

SELF-LOCALIZATION METHOD BY INTEGRATING SENSORS

T. Fuse^{a,*}, K. Matsumoto^b

^a Dept. of Civil Engineering, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656, Japan -
fuse@civil.t.u-tokyo.ac.jp

^b Rakuten, Inc., Higashi Shinagawa 4-12-3, Shinagawa-ku, Tokyo, 140-0002, Japan

Commission V, WG V/4

KEY WORDS: Sensor Fusion, Self-Localization, Mobile Devices, Augmented Reality, Visualization, Navigation

ABSTRACT:

Recently, development of high performance CPU, cameras and other sensors on mobile devices have been used for wide variety of applications. Most of the applications require self-localization of the mobile device. Since the self-localization is based on GPS, gyro sensor, acceleration meter and magnetic field sensor (called as POS) of low accuracy, the applications are limited. On the other hand, self-localization method using images have been developed, and the accuracy of the method is increasing. This paper develops the self-localization method by integrating sensors, such as POS and cameras, on mobile devices simultaneously. The proposed method mainly consists of two parts: one is the accuracy improvement of POS data filtering, and another is development of self-localization method by integrating POS and camera. The POS data filtering combines all POS data by using Kalman filter in order to improve the accuracy of exterior orientation factors. The exterior orientation factors with POS filtering are used as initial value of ones in image-based self-localization method. The image-based self-localization method consists of feature points extraction and tracking, relative orientation, coordinates estimation of the feature points, and orientation factors updates of the mobile device. The proposed method is applied to POS data and images taken in urban area. Through experiments with real data, the accuracy improvement by POS data filtering is confirmed. The proposed self-localization method with POS and camera make the accuracy more sophisticated by comparing with only POS data filtering.

1. INTRODUCTION

Recently, development of high performance CPU, cameras and other sensors on mobile devices have been used for wide variety of applications. One of the most popular applications is augmented reality (AR). The AR technique superimposes various data on the scene which the users see at the time instead of comprehensive and detailed three dimensional modelling such as virtual reality. The AR uses sequential images taken from same view points of users as environmental scene, and then reality of visualization increase compared with the virtual reality. AR requires self-localization, namely orientation of the mobile device.

Since the self-localization of the mobile device is based on GPS, gyro sensor, acceleration meter and magnetic field sensor (called as position and orientation system (POS) in this paper) of low accuracy, the applications are limited to tags superimposition on the sequential images.

On the other hand, self-localization method using images have been developed, and the accuracy of the method is increasing. The method, however, requires initial values of the orientation factors and distance of baseline. GCPs can be used to specify the absolute coordinates. But it takes a lot of time and effort, and the applicability is restrictive.

This paper develops the self-localization method by integrating sensors, such as POS and cameras, on mobile devices simultaneously. The proposed method mainly consists of two parts: one is the accuracy improvement of POS data filtering, and another is development of self-localization method by integrating POS data and images.

2. POS DATA FILTERING

2.1 Overview

In this study, an iPhone4s is used as a mobile device. The centre of the mobile device is defined as the origin of the coordinate system for the mobile device. In the coordinate system, z axis is set as the vertical direction of device monitor, and y axis is norm direction of the monitor plane, and x axis is defined according to right-handed coordinate system. The rotations with regard to the axis x , y , z (roll, pitch, yaw) are expressed as ω_g , ϕ_g , κ_g , respectively. For absolute coordinate system, origin is same as the device coordinate system. Z axis is set as the direction of gravitational force, and X axis is north-seeking. When the both coordinate systems coincide, the rotation angles are define as 0, namely $\omega_g=0$, $\phi_g=0$, $\kappa_g=0$.

As mentioned before, POS includes GPS, gyro sensor, acceleration meter and magnetic field sensor. The corresponding data are three dimensional coordinates of the sensor (longitude: E (degree), latitude: N (degree), ellipsoidal height: h (m)), angular velocity ($\dot{\omega}_g$, $\dot{\phi}_g$, $\dot{\kappa}_g$ (rad/s)), acceleration (\ddot{x}_a , \ddot{y}_a , \ddot{z}_a (G)) and azimuth angle (κ_i (rad)), respectively.

