Vision Sensor-Based Human Tracking Control of a Quadcopter UAV with Position Uncertainty

Taeseok Jin

Department of Mechatronics Engineering, Dongseo University
jints@dongseo.ac.kr

Abstract

In this paper, we describe a method for the attitude of a quadcopter UAV using image information of a moving object like a walking human. This method combines the observed position from GPS sensors and the estimated position from the images captured by a UAV's camera to localize a walking human. Using the a priori known path of a quadcopter UAV in world coordinates and a perspective camera model, we derive the geometric constraint equations which represent the relation between image frame coordinates for a moving object and the estimated quadcopter UAV's attitude. The proposed method utilizes the error between the observed and estimated image coordinates to localize the quadcopter UAV. The Kalman filter scheme is applied for this method. The performance of proposed algorithm is verified by the computer simulation and the experiments.

Keywords*: Quadcopter, multiple vision, tracking, Kalman filter, color recognition*

1. Introduction

As the minimization of unmanned aerial vehicles (UAVs) and performance improvement of sensors have made headway in recent times, research and commercialization also followed suit in various fields of application such as physical distribution services, disaster relief, observation work, and entertainment. Owing to the vertical mobility and hovering characteristics of small quadcopters, they are autonomous and able to fly indoors [1, 2].

To perform these functions, sensor-based position control is considered as an important research subject for quadcopter environment monitoring, position prediction, and tracking control. However, because UAVs such as small quadcopters cannot predict their own location based on the odometer information in their wheels (in contrast to land robots), this work has performed a basic study on the location estimation method based on the acquisition of images to solve these issues. An image sensor was installed in a UAV to acquire images of objects and obtain test results that estimate the location of a quadcopter position based on the acquired images in terms of uncertainty [3,9].

In this study, the subjects (walking human) were recognized through image processing in order for the quadcopter to move closely alongside them, and the subject, camera status, and geometric constraints between the image coordinates were applied to calculate the location. The route was predicted using a Kalman filter to determine an effective tracking method. The subjects performed color-recognition tasks from a camera installed in the quadcopter, and the location and distance were acquired. The Kalman filter, which is a nonlinear status estimator, was used to estimate the subject linear speed [4]. The linear speed and angular velocity of the moving objects were used in estimating the trajectory of the subjects; thus, they were effectively tracked [4-6] as the quadcopter moved along the tracked trajectory.

The system of the quadcopter used for the tracking experiments in this study was constructed as shown in Figure 1, and the mechanical parts were divided into the driving,

sensor, and control units. Some information on the system in this study were composed as shown in Figure 1.
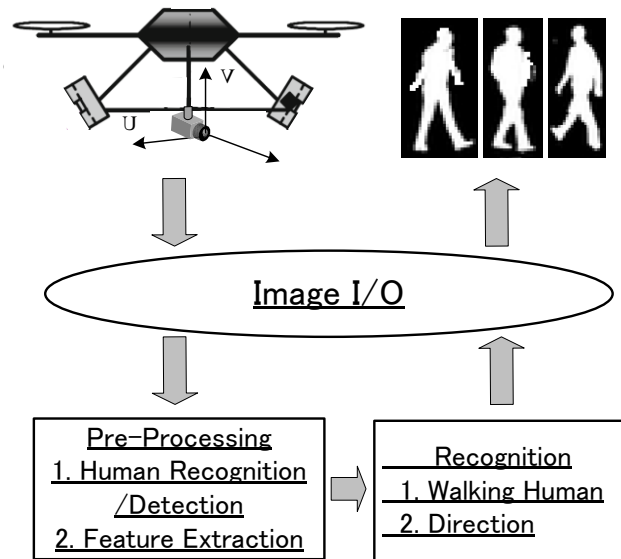

Figure 1. Process of Walking Human Tracking

2. Local Estimation of Walking Human

2.1. Walking Human Image Projection

As an experiment to recognize and estimate the moving objects, color information of the pedestrians was used to control the location and position of a plane body in this study. The location information of the pedestrians was provided to the quadcopter in advance, and when the quadcopter attempted to correct its location, the location estimation degree increased, and pedestrians were observed while the quadcopter was in flight to increase its flight mobility. The route of the walking pedestrians was linearly displayed on the image coordinates, and the current estimated location of the quadcopter was used to induce the geometric constraint equation by converting the coordinates [10,11].

