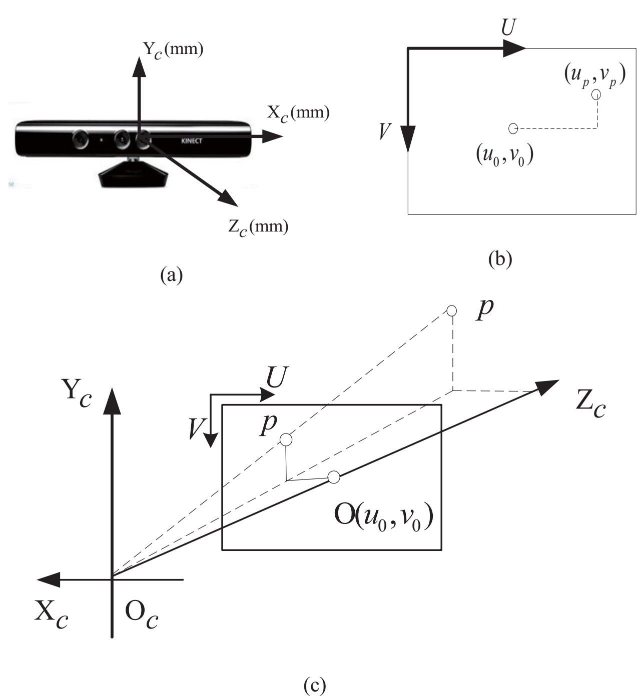


Fig. 5. Camera model, the image plane relationship diagra. (a) Camera model. (b) The image plane. (c) The relationship .

By indexing the pixel coordinates of the corresponding points, we are able to get the information, namely the distance from the target point to the camera plane.

Fig.5 shows relationship between camera model and the image plane. $(X_C Y_C Z_C)$ is the camera coordinate system. Fig.5 (a) shows the camera model. Origin of the coordinates is the center of infra-red (IR) camera. When we use depth image and color image at the same time, it is necessary to map the color image to the depth image, namely the image fusion. Fig.5 (b) shows the image plane.

From Fig. 5, it is easy to be derived the geometric relationship as follows

$$x_p = d_p,\ \frac{y_p}{d_p} = \frac{u_p - u_0}{f_u C_u},\ \frac{z_p}{d_p} = \frac{v_p - v_0}{f_v C_v} \tag{16}$$

where d_p represents the distance from the point to the camera plane. (u_p, v_p) and (u_0, v_0) represent the pixel coordinates of the point and the the pixel coordinates of the infrared camera in the image. (f_u, f_v) represents the focal distance vector of the infrared camera. C_u and C_v are fixed parameters of the camera, represent the length of the unit pixel, in pixel/mm.

By using GetDepthMap() function in OpenNI, depth image pixel values in mm can be obtained. And by using ConvertProjectToRealWorld() function we can transform the point of a projected coordinate system to the real world coordinate system and get the three-dimensional $[x_c, y_c, z_c]^{\mathrm{T}}$ in the camera coordinate system.

Then the three-dimensional coordinates in the body coordinate system is

$$[x_m, y_m, z_m]^{\mathrm{T}} = \boldsymbol{T}[x_c, y_c, z_c]^{\mathrm{T}}. \tag{17}$$

We can get the rotation matrix $\boldsymbol{T}$ through the correspondence between the two coordinate systems (see Fig.2 and see Fig.5 (a)).

$$\boldsymbol{T} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{18}$$

where $[x_m, y_m, z_m]^{\mathrm{T}}$ is the relative position between the leader and the follower, presents the three-dimensional of the leader in the body coordinate systems of the follower.

B. Camshift Tracking Frameworks

In this part, we track the leader by using CamShift tracing algorithm. The process of Camshift can be described as Fig.6.

Camshift uses the color feature to track the moving object. It uses cluster to find the target. The position and size of the moving object can be obtained. The search box of the next frame can be obtained according to size and position of the current frame. While the process is continuously done, the object can be continuously tracked. By using the size and position of the current window, the initial value of the current search box is set before the search begins. Because the moving object is searched around the possible position, much time is saved. Camshift has good robustness and real time [16].