Normally, each sensor is used separately (Daehne and Karigiannis, 2002), or some sensors are selected among them (Agrawal and Konolige, 2006; Kourogi and Kurata, 2005; Wagner *et al.*, 2009). The proposed method combines all sensors data in order to improve the accuracy of exterior orientation factors. Figure 1 shows the pipeline of the POS data filtering.

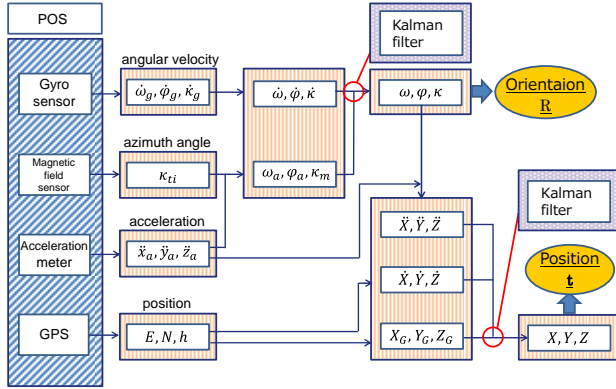


Figure 1. Flow of the POS data filtering

2.2 Filtering method of rotation

Initial rotation of the mobile device is calculated from acceleration meter and magnetic field sensor, and rotation during movement is calculated from the gyro sensor. The data of gyro sensor have drift error. In order to correct the drift error, filtering method with the data of the acceleration meter and the magnetic field sensor is applied. In this step, the data of such sensors are combined by using Kalman filter (Schall *et al.*, 2009). The assumption, that the mobile sensor moves linearly, is accepted here.

2.2.1 Calculation of rotation: When the mobile device remains stationary, the acceleration meter detect only acceleration of gravity g . The relationship between the acceleration and initial rotation (φ_a , θ_a , κ_m) can be expressed as follows.

$$\begin{bmatrix} \ddot{x}_a \\ \ddot{y}_a \\ \ddot{z}_a \end{bmatrix} = \begin{bmatrix} -g \sin \varphi_a \\ g \cos \varphi_a \sin \omega_a \\ g \cos \varphi_a \cos \omega_a \end{bmatrix} \quad (1)$$

From the equation (1), roll and pitch can be calculated.

$$\omega_a = \tan^{-1} \frac{\ddot{y}_a}{\ddot{z}_a}, \quad \varphi_a = \tan^{-1} \frac{-\ddot{x}_a}{\sqrt{\ddot{y}_a^2 + \ddot{z}_a^2}} \quad (2)$$

Yaw should be transformed according to the above defined coordinate system:

$$\kappa_m = \pi - \kappa_{ti} \quad (3)$$

For calculation of rotation during movement, data of the gyro sensor are added to the initial rotation. Derivatives of the rotation with respect to time ($\dot{\omega}$, $\dot{\varphi}$, $\dot{\kappa}$) are represented by using angular velocity ($\dot{\omega}_g$, $\dot{\varphi}_g$, $\dot{\kappa}_g$).

$$\begin{bmatrix} \dot{\omega} \\ \dot{\varphi} \\ \dot{\kappa} \end{bmatrix} = \frac{1}{\cos \varphi} \begin{bmatrix} \cos \varphi & \sin \omega \sin \varphi & \cos \omega \sin \varphi \\ 0 & \cos \omega \cos \varphi & -\sin \omega \cos \varphi \\ 0 & \sin \omega & \cos \omega \end{bmatrix} \begin{bmatrix} \dot{\kappa}_g \\ \dot{\varphi}_g \\ \dot{\omega}_g \end{bmatrix} = B \begin{bmatrix} \dot{\kappa}_g \\ \dot{\varphi}_g \\ \dot{\omega}_g \end{bmatrix} \quad (4)$$

By integrating equation (4), the rotation during movement can be acquired.

$$\begin{bmatrix} \omega \\ \varphi \\ \kappa \end{bmatrix} = \int \begin{bmatrix} \dot{\omega} \\ \dot{\varphi} \\ \dot{\kappa} \end{bmatrix} dt + \begin{bmatrix} \omega_a \\ \varphi_a \\ \kappa_m \end{bmatrix} \quad (5)$$

2.2.2 Filtering of rotation: According to integration, the error is accumulated. It is important to point out here that the acceleration meter and the magnetic field sensor are free from such cumulative error. With the acceleration meter and the magnetic field sensor, the rotation during movement is corrected. Filtering theory is applied in order to correct the cumulative error.

