TELKOMNIKA, Vol. 11, No. 2, February 2013, pp. 1093~1100
ISSN: 2302-4046

Vision-aided Navigation for Autonomous Aircraft Based on Unscented Kalman Filter

Junwei Yu[1], Junquan Yu[2], Xiaobo Jin[1], Guicai Wang[1]
[1]College of Information Science and Engineering, Henan University of Technology, Zhengzhou, Henan, 450001, China
[2]Henan Jinxing Beer Group Co., Ltd., Zhengzhou, Henan, 450009, China
Corresponding author, e-mail: yujunwei@126.com

Abstract

A vision-aided navigation system for autonomous aircraft is described in this paper. The vision navigation of the aircraft to the known scence is performed with a camera fixed on the aircraft. The location and pose of the aircraft are estimated with the corresponding control points which can be detected in the images captured. The control points are selected according their saliency and are tracked in sequential images based on Fourier-Melline transform. The simulation model of the aircraft dynamics and vision-aided navigation system based on Matlab/Simulink is built.The unscented Kalman filter is used to fuse the aircraft state information provided by the vision system and the inertial navigation system. Simulation results show that the vision-based navigation system provides satisfactory results of both accuracy and reliability.

Keywords*: vision location, aircraft navigation, unscented kalman filter, simulation*

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

The subject of autonomous aircrafts (which are also known as unmanned aerial vehicles, UAV) has drawn much research attention in both civilian and military areas since they have many potential uses in surveillance, remote sensing, target acquisition, firefighting, scientific research and other applications. Autonomous aircrafts often preferred for missions that are too dangerous for manned aircraft. The navigation system is equipped with various sensors such as inertial measurement unit (IMU), GPS and altimeter to increase their capability. The inertial navigation system (INS) is reliable, low-cost and commonly used with the greatest shortcoming of error accumulation. GPS receiver can provide highly accurate real-time position information to recalibrate the INS. However, GPS probably fail to work due to jamming, interference and terrain geometry. Recently, vision system is one of the most attractive aided navigation systems for autonomous aircrafts [1, 2]. Vision-aided system has many advantages of small size, low cost, high precision, and good autonomy. Vision-aided system can provide precise location and orientation information and determine control commands for aircraft controller.

Vision location and information fusion are two important processes of vision-aided navigation system. We can estimate the location and orientation of the aircraft using corresponding features. Resent survey [3] declares that many feature extractors are applied to different research areas like robot localization and image registration. The extractors include Hessian-Affine, Harris, SIFT, etc. However, it is time-consuming to handle numerous features in the processes of features detection, matching and pose estimation. We are focused on developing saliency descriptor of the features for visual navigation [4]. With a small number of most salient features selected as the control points, the efficiency and robustness of navigation system will be greatly improved. The vision scene changes rapidly in scale, rotation and illumination for vision-aided autonomous aircraft. Matching schemes must be defined with a proper measure of similarity between features in sequential images. Existing matching approaches such as patch classification and SIFT descriptor are computationally expensive. To get good robustness and runtime performance, we take a matching approach based on Fourier-Mellin transform [5].

Received December 8, 2012; Revised January 10, 2013; Accepted January 17, 2013

Figure 1 shows the vision-aided navigation system for autonomous aircraft. The integrated navigation system is composed of such measuring devices as IMU, GPS, altimeter and camera. The nonliear 6-DOF dynamical model of aircraft is built with the assumption that the aircraft is a rigid-body. To improve the precision and the stability of the vision-aided navigation system, an unscented Kalman filter(UKF) [6] is adopted to fuse the information of the vision system and the inertial navigation system in the integrated navigation system. As the kinematic equations of the autonomous aircraft and the vision-aided navigation system are highly non-linear, an unscented Kalman filter (UKF) is adopted to estimate the aircraft's location, velocity, and attitude using visual and inertial measurements. The integrated navigation system presented pose estimation and algorithm relies on tight integration information of the various sensors, thus leading to excellent estimation accuracy and robustness in the presence of modeling uncertainties, and nonlinearities. Simulation experiment is carried out to test the vision-aided navigation system based Matlab/Simulink.