Figure 2 shows that the quadcopter was equipped with a charge-coupled device (CCD) camera, and the camera was installed so that the optical axis below the quadcopter can be placed on the quadcopter X_R axis and a straight line. The camera's geometric model is a common pinhole model [7]. Here, (x_W, y_W, z_W) expresses the standard coordinates, (x_R, y_R, z_R) expresses the plane-body coordinates, and (u,v) expresses the image coordinates [7, 8].
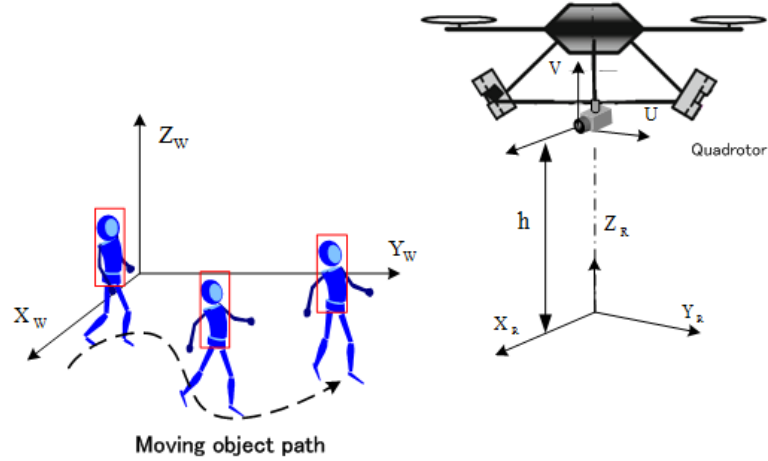
Figure 2. Coordinates of Walking Human and Quadcopter

The trajectory of the walking human on the $x_W - y_W$ plane of the reference coordinates, without loss of generality, is assumed to be as follows:

$$f(x_W, y_W) = 0 \tag{1}$$

$z_W (= z_0 \neq h)$ is also assumed to be constant and not equal to the camera height of quadcopter, h.

The walking human trajectory in the reference coordinates can then be transformed into the quadcopter coordinates, as follows:

$$\begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix} = \begin{bmatrix} \cos\hat{\theta}_r & \sin\hat{\theta}_r & 0 \\ -\sin\hat{\theta}_r & \cos\hat{\theta}_r & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_W - \hat{x}_r \\ y_W - \hat{y}_r \\ z_W \end{bmatrix}, \tag{2}$$

where $[\hat{x}_r \quad \hat{y}_r \quad \hat{z}_r]^T$ represents the current estimated position of the quadcopter and $[x_W \quad y_W \quad z_W]^T$ represents the position of the walking human.

This point ($[x_R \quad y_R \quad z_R]^T$) is mapped once more onto the image frame using perspective projection, as follows [7, 11]:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \lambda \dfrac{y_R}{x_R} \\ \lambda \dfrac{z_R - h}{x_R} \end{bmatrix}, \tag{3}$$

where λ represents the camera focal length and $z = \begin{bmatrix} u & v \end{bmatrix}^T$ is the position of the walking human on the image frame. Based on the Eq. (2) and Eq. (3), the geometric constraint equation can be generally represented as

$$f(z, \hat{x}) = 0, \tag{4}$$

where $\hat{x} = [\hat{x}_r, \hat{y}_r, \hat{z}_r]^T$ represents the current estimated position of the quadcopter.