For the efficient computation, Kalman filter is adopted as filtering method (Ristic *et al.*, 2004). The Kalman filter consists of dynamic equation and observation equation as follows:

$$\mathbf{x}_t = F_t \mathbf{x}_{t-1} + G_t \mathbf{v}_t \quad (6)$$

$$\mathbf{y}_t = H_t \mathbf{x}_t + \mathbf{w}_t \quad (7)$$

where \mathbf{x}_t = state vector at time t
 \mathbf{y}_t = observation vector at time t
 F_t , H_t , G_t = matrices of transition, system noise, observation at time t , respectively
 \mathbf{v}_t = system noise vector at time $t \sim N(0, Q_t)$
 \mathbf{w}_t = observation noise vector at time $t \sim N(0, R_t)$.

In this study, these vectors and matrices are constructed as follows:

$$\mathbf{x}_t = \begin{bmatrix} \omega \\ \varphi \\ \kappa \\ \dot{\kappa}_g \\ \dot{\varphi}_g \\ \dot{\omega}_g \end{bmatrix}, \quad \mathbf{y}_t = \begin{bmatrix} \omega_a \\ \varphi_a \\ \kappa_m \\ \dot{\kappa}_g \\ \dot{\varphi}_g \\ \dot{\omega}_g \end{bmatrix}, \quad F_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

The B_t is same as equation (4). Matrices H_t and G_t are set as identity matrices. In equation (8), suffix a , m , g express data of the acceleration meter, the magnetic field sensor, gyro sensor, respectively.

Since the Kalman filter is based on linear equation with normal distribution, the equations can be solved analytically. The equations (9) and (10) are related to forecast, and (11)–(13) filtering.

$$\mathbf{x}_{t|t-1} = F_t \mathbf{x}_{t-1|t-1} \quad (9)$$

$$\mathbf{V}_{t|t-1} = F_t \mathbf{V}_{t-1|t-1} F_t^T + G_t Q_t G_t^T \quad (10)$$

$$\mathbf{K}_t = \mathbf{V}_{t-1|t-1} H_t^T (H_t \mathbf{V}_{t-1|t-1} H_t^T + R_t)^{-1} \quad (11)$$

$$\mathbf{x}_{t|t} = \mathbf{x}_{t|t-1} + \mathbf{K}_t (\mathbf{y}_t - H_t \mathbf{x}_{t|t-1}) \quad (12)$$

$$\mathbf{V}_{t|t} = \mathbf{V}_{t|t-1} - \mathbf{K}_t H_t \mathbf{V}_{t|t-1} \quad (13)$$

State vector $\mathbf{x}_{t|t}$ is a result of filtering method for rotation.

2.3 Filtering method of position

Although the position of the mobile device is measured with GPS directly, the accuracy of the position can be improved by relating to the other sensors used for rotation calculation. Here the Kalman filter is also applied to all POS data to estimate the position.

2.3.1 Calculation of position: Since GPS data are originally longitude E , latitude N , ellipsoidal height h , the coordinates are transformed to the Plane Rectangular Coordinate System (X_G , Y_G , Z_G). The coordinates are used for initial position of the mobile device. For measurement of position during movement, data of GPS and the acceleration meter are considered.

