20th International Conference on Information Fusion
Xi'an, China - July 10-13, 2017

Fusion of Visual and Inertial Measurement Information for Unmanned Aerial Vehicles Autonomous Navigation in Unknown Environment

Li Jianguo, Liu Qiang, Liu Qiubing, Zhao Pengjiao
North Automatic Control Technology Institute
NO.351. Tiyu Road, Xiao Dian District, Taiyuan, China
E-mail: yxhewh@163.com

***Abstract*—An autonomous navigation scheme for unmanned aerial vehicles is presented based on visual and inertial measurement information fusion without the known ground cooperative target. The UAV relative translation and rotation motion parameters are estimated by inter-frame image feature detection and tracking. Then the relative motion parameters are considered to be the relative pose measurements of two consecutive images and the observation models explicitly related to the state are provided. In particular, the reduced-order state augmentation algorithm is used to account for the dependence of the measurements on the state at two different measurement time instants. The optimal fusion of visual and inertial measurement information is performed using the framework of the extended Kalman filter. Finally the feasibility and validity of the proposed navigation algorithm are demonstrated by mathematical simulation.**

Keywords—UAV; visual navigation; information fusion; autonomous navigation; features tracking

I. INTRODUCTION

In the last few years, unmanned aerial vehicles (UAVs) have attracted much attention in military and civilian applications with the development of the on-board processing capabilities [1-3]. Precise autonomous navigation capability is the necessary prerequisite to complete various missions for UAVs. Current navigation systems are mainly based on the dead reckoning from the Inertial Measurement Unit (IMU), however accumulated errors over time and uncertainties in initialization would result in large inertial navigation errors. The Global Position System (GPS) can be used as auxiliary navigation aids to fuse with inertial sensors, but GPS signal may be lost in complicated environments.

The vision sensors can obtain wealthy body motion information by external environment perception, and gather reconnaissance data as the onboard payload. As the development of the vision sensors and image processing technologies, autonomous visual navigation methods based on line-of-sight vector measurement have been successfully applied to many UAV navigation missions. However pure vision navigation algorithm has certain limitations including the requirement of at least three landmarks features at the same time and the low-rate pose estimation due to the time-consuming image processing.

As the vision sensors and IMU are complementary in terms of measurement information and update rate, the fusion of vision and inertial navigation can provide more precise and real-time navigation solution. According to whether prior knowledge of the landmark feature is available or not, vision-aided inertial navigation algorithms fall into two main categories. When the 3D positions of the landmark features are known, the navigation state estimation can be obtained by matching and tracking the features observed by the camera onboard with the landmarks in the map [4]. As the positions of the landmarks should be built prior to the mission, the navigation performance is bound to the knowledge of environment map. When the prior knowledge of the landmarks is not required, two types of algorithms are utilized in navigation: 1) fusion algorithm based on features epipolar constraint [5], and (2) fusion algorithm based on simultaneous localization and mapping (SLAM) [6-8]. Although SLAM can estimate the features positions and camera motion parameters simultaneously in unknown environment, the computation and storage complexity increase as the estimate proceed. The constraint-based fusion estimates the relative motion parameters using geometry constraint relationship from successive images, which eventually are utilized for correcting the inertial navigation results. The low computation complexity makes it more suitable for UVA real-time navigation application.

The relative motion parameters can be converted to different forms to update the inertial navigation. Hoffman et al. [9] transforms the relative motion estimate to the absolute position pseudo-measurement using the state prediction estimates from inertial navigation, however the attitude is assumed to be precisely known. In [10], the pairwise epipolar constraint is employed in conjunction with the dynamical model of aircraft to estimate the navigation state in an Implicit Extended Kalman Filter (IEKF) framework, where the white noise assumption may be invalid for the system process noise and observation noise. In [11], the relative pose between two images is regarded as the average velocity measurement to improve the navigation accuracy. Indelman et al. [12] investigates a new vision-aided navigation algorithm based on three-view geometry constraint by fusing features from three

This work is supported by the National Natural Science Foundation of China (Grant no. 61603006)

978-0-9964-5270-0©2017 ISIF

overlapping images with the concomitant navigation data. Compared with the inter-frame motion estimation method, the proposed three-view geometry method can efficiently improve the navigation observability and estimation precision. As the measurement update rate is decreased, the real-time performance which is necessary for UAV operation may not be maintained.

In order to improve the navigation capability in unknown environment and limited observation condition, in this paper we present a novel algorithm to fuse vision and IMU measurement information. The inter-frame motion estimation is considered to be the relative pose measurements, which can avoid the poorly conditioned problem introduced by the homography matrices decomposition. Systematic numerical simulations illustrate the optimal fusion of measurement information and high-precision autonomous navigation ability.

II. RELATIVE MOTION ESTIMATION ALGORITHM

Extracting and matching the feature points are the prerequisite work to finish for relative motion estimation, which directly relates to estimation precision of navigation algorithms. Harris feature point extracting algorithm and Kanade-Lucas-Tomasi feature point tracking algorithms [13] are only suitable for the case where inter-frame relative motion changes small. Scale Invariant Feature Transform (SIFT) algorithm [14], which has strong robustness to scale and illumination changes, is suitable to track feature points between the consecutive images frames. In order to eliminate mismatching, Random Sample Consensus (RANSAC) algorithm [15] is used to eliminate outliers and generate matched pairwise accurately.