2.2. Position Correction

There is some discrepancy between the calculated position of the walking human in the image frame, which is based on the estimated quadcopter position, and the actual value. Utilizing this error, the real position of the quadcopter can be corrected recursively. To overcome vague input information, *i.e.,* the noise of the human position in the image frame and uncertainty components of the position estimation of the quadcopter, the Kalman filtering technique is adopted to develop a robust observer [14, 15]. The geometric constraint equations between the human image coordinates and the quadcopter position are approximated to a linear equation, and the Kalman filtering technique is applied to estimate the quadcopter position.

It is assumed that the *i*-th measured vector, *i.e.,* the position of the walking human, $\hat{z}_i$, includes noise with the following average and variance:

$$\hat{z}_i = z_i + v_i,\qquad(5)$$

where $E[v_i] = 0$ and $E[v_i v_i^T] = S_i$.

Using the Taylor series expansion and ignoring the higher order nonlinear terms at the measured vector, $\hat{z}_i$, and the estimated position of the quadcopter, $\hat{x}_{i-1}$, the nonlinear constraint equations are approximated to linear equation as follows:

$$
\begin{aligned}
f(z_i, \hat{x}) &= 0 \\
&\approx f(\hat{z}_i, \hat{x}_{i-1}) + \left.\frac{\partial f}{\partial z}\right|_{z=\hat{z}_i} (z_i - \hat{z}_i) + \left.\frac{\partial f}{\partial \hat{x}}\right|_{\hat{x}=\hat{x}_{x-1}} (\hat{x} - \hat{x}_{i-1}),
\end{aligned}\qquad(6)
$$

where $\left.\dfrac{\partial f}{\partial z}\right|_{z=\hat{z}_i}$ and $\left.\dfrac{\partial f}{\partial \hat{x}}\right|_{\hat{x}=\hat{x}_{i-1}}$ represent the estimated value of $\dfrac{\partial f}{\partial z}$ and $\dfrac{\partial f}{\partial \hat{x}}$ at $\hat{z}_i$ and $\hat{x}_{i-1}$, respectively.

In a linear system, Eq. (6) can be rearranged as the following matrix equation [8]:

$$y_i = M_i \hat{x} + u_i,\qquad(7)$$

where $y_i = -f(\hat{z}_i, \hat{x}_{i-1}) + \left.\dfrac{\partial f}{\partial \hat{x}}\right|_{\hat{x}=\hat{x}_{i-1}} \hat{x}_{i-1}$, $M_i = \left.\dfrac{\partial f}{\partial \hat{x}}\right|_{\hat{x}=\hat{x}_{i-1}}$, and $u_i = \left.\dfrac{\partial f}{\partial z}\right|_{z=\hat{z}_i} (z_i - \hat{z}_i)$.

In this equation, y_i is the new measured vector, M_i linearly combines the measured vector and the quadcopter position, $\hat{x}$, and u_i is the error in linearization of the measured vector with the following average and variance values [6]:

$$E[u_i] = 0,\qquad(8)$$

$$E[u_i u_i^T] = W_i = \frac{\partial \hat{f}}{\partial z} S_i \frac{\partial \hat{f}}{\partial z}^T,\qquad(9)$$

Since M_i and y_i are a priori given values, if the average and variance of u_i are known, we can obtain the optimal estimated value of $\hat{x}$ with the new variance. The Kalman filter provides the estimated value, $\hat{x}$, which minimizes the expected squared error norm, $E[(\hat{x} - x)^T (\hat{x} - x)]$, as the linear combination of the measured vectors, $\{y_i\}$, as follows:

$$\hat{x}_i = \hat{x}_{i-1} + K_i(y_i - M_i \hat{x}_{i-1}), \tag{10}$$

$$K_i = P_{i-1} M_i^T (M_i P_{i-1} M_i^T + W_i)^{-1}, \tag{11}$$

$$P_i = (I - K_i M_i) P_{i-1}, \tag{12}$$

where K_i represents the Kalman gain, P_i is the zero-mean-variance matrix of the estimated error by the ith measured vector, and $\hat{x}_i$ is the estimated quadcopter position by the i-th measured vector.