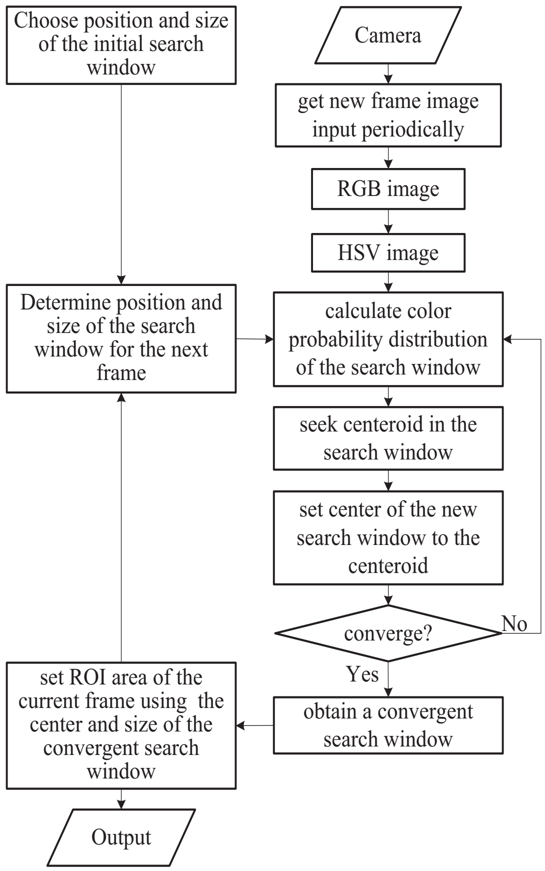


Fig. 6. The process of Camshift.

B. Formation Control Based on PID Control

In order to guarantee the stability of the formation, the robot's motion control is with position feedback control. By real-time detection of the leader robot, the relative position between the leader robot and the follower one can be obtained. Then get the error by comparing with the preset distance. By the error correction of the actuator of the follower robot, the closed-loop tracking control of the follower robot can be achieved.

As described before, we can get the relative position between the leader and the follower by using a Kinect sensor. While the leader robot marches along the specified path, by using the relative position, the tracking purposes can be achieved.

The input of the controller is the error of the robot relative distance $(\Delta x, \Delta y)$, while the output is velocity v and angular velocity w.

Assuming the preset distance is (x_d, y_d), represent the expected relative position, we can get a closed-loop control block diagram as shown in Fig. 7.

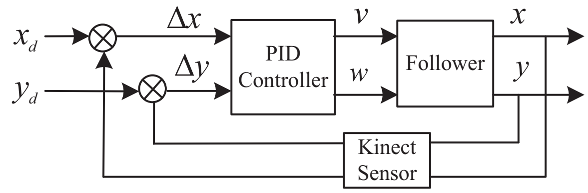


Fig.7. The block diagram of system control loop.

Obviously, the system is of multi-input and multi-output. However, because there is no tight coupling among the input and output components, it is able to be solved respectively.

V. Experimental Platform

A. Experimental Hardware

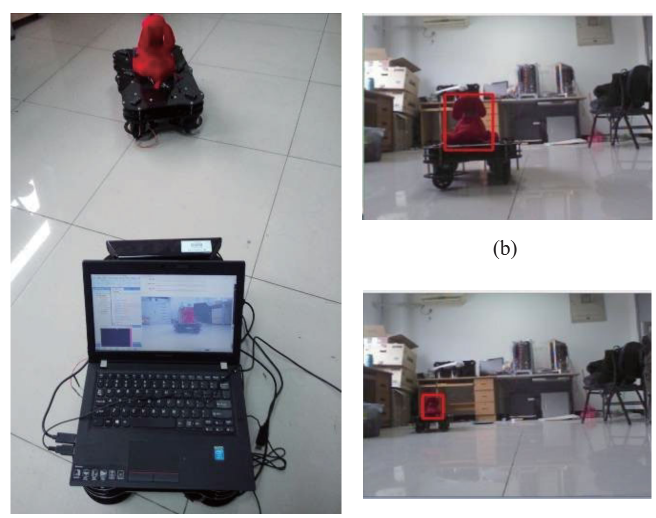
The robot features a real time embedded controller .A velocity encoder unit is also present, and contains a steering gear copper server in each wheel.

A control system was developed to evaluate the suitability of the Kinect sensor in dynamic environments by using the calibrated output from the depth camera. The visual distance controller runs on a laptop. The Kinect is mounted on the center of the robot and connected to the laptop via USB .

B. Tracking Experiment

In this paper, we have an experimental verification by using the VS2010 development platform, OpenCV (Open Source Computer Vision Library) and OpenNI (Open Natural Interaction), the tracking target is the leader robot, the moving range the target is about 300 mm to 2400 mm.

Fig.8 shows the tracing experiment. While the leader robot marches along the path, the follower one uses a Kinect sensor for real-time tracking.



(a) (b) (c)

Fig. 8. Tracing experiment. (a) The relative position between the leader and the follower. (b) (c) Real-time target tracking.

The three-dimensional coordinates of the center of mass of reference vehicle could be real-time output in the experiment the measured of the distance of the experimental data and actual measurement result were compared in Table 1.

TABLE I

RESULTS CONTRAST

Frame number	3d coordinate(*mm*) (x_m, y_m , z_m)	Actual depth (*mm*)
1	(292.814,-419.209,2426)	400
2	(61.353,-89.1535,552)	500
3	(297.009,-357.648,637)	650
4	(103.392,-124.107,810)	800
5	(35.1206,-180.62,983)	1000
6	(16.5331,-194.507,1180)	1200
7	(4.00914,-295.324,1504)	1500
8	(132.884,-375.74,1779)	1800
9	(11.5732,-385.783,1915)	1920
10	(94.5615,-436.702,2080)	2100
11	(175.369,-509.601,2356)	2400

The successful control of getting 3d coordinates by using the Kinect sensor demonstrates that the sensor is capable of operation in dynamic environments. Its high frame rate, low cost and absolute depth accuracy over a useful range make it suitable for use on a robotic platform.