The relationship between the measured acceleration and second derivatives of position (\ddot{X} , \ddot{Y} , \ddot{Z}) can be expressed as follows.

$$\begin{bmatrix} \ddot{X}_t \\ \ddot{Y}_t \\ \ddot{Z}_t \end{bmatrix} = - \left(R_k R_\phi R_\omega \begin{bmatrix} \ddot{x}_a \\ \ddot{y}_a \\ \ddot{z}_a \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \right) \quad (14)$$

Here, rotation matrices are represented as R . Additionally, velocity (\dot{X}_t , \dot{Y}_t , \dot{Z}_t) can be calculated from GPS data.

$$\begin{bmatrix} \dot{X}_t \\ \dot{Y}_t \\ \dot{Z}_t \end{bmatrix} = \frac{1}{\Delta t} \begin{bmatrix} X_{t-1} - X_{t-2} \\ Y_{t-1} - Y_{t-2} \\ Z_{t-1} - Z_{t-2} \end{bmatrix} \quad (15)$$

Where, Δt is GPS data acquisition interval. By using both of velocity and acceleration, position at time t is calculated.

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} X_{t-1} \\ Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \Delta t \begin{bmatrix} \dot{X}_{t-1} \\ \dot{Y}_{t-1} \\ \dot{Z}_{t-1} \end{bmatrix} + \frac{1}{2} \Delta t^2 \begin{bmatrix} \ddot{X}_{t-1} \\ \ddot{Y}_{t-1} \\ \ddot{Z}_{t-1} \end{bmatrix} \quad (16)$$

2.3.2 Filtering of position: In the similar fashion of filtering method of rotation, Kalman filter is applied for improving position. Dynamic and observation equations are same as equations (6), (7). State vector, observation vector and matrix F_t are set for position filtering.

$$\begin{aligned} \mathbf{x}_t &= [X_t \ Y_t \ Z_t \ \dot{X}_t \ \dot{Y}_t \ \dot{Z}_t \ \ddot{X}_t \ \ddot{Y}_t \ \ddot{Z}_t]^T \\ \mathbf{y}_t &= [X_{Gt} \ Y_{Gt} \ Z_{Gt} \ \dot{X}_t \ \dot{Y}_t \ \dot{Z}_t \ \ddot{X}_t \ \ddot{Y}_t \ \ddot{Z}_t]^T \\ F_t &= \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & \Delta t^2/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \Delta t^2/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \Delta t^2/2 \\ 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (17) \end{aligned}$$

Analytical solution of the dynamic and observation equations can be conducted in the same manner of equations (9) – (13).

3. IMAGE INTEGRATING METHOD

The position and rotation based on POS data filtering are used as initial value of the orientation factors in image-based self-localization method. The position is already represented in the geodetic coordinate, and then the distance of baseline is also supplied in the real scale. The image-based self-localization method consists of feature points extraction and tracking, relative orientation, three dimensional coordinates estimation of the feature points, and orientation factors updates of the mobile device. Figure 2 shows an overview of the proposed method.

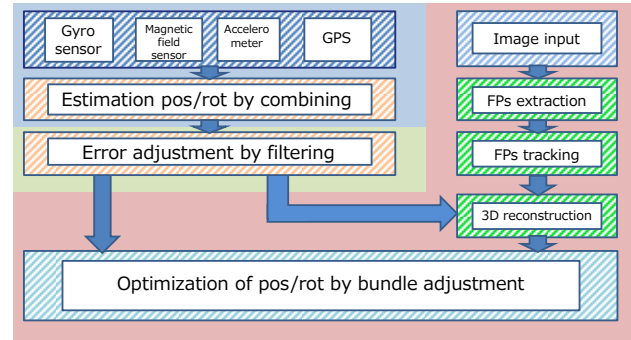


Figure 2. Flow of the POS and camera integration

3.1 Feature points extraction and tracking

Recently, sophisticated feature points extraction and tracking algorithms have been developed. One of the most reliable algorithms is SURF (Speeded-UP Robust Features) (Bay *et al.*, 2008). Figure 3 shows an example of feature points extraction by the SURF in our experiments.