The initial quadcopter position estimation and variance, $\hat{x}_0$ and P_0, can be obtained using the quadcopter driving model. Using n image frames from the image coordinates of the moving object, the final quadcopter position is recursively estimated as $\hat{x}_n$, with a variance of P_n.

2.3. Direction Correction According to Gaussian distribution

Figure 3 shows a two-dimensional Gaussian distribution that provides deviation information on a single observation of an object. The average of the central distribution is the estimated location of the pedestrian, and this axis displays the degree of standard deviation according to the maximum and minimum axes of the distribution relative to the estimation of the uncertainty of the observation. The distribution value for the specified location must be identical to the possibility that the pedestrian is actually at that location.

The pedestrian information for the specified location must be characterized by calculating the average, standard deviation, and angle of the fused distribution with regard to the observation of the indication point recognized by the quadcopter while hovering. The average, standard deviation, measured distribution angle received from the sensor, and sensor-error model regarding the recognized pedestrian can be accordingly calculated. A method of determining the parameters combined for the individual distributions is required, and such is generally being used as follows:
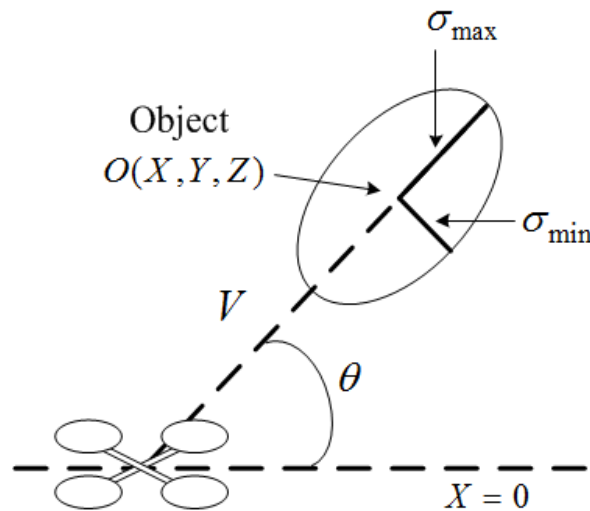
Figure 3. Gaussian Distribution Parameter: Average (X, Y), Ellipsoid Parameters were combined for Individual Distribution (σ $_{max}$, σ $_{min}$) Distance from Object *V*

The matrix pattern of the Kalman filter adopted by Smith and Cheeseman proposed several simple calculation methods [3]. These methods expressed a fused covariance

matrix pattern without considering the conversion factors, whereas the average, standard deviation, and direction of the maximum axis are independent of the conversion. A three-dimensional regular distribution of the Gaussian distribution is accordingly determined according to the standard deviation, σ, covariance matrix, C and average. Covariance matrix $\bar{C}$, relative to the observation indicates the existing research results as the maximum and minimum axes of the standard deviation, as expressed in Equation (13) [12, 13].

$$\bar{C} = \begin{bmatrix} \sigma_{max}^2 & 0 \\ 0 & \sigma_{min}^2 \end{bmatrix} \qquad (13)$$

Since the observation may be oriented arbitrarily with respect to the global coordinate frame, it must first be rotated to align with this frame:

$$C = R(-\theta)^T \underline{C} R(-\theta) \qquad (14)$$

where θ is the angle of the distribution's principal axis with respect to the global x-axis. This rotation accomplishes the transformation from observation parameters to the canonical form. Once the observations are in the canonical form, we continue to merge the observations into one.

