Journal of Robotics, Networking and Artificial Life, Vol. 3, No. 2 (September 2016), 102-106

Attitude/Position Estimation of Rigid-Body using Inertial and Vision Sensors

Shihao Sun
The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA)
Beijing, 100191, China

Yingmin Jia
The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA)
Beijing, 100191, China
E-mail: jxcrssh@126.com, ymjia@buaa.edu.cn
www.buaa.edu.cn

Abstract

This paper is concerned with the attitude/position estimation of a rigid-body using inertial and vision sensors. By employing the Newton-Euler method, a kinematic model is developed for the rigid-body by treating the inertial measurements as inputs. Based on the coordinate transformation, a nonlinear visual observation model is proposed by using the image coordinates of feature points as observations. Then the Extended Kalman filter (EKF) is utilized to estimate the attitude/position recursively. The effectiveness of the proposed algorithm is evaluated by simulation.

Keywords: attitude and position estimation, EKF, inertial sensor, vision sensor.

1. Introduction

Accurate attitude/position estimation of rigid-body has received considerable attention in the past decades[1-8]. This is partly due to the fact that it is generally required in many typical applications such as estimating the motion of a robot end-effector[9], navigation for small unmanned aerial vehicle[10], spacecraft relative navigation in rendezvous[11] etc. The widely used way of obtaining position and orientation is using an inertial measurement unit (IMU) as the navigation sensor, which consists of a tri-axis gyroscope and a tri-axis accelerometer. However, integration over a long time period may lead to unbounded estimation errors if noises, offsets, scale errors and uncertainty in navigation model are present. Using vision as a standalone sensor for attitude/position estimation is also quite a standard way[12], because of the ability to sense the actual attitude/position without accumulative errors. However, vision sensors can only sense the actual position but not the velocity and accelerate. Therefore, an underlying dynamic model for the motion of rigid-body is needed for accurate estimation when we use vision sensor alone.

Motivated by the discussions above, the combination of the vision and inertial sensors has been recognized as a promising choice for accurate attitude and position estimation. For example, rigid body pose estimation using inertial sensors and a monocular camera is considered in Ref. 13, and it is shown how rotation estimation can be decoupled from position estimation. In Ref. 14, Chen studied the problem of the pose estimation of robotic end-effector with inertial and $SE(3)$ measure−ments. Among these literature, the measurements of vision sensor are the actual attitude and position that can be obtained by using machine vision algorithms like PNP, stereo-vision. Although the

Published by Atlantis Press

observation model is linear in such way, the expression of observation noises is complicated according to the complicated machine vision algorithms and costs huge computations. An alternative way is applying the image coordinates of feature points as observations directly.

In this paper, we attempt to estimate the orientation of a rigid-body using inertial and vision sensors. The inertial measurements are treated as inputs for developing the kinematic model, and a novel non-linear observation model is proposed by using the image coordinates of feature points. The EKF is utilized to address the nonlinear filtering problem. Simulation results are provided to evaluate the performance of the proposed algorithm.

2. Problem Statement

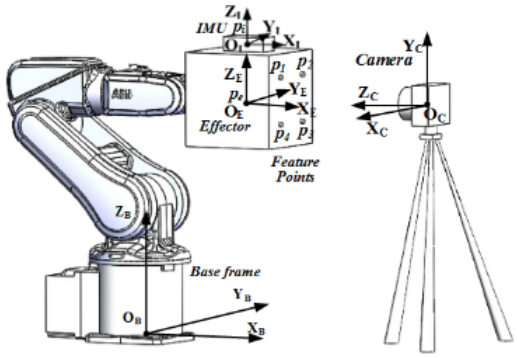

Fig. 1. System configuration and coordinate frames

In this paper, we consider the case in which the motion of a rigid-body (noted as effector) is controlled by the ABB-120 robot as shown in Fig.1. The effector is equipped with a strap-down IMU and four infrared LED feature points. The vision sensor is a camera equipped with an infrared filter, which is fixed at a certain location in the base frame.

2.1. *Effector Kinematics*

As shown in Fig. 1. Denote ${}^{B}\boldsymbol{p}_e \in R^3$ as the effector position in the base frame, evolving as:

$$ {}^{B}\dot{\boldsymbol{p}}_e = {}^{B}\boldsymbol{v}_e \quad {}^{B}\dot{\boldsymbol{v}}_e = {}^{B}\boldsymbol{a}_e \tag{1} $$

where ${}^{B}\boldsymbol{v}_e \in R^3$ and ${}^{B}\boldsymbol{a}_e \in R^3$ are referred as the translational velocity and acceleration of SCP in the base frame, respectively.