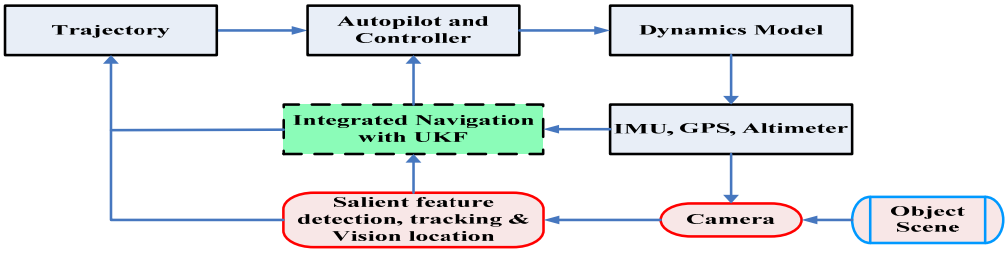


Figure 1. Vision-aided Navigation System for Autonomous Aircraft

The contents of this paper are organized as follows. In section 2, the vision location method with corresponding salient features is introduced. Section 3 presents the details of UKF algorithm and its application in the vision-aided navigation system. Section 4 gives the work environment, filter initializations and simulation results. Finally, conclusion and feature work are summarized.

2. Visual Location with Salient Features

2.1. Salient Feature Detection

Corner is one of the best features to be control point for vision location as it provides rich geometrical information of object. The control points are determined and loaded in the reference image. The corresponding control points can be detected and tracked in the real-time images. We can determine the location and orientation of the aircraft with these corresponding features. To improve the computational efficiency and environmental applicability of the visual navigation system, we select several most salient corners the control points. The location and pose of autonomous aircraft can be estimated with corresponding control points using the methods for perspective-n-point (PnP) problem.

We focus on the contour-based corners which define the main structure of objects. Existing contour-based corner detectors are mainly based on the curvature scale-space (CSS) [7][7]. As in (1), CSS corner detectors have two disadvantages in curvature estimation. First, the CSS curvature estimation involves second order derivatives which is sensitive to local variation and noise. Second, CSS requires appropriate Gaussian smoothing-scale selection which is a difficult task.

$$K = \frac{x_u y_{uu} - x_{uu} y_u}{(x_u^2 + y_u^2)^{1.5}} \quad (1)$$

Awrangjeb [8] proposed a corner detection technique based on the chord-to-point distance accumulation (CPDA) for the discrete curvature estimation. A chord is moved along a curve and the perpendicular distances from each point on the curve to the chord are summed to represent the curvature at that point. The CPDA discrete curvature at point P_k can be calculated

with equation (2). CPDA technique is completely based on the Euclidean distance and does not involve any derivative of the curve point. CPDA technique overcomes the problems associated with CSS corner detectors.

$$h_L(k) = \sum_{j=k-L}^{k} D_{k,j} \tag{2}$$

To save computational resources, we select several most salient corners as control points. Influence factors of feature saliency are scale, angle, gradient and local differences. The saliency response functions of these factors are built according their characteristics. Saliency descriptor of corner is defined as the linear combination of these functions. With the descending order of S_d , a certain amount of control points can be determined.

$$S_d = k_1 R_s + k_2 R_a + k_3 R_g + k_4 R_d \tag{3}$$

where, the S_d is the saliency descriptor of the corner. k_i, $i = 1, \cdots, 4$ are the weighting coefficients. R_s is the response function of scale. The important features are salient in multi-scales of image. R_s is calculated by the normalization of curvature product at the arc-length scale space. R_a is the response function of angle. The angle $\angle C$ at corner P_k is defined by two nearest candidate corners and R_a is defined by (2).

$$R_a = \begin{cases} \left|\sin \angle C\right|, & \frac{\pi}{4} < \angle C < \frac{3\pi}{4}, \quad \frac{5\pi}{4} < \angle C < \frac{7\pi}{4} \\ 0.5\left|\sin \angle C\right|, & otherwise \end{cases} \tag{4}$$

The gradient of the corner also reflects its significance. R_g is the response function of gradient which can be calculated by the normalization of gradient magnitude. R_d is the response function of local difference. The salient feature should be distinctive in its local region. R_d is proportional to the reciprocal of the similarity with neighbour corners.