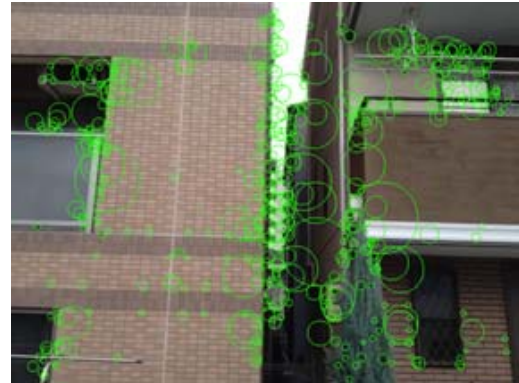


Figure 3. Feature points extraction by SURF

Even if the SURF is applied to feature points extraction and tracking, incorrect matching points are still exist. The feature points tracking are refined by using not only adjacent frames also sequential frames. Firstly, extracted feature points are searched in sequence between adjacent frames. After the tracking process within a certain number of frames, position of feature points are re-projected into first frames (back matching). If the displacement between first and last position of the points is larger than a threshold, the feature points are discarded. With the result of the matching, three dimensional coordinates of the feature points can be calculated. When the depth of the points is larger than a threshold, the feature points are also discarded.

After the above mentioned thresholding process, incorrect matching points still remain (Figure 4). Especially in the case

of application in urban area, similar textures make such incorrect matching. In this study, RANSAC (Random Sample Consensus) (Fischler and Bolles, 1981) is also applied (Figure 5). RANSAC algorithm is a method of outlier removal. Finally, the remaining points are accepted as feature points.

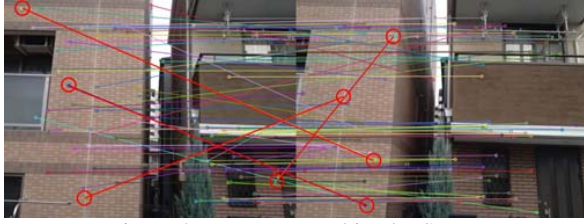


Figure 4. Incorrect matching by SURF



Figure 5. Correction of matching results by RANSAC

3.2 Three dimensional coordinates estimation

The position and rotation of a mobile sensor are already acquired through POS data filtering described in the previous chapter. With the position and rotation, and feature points tracking results, three dimensional coordinates of the feature points can be calculated. Camera calibration is conducted in advance. For the calibration, checkerboard is used.

First of all, an initial three dimensional coordinates is built based on intersection (Stewenius *et al.*, 2006). For the optimization of intersection, RANSAC algorithm (Fischler and Bolles, 1981) is also applied.

3.3 Orientation factors updates

Once initial value of the orientation factors and three dimensional coordinates of the feature points are acquired, the orientation factors updates are conducted by bundle adjustment (Triggs *et al.*, 2000; Luhmann *et al.*, 2014).

The feature points have coordinates $\mathbf{p}_i = (X_i, Y_i, Z_i)$. The each sequential frame has a three dimensional coordinates $\mathbf{q}_j = (X_j, Y_j, Z_j)$ as the camera position. At the frames j , the feature point i has camera coordinate system (u_{ij}, v_{ij}) . A transformation between the camera coordinate and the world coordinate systems represents collinearity equation.

$$\begin{aligned} \bar{u}_{ij} &= u_{ij} - \Delta u_{ij} = -c \frac{a_{11}(X_i - X_j) + a_{12}(Y_i - Y_j) + a_{13}(Z_i - Z_j)}{a_{31}(X_i - X_j) + a_{32}(Y_i - Y_j) + a_{33}(Z_i - Z_j)} \\ \bar{v}_{ij} &= v_{ij} - \Delta v_{ij} = -c \frac{a_{21}(X_i - X_j) + a_{22}(Y_i - Y_j) + a_{23}(Z_i - Z_j)}{a_{31}(X_i - X_j) + a_{32}(Y_i - Y_j) + a_{33}(Z_i - Z_j)} \end{aligned} \quad (18)$$

where c = focal length
 $\Delta u_{ij}, \Delta v_{ij}$ = factors of interior orientation
 a_{kl} = factors of rotation matrix

The position and rotation updates are computed iteratively by minimizing a robust objective function of the re-projection error.