Denote Euler (3-2-1) rotation angles $\boldsymbol{\Phi} = [\alpha \quad \beta \quad \gamma]^{\mathrm{T}}$ as the effector attitude observed in the base frame, according to attitude kinematics:

$$ \dot{\boldsymbol{\Phi}} = \boldsymbol{T}(\boldsymbol{\Phi})\,{}^{E}\boldsymbol{\omega}_e \tag{2} $$

where ${}^{E}\boldsymbol{\omega}_e \in R^3$ is the angular velocity of the effector as viewed in the effector frame and

$$ \boldsymbol{T}(\boldsymbol{\Phi}) = \begin{bmatrix} 0 & \sin\gamma/\cos\beta & \cos\gamma/\cos\beta \\ 0 & \cos\gamma & -\sin\gamma \\ 1 & \sin\gamma\tan\beta & \cos\gamma\tan\beta \end{bmatrix} \tag{3} $$

Discretization of equations (1) and (2) is needed to design a Kalman Filter. Assume that ${}^{B}\boldsymbol{a}_e$ and ${}^{E}\boldsymbol{\omega}_e$ are constants in each sampling interval $[kT_s,(k+1)T_s]$ with sampling period T_s.

The state transition can be approximated by:

$$ \begin{aligned} {}^{B}\boldsymbol{p}_e(k+1) &= {}^{B}\boldsymbol{p}_e(k) + T_s\,{}^{B}\boldsymbol{v}_e(k) + \frac{T_s^2}{2}\,{}^{B}\boldsymbol{a}_e(k) \\ {}^{B}\boldsymbol{v}_e(k+1) &= {}^{B}\boldsymbol{v}_e(k) + T_s\,{}^{B}\boldsymbol{a}_e(k) \\ \boldsymbol{\Phi}(k+1) &= \boldsymbol{\Phi}(k) + T_s\boldsymbol{T}(\boldsymbol{\Phi}(k))\,{}^{E}\boldsymbol{\omega}_e(k) \end{aligned} \tag{4} $$

2.2. *Transition models using IMU measurements*

The IMU in Fig. 1 consists of a tri-axis gyroscope and a tri-axis accelerometer that output:

$$ \begin{aligned} \bar{\boldsymbol{\omega}}_{imu} &= {}^{I}\boldsymbol{\omega}_i + {}^{I}\boldsymbol{m}_\omega \\ \bar{\boldsymbol{a}}_{imu} &= {}^{I}\boldsymbol{a}_i + {}^{I}\boldsymbol{g} + {}^{I}\boldsymbol{m}_a \end{aligned} \tag{5} $$

where $\bar{\boldsymbol{\omega}}_{imu}$ and $\bar{\boldsymbol{a}}_{imu}$ are IMU measurements, $\boldsymbol{m}_\omega$ and $\boldsymbol{m}_a$ are measurement noises, and $\boldsymbol{g}$ is the gravity vector.

According to the rigid kinematic theorem, the angular velocity and acceleration relations between point $\boldsymbol{p}_i$ and $\boldsymbol{p}_e$ are as follows:

$$ \begin{aligned} {}^{E}\boldsymbol{\omega}_e &= {}^{E}\boldsymbol{R}_I\,{}^{I}\boldsymbol{\omega}_i \\ {}^{B}\boldsymbol{a}_e &= {}^{B}\boldsymbol{R}_E\,{}^{E}\boldsymbol{R}_I\left({}^{I}\boldsymbol{a}_i + \boldsymbol{S}({}^{I}\boldsymbol{\alpha}_i)\,{}^{I}\boldsymbol{r}_{ei} + \boldsymbol{S}({}^{I}\boldsymbol{\omega}_i)\boldsymbol{S}({}^{I}\boldsymbol{\omega}_i)\,{}^{I}\boldsymbol{r}_{ei}\right) \end{aligned} \tag{6} $$