2.2. Feature Tracking

Matching control points between sequential images is an indispensable step in vision navigation. The visual scenes captured will change rapidly with the movement of autonomous aircraft. To obtain reliable matches, a feature similarity measurement based on Fourier-Mellin transform is defined to track the salient features in image sequence [5]. The Fourier-Mellin transform is a useful mathematical tool for image recognition with the operational characteristics of translation, rotation and scale. The Mellin transform on the conversion image at log-polar coordinates gives a transform-image that is invariant to translation, rotation and scale. The similarity of two features can be calculated using the amplitude spectrums of their local region as (5), where F_1 and F_2 are the features in two adjacent frames. To improve the matching efficiency, the search strategy of nearest neighbor(NN) is applied. The candidate control points are searched in the neighbor region (e.g. 20 pixels) of their corresponding points. In the adjacent frames, ratios of the corresponding distances between features are approximately equal. A consistency checking method based on the distance ratio is applied to eliminate the false matches.

$$C = \frac{\sum_{\omega_\xi, \omega_\phi} \left[\left|F_2(\omega_\xi, \omega_\phi)\right| - \overline{\left|F_2\right|}\right]\left[\left|F_1(\omega_\xi, \omega_\phi)\right| - \overline{\left|F_1\right|}\right]}{\left\{\sum_{\omega_\xi, \omega_\phi} \left[\left|F_2(\omega_\xi, \omega_\phi)\right| - \overline{\left|F_2\right|}\right]^2 \sum_{\omega_\xi, \omega_\phi} \left[\left|F_1(\omega_\xi, \omega_\phi)\right| - \overline{\left|F_1\right|}\right]^2\right\}^{0.5}} \tag{5}$$

2.3. Vision Location with SoftPOSIT

To determine the relative location and pose of the camera to the surrounding objects, coordinate systems including world coordinate system W, aircraft body coordinate system B, camera coordinate system C and image coordinate system I are established. The relationship between these coordinate systems is illustrated in Figure 2. According to the linear camera model, the optical center, the reference point and its corresponding image point compose a straight line, which is called collinear constraint. The principle of imaging and location is given by:

$$X_w = X_s + R_0 X_{cb} + \lambda R_0 R_c^b X_c \tag{6}$$

where, X_w is the world coordinates of the control point, and X_c is the image coordinate of the same point, X_{cb} is the offset between the origins in C and B, X_s is the world coordinates of the optical center, R_c^b is the rotation matrix from C to B which is determined by the camera fixing angle, R_0 is the rotation matrix from B to W which is determined by the attitude angles of the aircraft, λ is the scaling factor.

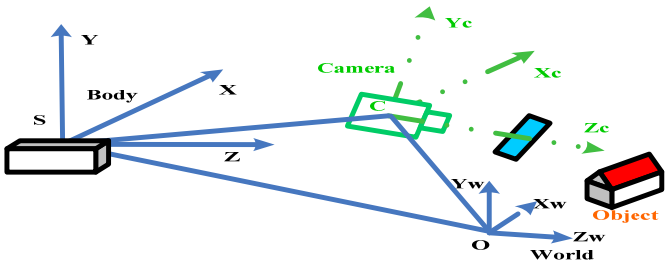


Figure 2. The Coordinate Systems of Imaging and Location

Once the control points on the object have been identified, it is possible to determine the location and pose of the aircraft by solving a set of equations of imaging. This pose estimation problem with correspondence points is known as PnP. Many techniques have been developed to solve this problem including direct linear transform, iterative algorithms, and non-iterative approaches. The linear algorithm is the simplest way to recover the pose of camera. However, we don't use linear algorithm in most applications as it is noise sensitive. Iterative algorithms such as SoftPOSIT [9] involve starting with a scaled orthographic projection model and then iteratively refining this initial estimate. As the number of control points is controllable with the saliency descriptor, we can get fast and accurate result using the iterative algorithm with the following pseudocode.