The covariance matrices of two distributions, C1 and C2, can be combined into a single covariance matrix, C, as

$$C = C_1 - C_1[C_1 + C_2]^{-1}C_1. \qquad (15)$$

Now the mean of the resulting merged distribution, X, is computed from the individual distribution means and covariance matrices as follows:

$$\hat{X} = \hat{X}_1 + C_1[C_1 + C_2]^{-1}(\hat{X}_2 - \hat{X}_1). \qquad (16)$$

The hovering position of the quadcopter can be corrected through image recognition of the pedestrian, and the three-dimensional uncertainty ellipsoid relative to the pedestrian movements can be displayed as follows.

3. Experiments

3.1. Experimental Conditions

First, the UAV body was set so that it could be tracked while in a hovering state to maintain strong recognition of the pedestrian. The pedestrian's movements proceeded in a curved path like a parabola to express a complicated pattern of movement. The experimental conditions that considered the actual parameters were as follows:

Camera height: $h \cong 300cm$

Focal length: $\lambda = 1.55cm$

CCD size: $0.60cm(H) \times 0.52cm(V)$

The covariance of the estimated vectors is independent of each other, and the data acquired through the experiment, as expressed in Equation (17), were used considering a standard deviation of approximately three pixels.

$$S = \begin{bmatrix} 0.005^2 & 0 \\ 0 & 0.005^2 \end{bmatrix} \qquad (17)$$

When the quadcopter's control period was set to 100 s and the input speed error of the quadcopter's four rotor blades were set to 3% of actual input, the Kalman filter's initial estimated value of location and covariance of error were the values based on the experimental results of the quadcopter's flight model. The pedestrian decided to walk along an arc-shaped route expressed in Equation (18).

$$Y_W = -0.01(X_W - 500)^2 + 300,$$
$$X_W = 200 - 500, \tag{18}$$
$$Z_W = 170$$

3.2. Results

Several experiments were performed with the human walking in *S*-curve directions, making turns. In all cases, the quadcopter is able to closely follow the walking human. Figure 4 shows a simulation of the walking human with varying speeds and patterns of movement with changes in direction. At all times, the quadcopter satisfactorily follows the walking human.
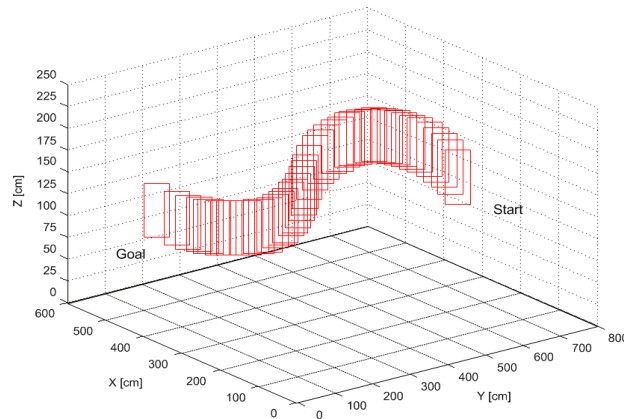

Figure 4. Motion Path of a Walking Human Captured By Simulation

Figures 5(a) and (b) show the tracing movement trajectory of the pedestrian, and the bounding box that indicates the recognition status regarding the rear state obtained while the subject moved from the point of departure to the destination in a forward direction respectively. Here, the error intervals of the bounding box in Figure 5(b) are shown to be slightly narrower than those in the bounding box in Figure 5(a). The error interval of the bounding box may look narrower according to the decrease in uncertainty when the pedestrian arrived at the location compared with the increase in uncertainty at the point of departure.