where ${}^{E}\boldsymbol{R}_I$ is a constant orientation matrix referred as the IMU frame attitude observed in the effector frame; ${}^{I}\boldsymbol{r}_{ei}$ is the position vector of $\boldsymbol{p}_e$ relative to $\boldsymbol{p}_i$ in IMU frame; ${}^{I}\boldsymbol{\alpha}_i$ is the angular acceleration of the IMU, according to the assumption that ${}^{E}\boldsymbol{\omega}_e$ is constant in each sampling interval, then ${}^{I}\boldsymbol{\alpha}_i(k) = \boldsymbol{0}$; ${}^{B}\boldsymbol{R}_E$ is the orientation matrix referred as the end-effector attitude observed in the base frame, which is given by

$$ {}^{B}\boldsymbol{R}_E = \begin{bmatrix} \cos\alpha\cos\beta & -\sin\alpha\cos\gamma+\cos\alpha\sin\beta\sin\gamma & \sin\alpha\sin\gamma+\cos\alpha\sin\beta\cos\gamma \\ \sin\alpha\cos\beta & \cos\alpha\cos\gamma+\sin\alpha\sin\beta\sin\gamma & -\cos\alpha\sin\gamma+\sin\alpha\sin\beta\cos\gamma \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma \end{bmatrix}; $$

and $\boldsymbol{S}(\bullet)$ is the cross product operator transforms a vector $\boldsymbol{c} = [c_1 \quad c_2 \quad c_3]^{\mathrm{T}}$ to a skew-symmetric matrix

$$ \boldsymbol{S}(\boldsymbol{c}) = \begin{bmatrix} 0 & -c_3 & c_2 \\ c_3 & 0 & -c_1 \\ -c_2 & c_1 & 0 \end{bmatrix}. $$

Given (5) and (6) at the sampling interval, we have

$$
\begin{aligned}
{}^{E}\boldsymbol{\omega}_e(k) &= {}^{E}\boldsymbol{R}_I\,\bar{\boldsymbol{\omega}}_{imu}(k) - {}^{E}\boldsymbol{R}_I\,{}^{I}\boldsymbol{m}_{\omega}(k)\\
{}^{B}\boldsymbol{a}_e(k) &= {}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\,\bar{\boldsymbol{a}}_{imu}(k) + {}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\,{}^{I}\boldsymbol{r}_{ei} - {}^{B}\boldsymbol{g}\\
&\quad - {}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\Big(\boldsymbol{S}\big(\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\,{}^{I}\boldsymbol{r}_{ei}\big) + \boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\boldsymbol{S}\big({}^{I}\boldsymbol{r}_{ei}\big)\Big)\,{}^{I}\boldsymbol{m}_{\omega}(k)\\
&\quad + {}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\boldsymbol{S}\big({}^{I}\boldsymbol{m}_{\omega}(k)\big)\boldsymbol{S}\big({}^{I}\boldsymbol{m}_{\omega}(k)\big)\,{}^{I}\boldsymbol{r}_{ei} - {}^{B}\boldsymbol{m}_a
\end{aligned}
\tag{7}
$$

Then the state transition equation (4) can be rewritten as

$$
\boldsymbol{X}(k+1) = \boldsymbol{\varphi}\big[\boldsymbol{X}(k),\boldsymbol{U}(k)\big] + \boldsymbol{\Gamma}\big[\boldsymbol{X}(k),\boldsymbol{U}(k)\big]\boldsymbol{W}(k) \tag{8}
$$

where $\boldsymbol{X}(k) = \big[{}^{B}\boldsymbol{p}_e(k)\ {}^{B}\boldsymbol{v}_e(k)\ \boldsymbol{\Phi}(k)\big]^{\mathrm{T}}$, $\boldsymbol{U}(k) = \big[\bar{\boldsymbol{\omega}}_{imu}(k)\ \bar{\boldsymbol{a}}_{imu}(k)\big]^{\mathrm{T}}$

$\boldsymbol{W}(k) = \big[{}^{I}\boldsymbol{m}_{\omega}(k)\ {}^{I}\boldsymbol{m}_{a}(k)\ \boldsymbol{S}\big({}^{I}\boldsymbol{m}_{\omega}(k)\big)\boldsymbol{S}\big({}^{I}\boldsymbol{m}_{\omega}(k)\big)\,{}^{I}\boldsymbol{r}_{ei}\big]^{\mathrm{T}}$,