Pose and location estimation with SoftPOSIT

1) Initialization

Initialization of the homogeneous coordinates: $w_k = 1$;

Description of the control point coordinates $A = [X_w \; w_k]$ and its pseudo inverse matrix B。

2) Do the iteration

Compute two vectors of rotation： r_1, r_2;

Compute the third pose vector by $r_3 = r_1 \times r_2$;

Compute the image center: X_s, Y_s;

3) Parameters update

Coefficient of homogeneous coordinate： $w_k = A \cdot r_3 / Z_s$;

Update the image coordinates: $x_k = w_k \cdot x_k$, $y_k = w_k \cdot y_k$;

4）End the iteration

End the iteration if the variation of the image coordinates are less than the threshold.

5）Output the roation matrix and translation vector.

3. Vision-aided Navigation System based on UKF

3.1. 6. DOF Model of the Aircraft

The aircraft can be treated as a rigid body. Regardless the control strategies, we emphasize the state parameters of the aircraft such as north position, east position, altitude, pitch angle, yaw angle, roll angle and total velocity. Motion of the six degree-of-freedom(DOF) aircraft can be described by some nonlinear kinematical equations [10].

$$m\frac{d\mathbf{V}}{dt}=\mathbf{F},\quad \frac{d\mathbf{H}}{dt}=\mathbf{M} \tag{7}$$

where, $\mathbf{V}$, $\mathbf{H}$, $\mathbf{F}$, $\mathbf{M}$ are the aircraft' velocity, angular momentum, aggregate external force, total torque.

With the thrust, aerodynamics and gravity projected to the trajectory coordinate system, dinymical equations of the centroid of aircraft are (8).

$$\begin{aligned} m\frac{dV}{dt} &= P\cos\alpha\cos\beta - X - mg\sin\theta \\ mV\frac{d\theta}{dt} &= P(\sin\alpha\cos\gamma_v + \cos\alpha\sin\beta\sin\gamma_v) + Y\cos\gamma_v - Z\sin\gamma_v - mg\cos\theta \\ -mV\cos\theta\frac{d\psi_v}{dt} &= P(\sin\alpha\sin\gamma_v - \cos\alpha\sin\beta\cos\gamma_v) + Y\sin\gamma_v + Z\cos\gamma_v \end{aligned} \tag{8}$$

where, m, V, P, X, Y, Z are the aircraft' mass, velocity, thrust, resistance, lift and lateral force.

Dinymical equations of the rotation around the centroid of aircraft are described by (9).

$$\begin{aligned} I_x\frac{d\omega_x}{dt} + (I_z - I_y)\omega_y\omega_z &= M_x \\ I_y\frac{d\omega_y}{dt} + (I_x - I_z)\omega_x\omega_z &= M_y \\ I_z\frac{d\omega_z}{dt} + (I_y - I_z)\omega_x\omega_y &= M_z \end{aligned} \tag{9}$$

where, M_x, M_y, M_z are the torque components to the centroid. ω_x, ω_y, ω_z are three angular velocity components. I_x, I_y, I_z are inertia components to different coordinate axis.