According to the experimental results, the recognition uncertainty with regard to the pedestrian decreased the closer he/she approached his/her destination. The uncertainty ellipsoid on the estimated values of the location and covariance before and after correction is shown in Figure 6.
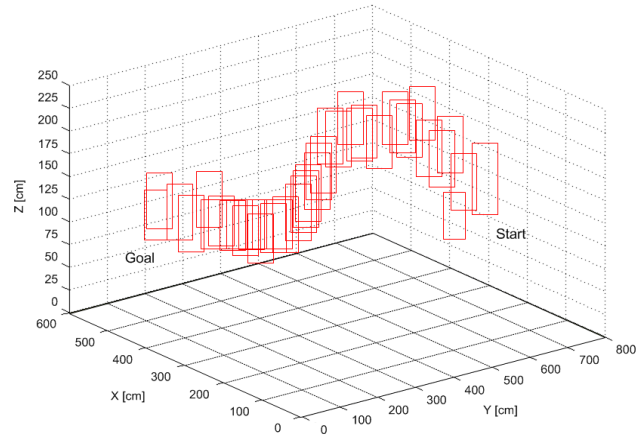
(a) Motion Path of a Walking Human Captured by Hovering Quadcopter
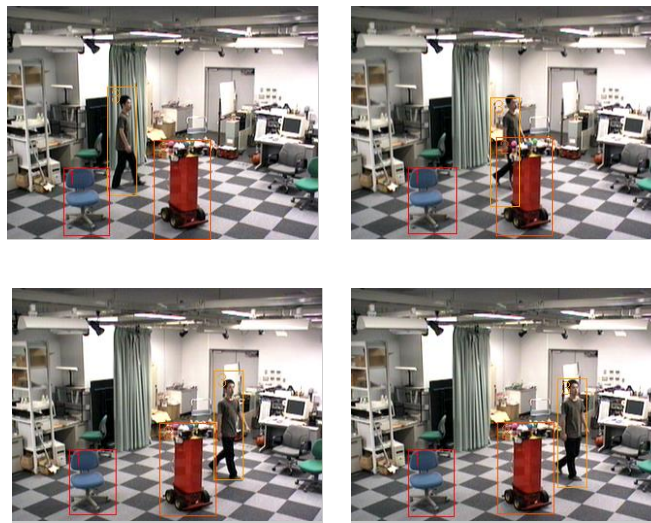
(b) Reverse Direction (goal–start)
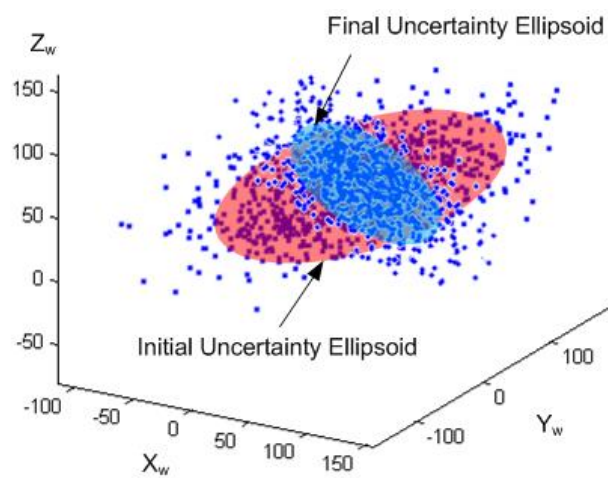
Figure 5. Pedestrian Tracking from the Video

Figure 6. Uncertainty Ellipsoid of the Location Estimation

4. Conclusion

In this paper, a method has been presented that estimates the absolute location of a pedestrian through an estimation method of tracking pedestrians while a quadcopter is hovering. The uncertainty ellipsoid was used to quantitatively show the location estimation uncertainty of the quadcopter. The quadcopter location was shown to be correctable by applying the Kalman filter and the geometric constraints based on the images through coordinate conversions.

To examine the presented method, a pedestrian that moved along a specified route was projected on a video camera loaded in a UAV, and the constraints generated between the object image coordinates and the quadcopter current estimated location were induced. Because this information was calculated based on the quadcopter-estimated location, the pedestrian predicted image coordinates were different from those actually observed. This difference was used to inversely determine the quadcopter location. Because the pedestrian image coordinates included static noise during processing, an uncertain component existed in the quadcopter's location estimate.