$Cov\big\{\boldsymbol{W}(k),\boldsymbol{W}(j)\big\} = \boldsymbol{Q}_k\delta_{kj}$

$$
\boldsymbol{\varphi}\big[\boldsymbol{X}(k)\big] = \begin{bmatrix}
{}^{B}\boldsymbol{p}_e(k) + T_s\,{}^{B}\boldsymbol{v}_e(k) + \frac{T_s^2}{2}\Big({}^{B}\tilde{\boldsymbol{R}}_E(k)\,{}^{E}\boldsymbol{R}_I\big(\bar{\boldsymbol{a}}_{imu}(k) + \boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\,{}^{I}\boldsymbol{r}_{ei}\big) - {}^{B}\boldsymbol{g}\Big)\\
{}^{B}\boldsymbol{v}_e(k) + T_s\Big({}^{B}\tilde{\boldsymbol{R}}_E(k)\,{}^{E}\boldsymbol{R}_I\big(\bar{\boldsymbol{a}}_{imu}(k) + \boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\,{}^{I}\boldsymbol{r}_{ei}\big) - {}^{B}\boldsymbol{g}\Big)\\
\boldsymbol{\Phi}(k) + T_s\boldsymbol{T}\big(\boldsymbol{\Phi}(k)\big)\,{}^{E}\boldsymbol{R}_I\bar{\boldsymbol{\omega}}_{imu}(k)
\end{bmatrix}
$$

$$
\boldsymbol{\Gamma}\big[\boldsymbol{X}(k)\big] = \begin{bmatrix}
-\frac{T_s^2}{2}T_s^2\,{}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\Big(\boldsymbol{S}\big(\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\,{}^{I}\boldsymbol{r}_{ei}\big) + \boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\boldsymbol{S}\big({}^{I}\boldsymbol{r}_{ei}\big)\Big) & -\frac{T_s^2}{2} & \frac{T_s^2}{2}\,{}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\\
-T_s\,{}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\Big(\boldsymbol{S}\big(\boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\,{}^{I}\boldsymbol{r}_{ei}\big) + \boldsymbol{S}\big(\bar{\boldsymbol{\omega}}_{imu}(k)\big)\boldsymbol{S}\big({}^{I}\boldsymbol{r}_{ei}\big)\Big) & -T_s & T_s\,{}^{B}\boldsymbol{R}_E(k)\,{}^{E}\boldsymbol{R}_I\\
T_s\boldsymbol{T}\big(\boldsymbol{\Phi}(k)\big)\,{}^{E}\boldsymbol{R}_I & 0 & 0
\end{bmatrix}.
$$

2.3. *Observation models of Camera*

The rigid body in Fig. 1 contains four feature points noted as ${}^{E}\boldsymbol{p}_i = \big[{}^{E}p_{ix}\ {}^{E}p_{iy}\ {}^{E}p_{iz}\big]^{\mathrm{T}}$, $i = 1\cdots4$, and the relation between the fixed camera and the base frame are noted as the orientation matrix ${}^{C}\boldsymbol{R}_B$ and the translation vector ${}^{C}\boldsymbol{r}_{bc}$. By using the homogeneous transform matrices between the frames of Base, Effector and Camera, we obtain

$$
\begin{bmatrix} {}^{C}\boldsymbol{p}_i \\ 1 \end{bmatrix} = \begin{bmatrix} {}^{C}\boldsymbol{R}_B & {}^{C}\boldsymbol{r}_{bc} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} {}^{B}\boldsymbol{R}_E & {}^{B}\boldsymbol{p}_e \\ 0 & 1 \end{bmatrix}\begin{bmatrix} {}^{E}\boldsymbol{p}_i \\ 1 \end{bmatrix} \quad i = 1\cdots4 \tag{9}
$$

According to the camera pinhole model with the projective geometry, we obtain

$$
\begin{aligned}
u_i &= f_x\frac{{}^{C}p_{ix}}{{}^{C}p_{iz}} + u_0\\
v_i &= f_y\frac{{}^{C}p_{iy}}{{}^{C}p_{iz}} + v_0
\end{aligned}
\qquad i = 1\cdots4 \tag{10}
$$

where f_x and f_y are pixel magnification factors, (u_0,v_0) denotes the image coordinate of the camera's principal point, and (u_i,v_i) is the coordinate of the feature point $\boldsymbol{p}_i$ in the image plane. In view of (9) and (10), denote $\boldsymbol{Z}_i(k)$ as the image coordinate $\big(u_i(k),v_i(k)\big)$ at time k , then the observation model of camera can be written as