The location (x, y, z) and pose $(\psi, \vartheta, \gamma)$ of the aircraft can be infered from the above equations.

$$\begin{aligned} \frac{dx}{dt} &= V\cos\theta\cos\psi_v \\ \frac{dy}{dt} &= V\sin\theta \\ \frac{dz}{dt} &= -V\cos\theta\sin\psi_v \end{aligned} \tag{10}$$

$$\begin{aligned}
\frac{d\vartheta}{dt} &= \omega_y \sin\gamma + \omega_z \cos\gamma \\
\frac{d\psi}{dt} &= \frac{1}{\cos\vartheta}(\omega_y \cos\gamma - \omega_z \sin\gamma) \\
\frac{d\gamma}{dt} &= \omega_x - \tan\vartheta(\omega_y \cos\gamma - \omega_z \sin\gamma)
\end{aligned} \tag{11}$$

3.2. Vision-aided Navigation based on UKF

Many filters are emploied to fuse different sensors in the integrated navigation system [11]. To improve the precision and stability of the vision aided navigation system, an unscented Kalman filter (UKF) is adopted to fuse the information of vision system and INS as the aircraft model is highly nonlinear [12]. The UKF [6] uses a deterministic sampling technique known as the unscented transform to pick a minimal set of sample points around the mean. These sigma points are then propagated through the non-linear functions, from which the mean and covariance of the estimate are then recovered. The unscented Kalman filter (UKF) has several advantages over other nonlinear filters. While maintaining the same order of magnitude in computational complexity, the UKF provides at least second-order accuracy compared to the first-order accuracy of the extended Kalman Filter (EKF). UKF does not necessitate the computation of the messy Jacobian matrices, which is a difficult task for EKF. In addition, UKF has the computational complexity advantage over the particle filter (PF).

The general nonlinear discrete process model and measurement model of vision-aided navigation system for autonomous aircraft are given by:

$$\begin{aligned}
x_{k+1} &= f(x_k, u_k, w_k) \\
y_k &= h(x_k, v_k)
\end{aligned} \tag{12}$$

where, x_k is the state vector of aircraft location, pose and velocity. y_k is the observation vector. u_k is the control vector. w_k and v_k are process noise and measurement noise respectively.

The initial conditions and error variance matrix are known:

$$\begin{aligned}
\hat{x}_0 &= E[x_0] \\
P_0 &= E\left[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T\right]
\end{aligned} \tag{13}$$

There are two important processes of time update and measurement update to apply the UKF. The construction and spread of sigma points are determined by:

$$\begin{aligned}
\chi^a_{k-1} &= \left[\hat{\chi}^a_{k-1} \quad \hat{\chi}^a_{k-1} + \sqrt{(n+\lambda)P^a_{k-1}} \quad \hat{\chi}^a_{k-1} - \sqrt{(n+\lambda)P^a_{k-1}}\right] \\
\chi^a_{k|k-1} &= f(\chi^a_{k-1}) \\
\hat{x}^-_k &= \sum_{i=0}^{2n} W_i^{(m)} \chi^a_{k|k-1} \\
P^-_k &= \sum_{i=0}^{2n} W_i^{(c)} (\chi^a_{k|k-1} - \hat{x}^-_k)(\chi^a_{k|k-1} - \hat{x}^-_k)^T
\end{aligned} \tag{14}$$

Time-update equations of measurement model are:

$$\begin{aligned}
\mathrm{y}_{\mathrm{k|k\text{-}1}} &= h(\chi^a_{k|k-1}) \\
\hat{\mathrm{y}}^-_k &= \sum_{i=1}^{2n} W_i^{(m)} \mathrm{y}_{\mathrm{k|k\text{-}1}}
\end{aligned} \tag{15}$$

The measurement-update equations are given by:

$$
\begin{aligned}
\mathrm{P}_{\mathrm{y}} &= \sum_{i=0}^{2n} W_i^{(c)} (\mathrm{y}_{k|k\text{-}1} - \hat{\mathrm{y}}_k^-)(\mathrm{y}_{k|k\text{-}1} - \hat{\mathrm{y}}_k^-)^T \\
\mathrm{P}_{\mathrm{xy}} &= \sum_{i=0}^{2n} W_i^{(c)} (\chi_{k|k-1}^a - \hat{x}_k^-)(\mathrm{y}_{k|k\text{-}1} - \hat{\mathrm{y}}_k^-)^T \\
\hat{x}_k &= \hat{x}_k^- + K_k (y_k - \hat{y}_k^-) \\
K_k &= P_{xy} P_y^{-1} \\
P_k &= P_k^- - K_k P_y K_k^T
\end{aligned}
\tag{16}
$$