$$\begin{aligned} E(\mathbf{p}_1, \dots, \mathbf{p}_M, \mathbf{q}_1, \dots, \mathbf{q}_N) &= \frac{1}{2} \|\mathbf{e}\|^2 \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \left\{ (u_{ij} - \bar{u}(\mathbf{p}_i, \mathbf{q}_j))^2 + (v_{ij} - \bar{v}(\mathbf{p}_i, \mathbf{q}_j))^2 \right\} \end{aligned} \quad (19)$$

$$\mathbf{x} = \arg \min \sum \text{Obj} \left(\frac{|e_k|}{\sigma_k}, \sigma_T \right) \quad (20)$$

Tukey bi-weight objective function $\text{Obj}(\sigma_T)$ is applied as a robust objective function and \mathbf{x} is a set of parameters. Iteration of reweighted least squares method is used to allow the robust estimator to converge.

There are two types of the bundle adjustment: full bundle adjustment and local bundle adjustment. The local bundle adjustment uses only some recent frames. The full bundle adjustment is more accurate than the local bundle adjustment, but computational load is more expensive. In the sense of computation, the local bundle adjustment is more preferable at the expense of accuracy (Arth *et al.*, 2009). The local bundle adjustment can be applied recursively (McLauchlan, 2000).

$$E_{1:j+1} = E_{1:j} + E_{j+1} \quad (21)$$

$E_{1:j}$ expresses the objective function by using from 1st frame to j th frame. According to the recursive form, bundle adjustment can be conducted effectively. It is important to point out here that the accuracy depends on the number of frames with the recursive form. We examined the relationships between number of frames and computation time / sum of squared error. In this study, based on the length of baseline, the local bundle adjustment is applied for improving the accuracy.

In order to solve the bundle adjustment problem, Levenberg-Marquardt method (Hartley and Zisserman, 2004) is applied. The objective function E is approximated by the following formula:

$$E(\mathbf{x} + \delta \mathbf{x}) \approx E(\mathbf{x}) + \mathbf{g}^T \delta \mathbf{x} + \frac{1}{2} \delta \mathbf{x}^T \mathbf{H} \delta \mathbf{x} \quad (22)$$

$$(\mathbf{H} - \lambda \mathbf{I}) \delta \mathbf{x} = \mathbf{g} \quad (23)$$

where

$$\begin{aligned} \mathbf{g} &= \left. \frac{dE}{d\mathbf{x}} \right|_{\mathbf{x}} \quad (\text{gradient}) \\ \mathbf{H} &= \left. \frac{d^2 E}{d\mathbf{x}^2} \right|_{\mathbf{x}} \quad (\text{Hessian}) \\ \lambda &= \text{damping factor} \end{aligned}$$

4. EXPERIMENTS

The proposed method is applied to POS data and images taken in urban area. An iPhone4s is used as a mobile device, which is equipped with GPS, gyro sensor, acceleration meter, magnetic field sensor (POS) and a camera. Camera calibration is conducted in advance. In order to evaluate the accuracy, orientation data of Mobile Mapping System (MMS) of Trimble are compared. The MMS is also equipped with the iPhone4s.

The experimental site is in urban area (Figure 6), at which condition of GPS receiver is appropriate.



Figure 6. Experimental site

4.1 Experiment of POS data filtering

First experiment is application of POS data filtering. Figure 7 compares estimation results of roll, pitch and yaw with MMS data.