$$
\boldsymbol{Z}_i(k) = \boldsymbol{h}_i\big[\boldsymbol{X}(k),{}^{E}\boldsymbol{p}_i\big] + \boldsymbol{V}_i(\mathrm{k}) \quad i = 1\cdots4 \tag{11}
$$

where $\boldsymbol{h}_i$ is a 2-dim function derived by replacing ${}^{C}\boldsymbol{p}_i$ in equation (10) using equation (9), $\boldsymbol{V}(\mathrm{k})$ is the measurement noise and its covariance is $Cov\big\{\boldsymbol{V}_i(k),\boldsymbol{V}_i(j)\big\} = \boldsymbol{R}_{ki}\delta_{kj}$.

3. Estimation based on EKF

Since the state transition model (8) and the observation model (11) are both nonlinear, the EKF is used to update the effector motion state estimate $\hat{\boldsymbol{X}}(k|k)$ and its estimation error covariance matrix $\boldsymbol{P}(k|k)$.

The EKF is implemented as follows:

1) *Prediction*. Denote $\hat{\boldsymbol{X}}(k+1|k)$ $\boldsymbol{P}(k+1|k)$ $\hat{\boldsymbol{Z}}_i(k+1|k)$ as the one step predictions of estimation, covariance matrix and observation at the time $k+1$.

Then as to (8), they can be obtained

$$
\hat{\boldsymbol{X}}(k+1|k) = \boldsymbol{\varphi}\big[\hat{\boldsymbol{X}}(k|k),\boldsymbol{U}(k)\big] \tag{12}
$$

$$
\begin{aligned}
\boldsymbol{P}(k+1|k) &= \boldsymbol{\Psi}\big[k+1|k\big]\boldsymbol{P}(k|k)\boldsymbol{\Psi}^{\mathrm{T}}\big[k+1|k\big]\\
&\quad + \boldsymbol{\Gamma}\big[\hat{\boldsymbol{X}}(k|k),\boldsymbol{U}(k)\big]\boldsymbol{Q}_k\boldsymbol{\Gamma}^{\mathrm{T}}\big[\hat{\boldsymbol{X}}(k|k),\boldsymbol{U}(k)\big]
\end{aligned}
\tag{13}
$$

where $\boldsymbol{\Psi}\big[k+1|k\big] = \left.\dfrac{\partial\boldsymbol{\varphi}\big[\boldsymbol{X}(k),\boldsymbol{U}(k)\big]}{\partial\boldsymbol{X}(k)}\right|_{\boldsymbol{X}(k)=\hat{\boldsymbol{X}}(k|k)}$

$$
\hat{\boldsymbol{Z}}(k+1|k) = \boldsymbol{h}\big[\hat{\boldsymbol{X}}(k+1|k),{}^{E}\boldsymbol{p}_i\big] \tag{14}
$$

where $\hat{\boldsymbol{Z}} = \big[\hat{\boldsymbol{Z}}_1\ \hat{\boldsymbol{Z}}_2\ \hat{\boldsymbol{Z}}_3\ \hat{\boldsymbol{Z}}_4\big]^{\mathrm{T}}$ and $\boldsymbol{h} = \big[\boldsymbol{h}_1\ \boldsymbol{h}_2\ \boldsymbol{h}_3\ \boldsymbol{h}_4\big]^{\mathrm{T}}$ represent four feature points.

2) *Observation*. As the feature points are infrared LED and the camera is equipped with an infrared filter, we can easy obtain the coordinates $\boldsymbol{Z}_i(k+1)$ of the feature point $\boldsymbol{p}_i$ in the image plane at the time $k+1$.