VI. CONCLUSION

In this paper, we used multiple sensors fusion system to present a method to estimate localization for leader-follower formations. It can be applied into formation control for multiple mobile robots in the future.

Odometry has accumulated heading angle and position errors. As the operating time and moving distance increasing, errors constantly increased. Therefore, accurate RTK-GPS positions are used as the measurements to produce more accurate vehicle position using Extended Kalman Filter.

The paper also presented a theoretical and experimental analysis of depth data acquired by the Kinect sensor. From the results t the following main conclusions can be drawn:

An accurate calibration of the RGB camera and the IR camera is necessary to eliminate misalignments between the color and depth data;

With distance increasing, the depth resolution decreases seriously and the random error of depth data increases;

The depth data should be acquired within 0.5–3 m distance to the sensor. At smaller or larger distances, the quality of the data is degraded seriously because of the noise and low resolution of the depth measurements.

Further work will involve integrating the Kinect sensor into the navigation layer of the system. The depth and RGB images combined in the manner described in Section V-B will require more robust methods for this purpose.

VII. REFERENCES

[1] Zhuang Y, Gu M W, Wang W, et al. "Multi-robot cooperative localization based on autonomous motion state estimation and laser data interaction", Science China Information Sciences, 2010, 53(11): 2240-2250.

[2] De Silva O, Mann G K I, Gosine R G. "Pairwise observable relative localization in ground aerial multi-robot networks", Control Conference (ECC), 2014 European. IEEE, 2014: 324-329.

[3] Ahmad A, Tipaldi G D, Lima P, et al. "Cooperative robot localization and target tracking based on least squares minimization", Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013: 5696-5701.

[4] V.DungHoang,M.Hale, K.Hyun."Localization estimation based on Extended Kalman filter using multiple sensors", Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE, pp. 5498–5503, Nov. 2013.

[5] Rekleitis I M, Dudek G, Milios E E. "Multi-robot cooperative localization: a study of trade-offs between efficiency and accuracy", Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on. IEEE, 2002, 3: 2690-2695.

[6] M. Spangenberg, V. Calmettes and J.-Y. Tourneret, "Fusion of GPS ,INS and Odomtric Data for Automotive Navigation", 15th European Signal Processing Conference (EUSIPCO 2007), Poznan, Poland, September 3-7, 2007.

[7] L. Junwei, C. Cappelle, Y. RUICHEK, "Intelligent Vehicle Localization in Urban Enviroments Using EKF-based Visual Odometry and GPS Fusion", Preprints of the 18th IFAC World Congress, Milano (Italy) August 28 ,Sepetmber2,2011.

[8] F. Fraundorfer and D. Scaramuzza, "Visual Odometry:Part II: Matching, Robustness, Optimization, Applications", IEEE Robotics&Automation Magazine, vol. 19, pp. 78-90, 2012.

[9] Mariottini G L, Pappas G, Prattichizzo D, et al. "Vision-based localization of leader-follower formations" [C]//Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on. IEEE, 2005: 635-640.

[10] B.Su Cho, W.Jin Seo, W.sung Moon, and K.Ryul Baek, "Positioning of a Mobile Robot Based on Odometry and a New Ultrasonic LPS", International Journal of Control, Automation, and Systems, pp.333-345, 2013.

[11] Chen C, Chai W, Zhang Y, et al. "A RGB and D vision aided multi-sensor system for indoor mobile robot and pedestrian seamless navigation",Position, Location and Navigation Symposium-PLANS 2014, 2014 IEEE/ION. IEEE, 2014: 1020-1025.

[12] E. Royer, M. Lhuillier, M. Dhome, and J.-M. Lavest, "MonocularVision for Mobile Robot Localization and Autonomous Navigation", International Journal of Computer Vision, vol. 74, pp. 237-260, 2007.

[13] Cunha J, Pedrosa E, Cruz C, et al. "Using a depth camera for indoor robot localization and navigation", DETI/IEETA-University of Aveiro, Portugal, 2011.

[14] Stowers J, Hayes M, Bainbridge-Smith A. "Altitude control of a quadrotor helicopter using depth map from Microsoft Kinect sensor", Mechatronics (ICM), 2011 IEEE International Conference on. IEEE, 2011: 358-362.

[15] Khoshelham K, Elberink S O. "Accuracy and resolution of kinect depth data for indoor mapping applications", Sensors, 2012, 12(2): 1437-1454.

[16] Allen, John G., Richard YD Xu, and Jesse S. Jin. "Object tracking using camshift algorithm and multiple quantized feature spaces." Proceedings of the Pan-Sydney area workshop on Visual information processing. Australian Computer Society, 2004: 3-7.