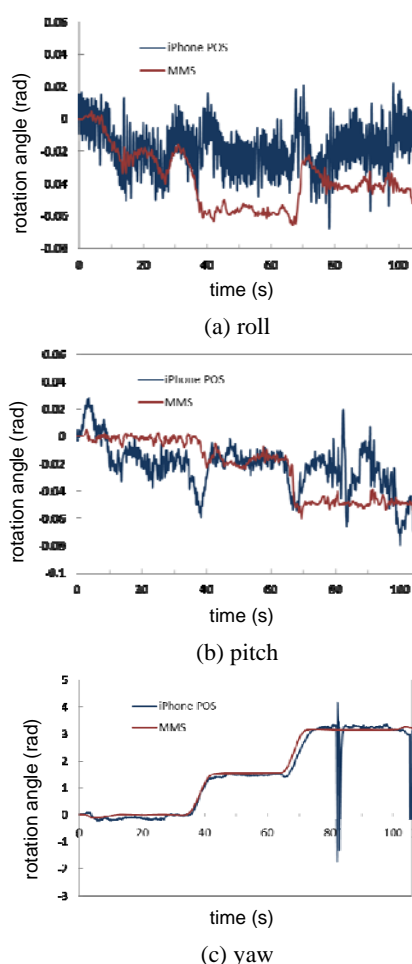


Figure 7. Results of rotation angle with POS sensor fusion

The displacement between estimated value and MMS data with respect to roll, pitch, and yaw is 0.024 (rad), 0.031 (rad), and 0.17 (rad) on average, respectively. In the result of yaw, the gap was large, because of limitation of function in iPhone, that is negative value of yaw angle is converted to zero automatically.

Position accuracy is also examined (Figure 8). The root mean square of the positions displacement is 5.67 (m)

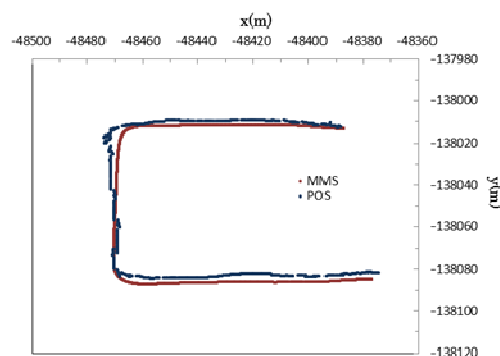


Figure 8. Result of position with POS sensor fusion

4.2 Experiment of POS and camera integration

As shown in Figure 8, accuracy of POS data filtering became worse at the area in a curve. Focusing on such area, result of POS and camera integration was examined.

Figure 9 shows the accuracies of rotation angle by all sensors (POS and camera) integration. From the comparison between the accuracies of the proposed method and ones of POS data filtering, clear improvement was not confirmed.

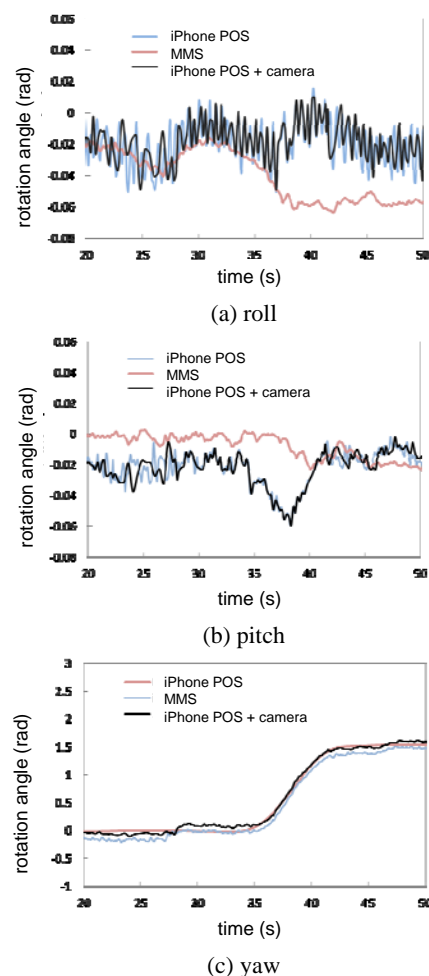


Figure 9. Results of rotation angle with all sensors integration

Figure 10 shows the result of all sensors integration. Compared with the previous result, the accuracy improvement can be confirmed. During the application area, the root mean square of the positions displacement with only POS data filtering is 3.59 (m). On the other hand, one with POS and camera integration improved to 3.26 (m).