3) *Update*. When the new image is obtained at time $k+1$, the filter can be computed as

$$
\boldsymbol{K}(k+1) = \boldsymbol{P}(k+1|k)\boldsymbol{H}^{\mathrm{T}}(k+1)\boldsymbol{O}(k+1) \tag{15}
$$

where $\boldsymbol{O}(k+1) = \Big[\boldsymbol{H}(k+1)\boldsymbol{P}(k+1|k)\boldsymbol{H}^{\mathrm{T}}(k+1) + diag\big[\boldsymbol{R}_{ki}\big]_{i=1\cdots4}\Big]^{-1}$

$$
\boldsymbol{H}(k+1) = \left.\frac{\partial\boldsymbol{h}\big[\boldsymbol{X}(k+1),{}^{E}\boldsymbol{p}_i\big]}{\partial\boldsymbol{X}(k+1)}\right|_{\boldsymbol{X}(k+1)=\hat{\boldsymbol{X}}(k+1|k)}.
$$

Then the state estimate $\hat{\boldsymbol{X}}(k+1|k+1)$ and its estimation error covariance matrix $\boldsymbol{P}(k+1|k+1)$ at time $k+1$ can be computed by

$$
\hat{\boldsymbol{X}}(k+1|k+1) = \hat{\boldsymbol{X}}(k+1|k) + \boldsymbol{K}(k+1)\big[\boldsymbol{Z}(k+1) - \hat{\boldsymbol{Z}}(k+1|k)\big] \tag{16}
$$

$$
\boldsymbol{P}(k+1|k+1) = \big[\boldsymbol{I} - \boldsymbol{K}(k+1)\boldsymbol{H}(k+1)\big]\boldsymbol{P}(k+1|k) \tag{17}
$$

Shihao Sun, Yingmin Jia

4. Simulation results

As stated in section 2, the effector is controlled by the robot, in which the revolute joints are programmed as $\boldsymbol{q} = \frac{\pi}{180}[40 \quad 10 \quad -5 \quad 60 \quad 8 \quad 10]\cos\left(\frac{t}{100}\right)(rad) \quad t \in [0, 100\pi]$. And the robot forward kinematics with nominal D-H parameters is used to provide the attitude /position of the effector relative to base frame, which serve as the real. The sampling period of IMU and camera are both set as $0.1s$, ignored the different sampling rate of the sensors. The noises of the IMU and camera sensors are assumed as $\boldsymbol{m}_{\omega} \sim \boldsymbol{N}(0, 0.1^2)$ $\boldsymbol{m}_{a} \sim \boldsymbol{N}(0, 2.5^2)$ $\boldsymbol{V} \sim \boldsymbol{N}(0, 5^2)$. The plots in Fig.2 and Fig.3 show the estimate results based on EKF and transition models only use IMU, respectively. The simulation results suggest the effectiveness of the proposed method.

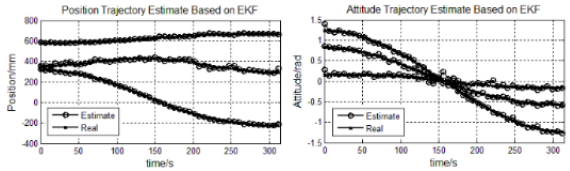

Fig. 2.Estimate of Effector Motion Based on EKF

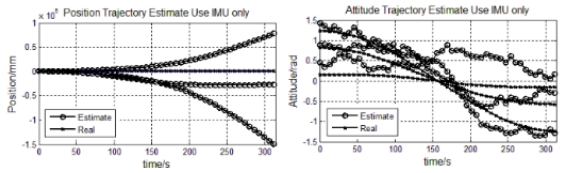

Fig. 3. Estimate of Effector Motion Use IMU Only

Notice from (7) that the IMU's acceleration measurement is affected by the effector acceleration and the gravity, and the effector acceleration is much smaller than gravity in this simulation. Thus, it is very sensitive to noise, and lead to the estimate errors of position use IMU only are very large as shown in Fig.3.

The measure model $\boldsymbol{Z}(k) = \boldsymbol{X}(k) + \boldsymbol{V}(k)$, which can be obtained by using the PNP algorithm (refer to solvepnp function in OpenCV), is also applied in EKF to compare against the proposed observation models noted as Pixels Measure. To illustrate the performance of the observation models, the root mean square error (RMSE) in position and attitude are shown and the simulation results are derived from 100 Monte Carlo runs. The RMSE in position and attitude are shown in Fig. 4. The simulation results suggest that the performance of the EKF can be improved by using the image coordinates (Pixels) of feature points as observations.

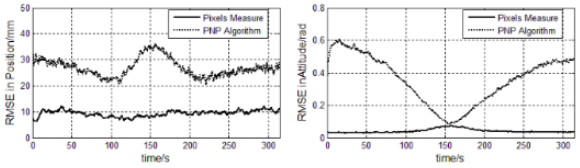

Fig. 4. The RMSE in position and attitude

Because of the complex computation, the measure- ment noises covariance matrices in measurement model obtained by PNP algorithm are not calculated based on the pixels noises of the camera, it is set based on our experience in this simulation. Besides, the errors of the attitude measure are shown in Fig.5. It's clear from the figure that the error is non-Gaussian.