4. Simulation Results

The proposed methods are tested on real image sequences and simulated scenes. Salient features detection and matching results of indoor robot navigation sequence are show in Figure 3. The “Corridor” sequence is from the visual geometry group of Oxford University. Twenty most salient corners based on CPDA are detected on each frame using the proposed saliency descriptor. The corresponding features on the adjacent frames are matched using Fourier-Mellin transform. The proposed method has shown better performance on both match accuracy and computational complexity than the commonly used SIFT descriptor.

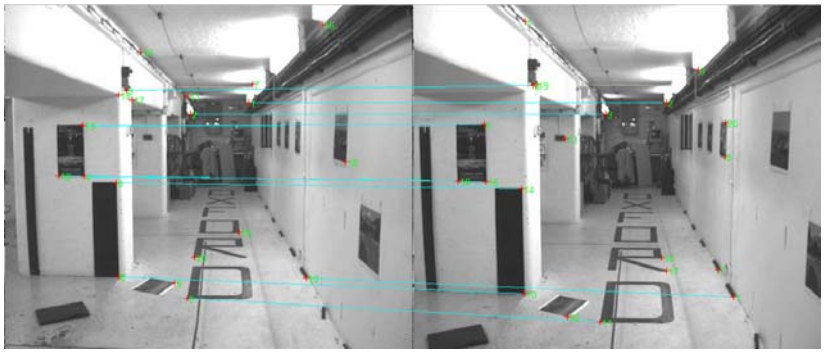


Figure 3. Results of Salient Features Detection and Matching

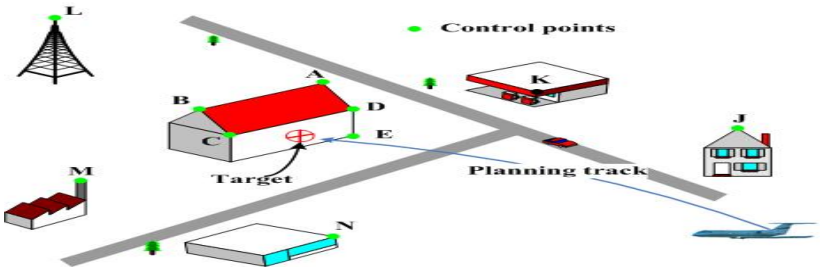
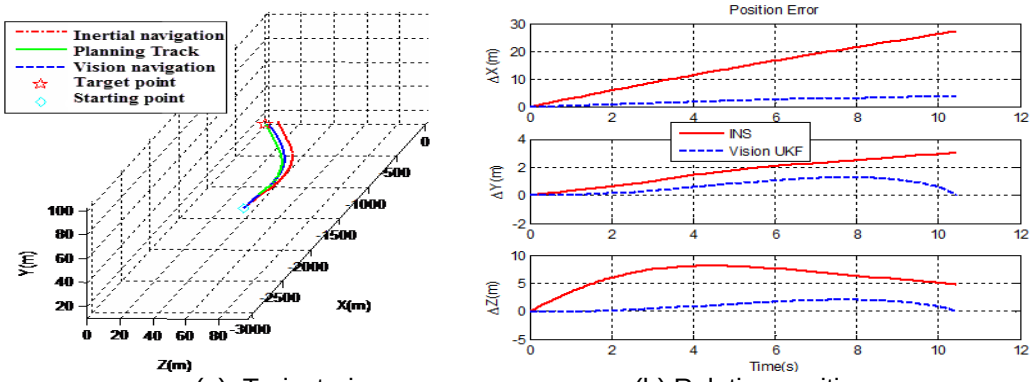