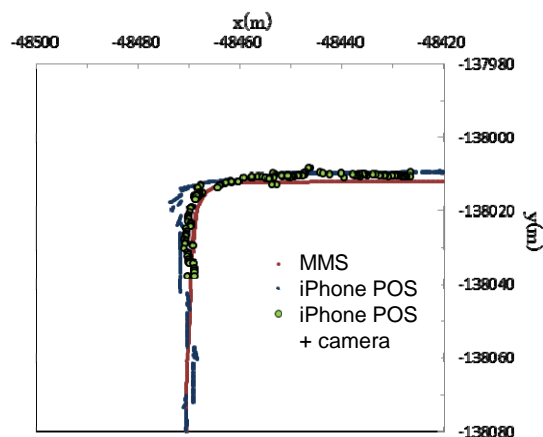


Figure 10. Result of position with all sensors integration

5. CONCLUSIONS

This paper develops the self-localization method using sensors, such as POS and cameras, on mobile devices simultaneously. The proposed method applies the Kalman filter in order to combine all sensors data except for camera (POS data filtering). Additionally, by using the position and rotation from POS data filtering as initial value of bundle adjustment, POS and camera integration method is achieved.

Through experiments with real data, the accuracy improvements of position and rotation by POS data filtering were confirmed. The results of final integration method improved the accuracy. It means that proposed self-localization method with POS and camera make the accuracy more sophisticated compared with only POS data filtering. Especially, the improvement at the area in a curve is noticeable. According to the experiments, the significance of the proposed method is confirmed.

As a further work, accuracy of three dimensional coordinates estimation of feature points will be evaluated by comparing with laser scanner data on MMS. Additionally, integrated filtering method between POS data filtering and bundle adjustment will become challenging investigation. As a result, promising method can be constructed, and then more impressive visualization will be accomplished.

REFERENCES

- Agrawal, M. and Konolige, K., 2006. Real-time localization in outdoor environments using stereo vision and inexpensive GPS. *Proceedings of the 18th International Conference on Pattern Recognition*, pp.1063-1068.
- Arth, C., Wagner, D., Klopschitz, M., Irschara, A., Schmalstieg, D., 2009. Wide area localization on mobile phones.

Proceedings of 8th IEEE International Symposium on Mixed and Augmented Reality, pp.73-82.

Bay, H., Ess, A., Tuytelaars, T. and Vangool, L., 2008. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3), pp.346-359.

Daehne, P. and Karigiannis, J., 2002. Archeoguide: system architecture of a mobile out door augmented reality system. *Proceedings of 1st IEEE International Symposium on Mixed and Augmented Reality*, pp. 506-514.

Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), pp.381-395.

Hartley, R. and Zisserman, A., 2004. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge.

Kourogi, M. and Kurata, T., 2005. Personal positioning based on walking locomotion analysis with self-contained sensors and wearable camera. *Proceedings of IEEE/ACM International Symposium on Virtual Reality Conference*, pp. 71-78.

Luhmann, T., Robson, S., Kyle, S. and Boehm, J., 2014. *Close Range Photogrammetry and 3D Imaging, 2nd Edition*. Walter de Gruyter GmbH, Berlin.

McLauchlan, P.F., 2000. A batch / recursive algorithm for 3D scene reconstruction. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 2738-2743.

Reitmayr, G. and Drummond, T.W., 2005. Going out: robust model-based tracking for outdoor augmented reality. *Proceedings of IEEE on Virtual Reality Conference*, pp. 71-78.

Ristic, B., Arulampalam, S., Gordon, N., 2004. *Beyond the Kalman Filter: Particle Filter for Tracking Applications*. Artech House Publishers, Boston.

Schall, G., Wagner, D., Reitmayr, G., Taichmann, E., Wieser, M., Schmalstieg, D., Hofmann, W.B., 2009. Global pose estimation using multi-sensor fusion for outdoor augmented reality. *Proceedings of 8th IEEE International Symposium on Mixed and Augmented Reality*, pp.153-162.

Stewenius, H., Engels, C. and Nister, D., 2006. Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60, pp284-294.

Triggs, B., McLauchlan, P., Hartley, R. and Fitzgibbon, A., 2000. Bundle adjustment – a modern synthesis. In Triggs, B., Zisserman, A. and Szeliski, R. eds., *Vision Algorithms: Theory and Practice*, Springer-Verlag, Berlin.

Wagner, D., Schmalstieg, D., Bischof, H., 2009. Multiple target detection and tracking with guaranteed framerates on mobile phones. *Proceedings of 8th IEEE International Symposium on Mixed and Augmented Reality*, pp.57-64.