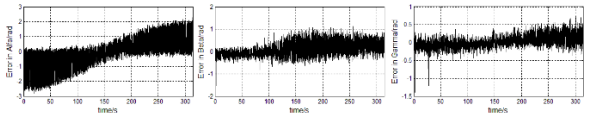

Fig. 5. The errors of the attitude measure

5. Conclusion

This paper investigated the attitude/position estimation of a rigid-body by using a measurement system consist of an inertial sensor and a vision sensor, and an EKF is applied to fuse these measurements. The motion states are propagated in time using the inertial measurements processed through the Newton-Euler equations, and the pixels coordinates of feature points are used as observation directly, which is different from using position and attitude coordinates in the camera frame. Simulation results suggest that the performance of the EKF can be improved by using the pixels coordinates of feature points as observations than calculating position and attitude coordinates through PNP algorithm. Future work would focus on the on-line tracking estimation experiments.

Acknowledgements

This work was supported by the National Basic Research Program of China (973 Program: 2012CB821200, 2012CB821201) and the NSFC (61134005, 61327807, 61521091, 61520106010).

References

1. L Chen, and Y Jia, Variable-poled Tracking Control of a Two-wheeled Mobile Robot Using Differential Flatness, *Journal of Robotics, Networking & Artificial Life*, 1(1) (2014), 12-16.
2. J He, and Y Jia, Adaptive Sliding Mode Control for Magnetic Levitation Vehicles, *Journal of Robotics, Networking & Artificial Life*, 1(2) (2014)169-173.
3. X Lu, and Y Jia, Attitude Reorientation of Spacecraft with Attitude Forbidden Zones, *Journal of Robotics, Networking & Artificial Life*, 2(1) (2015)13-16.
4. C Yang, and Y Jia, Adaptive Multiple-Model Control of A Class of Nonlinear Systems, *Journal of Robotics, Networking & Artificial Life*, 2(2) (2015)69-72.
5. M Duan, and Y Jia, Adaptive Sliding Mode Control for A 2 DOF Magnetic Levitation System with Uncertain Parameters, *Journal of Robotics, Networking & Artificial Life*, 2(4) (2016)263-267
6. Y Jia, Robust Control with Decoupling Performance for Steering and Traction of 4WS Vehicles under Velocity-Varying Motion, *IEEE Transactions on Control Systems Technology*, 8(3)(2000)554-569.
7. Y Jia, Alternative Proofs for Improved LMI Representations for the Analysis and the Design of Continuous-Time Systems with Polytopic Type Uncertainty: A Predictive Approach, *IEEE Transactions on Automatic Control*, 48(8)(2003) 1413-1416.
8. Y Jia, General Solution to Diagonal Model Matching Control of Multi‐Output-Delay Systems and Its Applications in Adaptive Scheme, *Progress in Natural Science*, 19(1)(2009) 79-90.
9. Wang C, Chen W, Tomizuka M., Robot end-effector sensing with position sensitive detector and inertial sensors, in *Proc. IEEE Int. Conf. Robotics and Automation.* (Minnesota, USA, 2012), 5252-5257.
10. Williamson, Walton R., et al, Sensor Fusion Applied to Autonomous Aerial Refueling, *J. Guid Control Dynam.* **32**(1) (2009) 262-275.
11. Zhang G, M. Kontitsis, et al, Cooperative Relative Naviga-tion for Space Rendezvous and Proximity Operations using Controlled Active Vision, *J. Field Robot.* **00**(0) (2015) 1–24.
12. D. B. Gennery. Visual Tracking of Known Three Dimensional Objects, *Int J Comput Vision,* **7**(3) (1992), 243-270.
13. H Rehbinder, B.K. Ghosh, Pose estimation using line based dynamic vision and inertial sensors. *IEEE T Automat Contr*, **48**(2) (2003) 186 - 199.
14. Chen X, et al. Pose estimation of robotic end-effectors under low speed motion using EKF with inertial and SE(3) measurements, in *Proc. IEEE Int. Conf. Advanced Intellig-ent Mechatronics*,(Busan, Korea 2015) 1585-1590.