Figure 4. Simulation Scene With Ten Control Points

The model of the aircraft and vision-aided navigation system are built based on Matlab/Simulink. Figure 4 shows the simulation scene of vision-aided navigation for autonomous aircraft. The aircraft is planned to fly toward a house from 3Km away. Ten salient features are selected as the control points for pose estimation. Figure 5 depicts the comparison results of aircraft trajectories with different navigation systems. The trajectories include planning trajectory, trajectory with inertial system, and trajectory with vision aided navigation system based on UKF. Figure 6 shows the relative postion error results of INS and vision aided system. Aircraft track with traditional navigation system tends deviating from the planned track with the

accumulative error of inertial system. The track with vision-aided navigation system is more close to the planned track. After 300 Monte Carlo simulations, the statistical distance errors of inertial navigation and vision navigation are 14.0m with 0.4m respectively.



(a) Trajectories (b) Relative position error

Figure 5. Comparison Result of Aircraft

5. Conclusion

Salient features are helpful to improve the environmental applicability of vision navigation system. We present a feature saliency descriptor to choose control points for autonomous aircraft visual navigation. The location and pose of aircraft can be estimated with the corresponding control points. The unscented Kalman filter is used to fuse the state information and eliminate the accumulated error of the inertial navigation system. Experimental results show that the accuracy and reliability of the autonomous aircraft navigation system are significantly increased using the proposed method. In future work, we will optimize the algorithm complexity to use the vision-aided navigation system in real-time enviroment.

References

[1] Kaiser M, Gans N, Dixon W. Vision-Based Estimation for Guidance, Navigation, and Control of an Aerial Vehicle. *IEEE Transactions on Aerospace and Electronic Systems*. 2010; 46(3): 1064-1077.
[2] Mourikis A, Trawny N, Roumeliotis S, et al. Vision-Aided Inertial Navigation for Spacecraft Entry, Descent, and Landing. *IEEE Transactions on Robotics*, 2009; 25(2): 264-280.
[3] Mikolajczyk K, Schmid C. A Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005; 27(10): 1615-1630.
[4] Kadir T, Brady M. Scale, Saliency and Image Description. *International Journal of Computer Vision*. 2001; 45(2): 83-105.
[5] Chen QS, Defrise M, Deconinck F. Symmetric Phase-only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition. *IEEE Transactions Pattern Analysis and Machine Intelligence*. 1994; 16(12): 1156-1168.
[6] Van der Merwe R, Wan EA, Julier S. *Sigma-point Kalman Filters for Nonlinear Estimation and Sensor-fusion Application to Integrated Navigation*. AIAA Guidance, Navigation & Control Conference, 2004.
[7] He XC, Yung C. *Curvature Scale Space Corner Detector with Adaptive Threshold and Dynamic Region of Support*. Proceedings of the 17th International Conference on Pattern Recognition, 2004; 2: 791-794.
[8] Awrangjeb M, Lu G. Robust Image Corner Detection Based on the Chord-to-point Distance Accumulation Technique. *IEEE Transaction on Multimedia*. 2008; 10(6): 1059-1072.
[9] David P, DeMenthon D, Duraiswami R, et al. SoftPOSIT: Simultaneous Pose and Correspondence Determination. *International Journal of Computer Vision*. 2004; 59(3): 259-284.
[10] Sonneveldt L. *Nonlinear F-16 Model Description*. Technical report, Delft University of Technology. The Netherlands. 2006.
[11] Yanran Wang, Hai Zhang, Qifan Zhou. Adaptive Integrated Navigation Filtering based on Accelerometer Calibration. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(7): 1869-1878.
[12] Oh SM, Johnson EN. *Relative Motion Estimation for Vision-based Formation Flight using Unscented Kalman Filter*. AIAA Guidance, Navigation & Control Conference and Exhibit. AIAA 2007; 2007-6866.