The Kalman filter that robustly estimates the status was used to track the moving objects. The tests results showed that the accuracy of the quadcopter location estimation was further influenced by the arc-shaped path of the movement of the pedestrian rather than a linear path.

Acknowledgements

References

[1] M. Abdolhosseini, Y.M. Zhang and C.A. Rabbath, "An efficient model predictive control scheme for an unmanned quadrotor helicopter", Journal of Intelligent & Robotic Systems, vol. 70, (2013), pp.27–38.

[2] A. Razinkova, B.J. Kang, H.C. Cho and H.T. Jeon, "Constant Altitude Flight Control for Quadrotor UAVs with Dynamic Feedforward Compensation", International journal of fuzzy logic and intelligent systems, vol. 14, no.1, (2014), pp. 26-33.

[3] R.C. Smith and P. Cheeseman, "On the Representation and Estimation of Spatial Uncertatinty", The International Journal of Robotics Research, vol. 5, no.4, (1986), pp. 56-68.

[4] A. Alaimo, V. Artale, C. Milazzo and A. Ricciardello, "Pid controller applied to hexacopter flight", Journal of Intelligent & Robotic Systems, vol. 73, (2014), pp. 261–270.

[5] R. Ritz, Mark W. Muller, M. Hehn and R. D'Andrea, "Cooperative Quadrocopter Ball Throwing and Catching", in IEEE/RSJ International Conference on Intelligent Robots and Systems, (2012), pp. 4972-4978.

[6] F. Beltramini, M. Bergamasco and M. Lovera, "Experiment design for mimo model identification, with application to rotorcraft dynamics", In World Congress, vol. 8, (2011), pp. 14392–14397.

[7] M. Achtelik, A. Bachrach, R. He, S. Prentice and N. Roy, "Autonomous navigation and exploration of a quadrotor helicopter in GPS-denied indoor environments", in Robotics: Science and Systems Conference, (2008).

[8] Y.C. Bae, Y. Kim and C.S. Kim, "Obstacle Avoidance Method in the Chaotic Unmanned Aerial Vehicle", Journal of the korean institute of intelligent systems, vol. 14, no. 7, (2004), pp. 883-888.

[9] T.S. Jin, "Kinematic Method of Camera System for Tracking of a Moving Object", vol.8, no.2, (2010), pp. 145-149.

[10] W. F. Phillips, "Mechanics of Flight", John Wilet & Sons, (2004).

[11] J.M. Rico-Martinez and J. Gallardo-Alvarado, "A simple method for the determination of angular velocity and acceleration of a spherical motion through quaternions", Meccanica, vol. 35, (2000), pp.111-118.

[12] T.S. Jin, "Command Fusion for Navigation of Mobile Robots in Dynamic Environments with Objects", Journal of information and communication convergence engineering, vol.11, no.1, (2013), pp. 24-29.

[13] R. F. Stengel, "Flight Dynaimcs", Princeton University Press, **(2004)**.

[14] F. Han, Y. Shan, R. Cekander, H. Sawhney and R. Kumar, "A two-stage approach to people and vehicle detection with hog-based SVM", In Performance Metrics for Intelligent Systems Workshop in conjunction with the IEEE Safety, Security, and Rescue Robotics Conference, **(2006)**, pp. 133-140.

[15] R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, V. Manohar, M. Boonstra and V. Korzhova, "Performance evaluation protocol for face, person and vehicle detection & tracking in video analysis and content extraction (vace-ii)", In Workshop of Classification of Events, Activties and Relationships, **(2006)**.

Authors

Taeseok Jin, he received his Ph.D. degree from Pusan National University, Busan, Korea, in 2003, in Electronics engineering.

He is currently an associate professor at Dongseo University. From 2004 to 2005, he was a Postdoctoral Researcher at the Institute of Industrial Science, University of Tokyo, Japan. His research interests include network sensors fusion, mobile robots, computer vision, and intelligent control. Dr. Jin is a Member of the KIIS, KIEE, ICROS, and JRS.