Traditional relative motion estimation takes use of two consecutive images measured by optical cameras to solve homography matrix which is then decomposed to obtain the relative position and attitude estimates. However the homography decomposition is a poorly conditioned problem. To obtain correct relative motion estimates, additional approach need be used to eliminate ambiguous solutions. In this paper relative motion parameters are computed directly, which can efficiently reduce calculation complexity. The coordinate systems involved include the camera coordinate system (C), the UAV body coordinate system (B) and the navigation coordinate system (N). It is assumed that the measurement period of camera is t_m , two consecutive camera sampling instances are t_k and t_{k+m} respectively, the corresponding coordinate systems are C_1 and C_2 respectively. Inter-frame rotation and translation motion between consecutive sampling are respectively given by $\bar{\boldsymbol{q}}_2^1$ and $\boldsymbol{p}_2^1$, feature points coordinates in two focal planes are denoted by $\boldsymbol{u}_i = (u_i, v_i)^{\mathrm{T}}$ and $\boldsymbol{u}_i' = (u_i', v_i')^{\mathrm{T}}$ respectively. Given the depth z_i between feature point and image plane and precisely calibrated internal parameters of camera, the line-of-sight measurement of the feature point in the frame C_1 can be given by

$$\boldsymbol{X}_i = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} u_i z_i \\ v_i z_i \\ z_i \end{bmatrix} \tag{1}$$

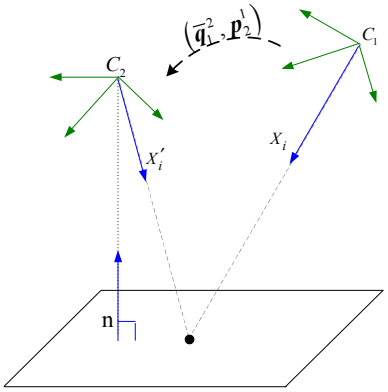

Fig.1 Sketch of relative motion between the frames

According to the inter-frame relative motion relationship shown in Fig.1 [16], the line-of-sight vector measurement of the feature point in the frame C_2 follows

$$\boldsymbol{X}_i' = \begin{bmatrix} x_i' & y_i' & z_i' \end{bmatrix}^{\mathrm{T}} = A\left(\bar{\boldsymbol{q}}_1^2\right)\boldsymbol{X}_i + \boldsymbol{p}_1^2 \tag{2}$$

Then the position of the feature point at t_k expressed in the camera coordinate system at t_{k+m} can be written as

$$\boldsymbol{f}_i\left(\boldsymbol{u}_i, \bar{\boldsymbol{q}}_1^2, \boldsymbol{p}_1^2\right) = \begin{bmatrix} \dfrac{x_i'}{z_i'} & \dfrac{y_i'}{z_i'} \end{bmatrix}^{\mathrm{T}} \tag{3}$$

By minimizing the position errors of feature point between measurement and projection in image planes, relative motion parameters can be solved. Therefore the cost function can be written as follow

$$J\left(\bar{\boldsymbol{q}}_1^2, \boldsymbol{p}_1^2\right) = \sum_i \left\| \boldsymbol{u}_i' - \boldsymbol{f}_i \right\|^2 \tag{4}$$

In view of the unit norm constraint of the attitude quaternion, by introducing the lagrange multiplier λ the cost function gives

$$J'\left(\bar{\boldsymbol{q}}_1^2, \boldsymbol{p}_1^2, \lambda\right) = \sum_i \left\| \boldsymbol{u}_i - \boldsymbol{f}_i \right\|^2 + \lambda\left(\bar{\boldsymbol{q}}^{\mathrm{T}}\bar{\boldsymbol{q}} - 1\right) \tag{5}$$

In consideration of nonlinear least square problem above, relative motion parameters can be solved by Levenberg-Marquardt (LM) algorithm [17]. In order to improve LM algorithm's convergence speed, relative displacements and initial attitude estimates are provided by inertial navigation algorithms. Finally camera's depth information is iterated and updated by relative motion parameter estimates from LM algorithm.

The drawback of optical navigation algorithm is that the scale of image can not be obtained. In other words, the camera is unable to distinguish large targets in remote range and small targets in close range. Thus, 5 degrees of freedom direction of translation and relative attitude need to be combined with information measured by an altimeter to recover relative

displacement. Assuming that the altimeter is installed along camera's optical axis and the altimeter measurements are l_1 and l_2 at t_1 and t_2 respectively, the components along normal line of corresponding plane are l_{1d} and l_{2d} . Thus the difference between two altitude measurements equals to the component of relative displacement on normal line given by

$$n \cdot \left\| \boldsymbol{p}_2^1 \right\| = l_{1d} - l_{2d}$$

The outputs of LM algorithm combined with the measurements from altimeters are relative pose estimate with 6 degrees of freedom and its error covariance matrix.

When relative motion parameters converge to the optimal estimates, the algorithm stops iterative loops. Based on the definition of Fisher matrix given by

$$\boldsymbol{P}_{xx}^{-1} \approx \boldsymbol{F}_{xx} = \frac{\partial^2 J(\boldsymbol{x})}{\partial \boldsymbol{x} \partial \boldsymbol{x}}$$

The error covariance matrix of relative motion estimate can be obtained.

As the UAV and camera have the same state changes. Thus relative displacement and rotation estimates of UAV are obtained by camera's relative motion estimation algorithm and then are regarded as relative state measurement to update inertial navigation. In order to built the explicit relationship between measurements and navigation states, the measurement model can be written as

$$\boldsymbol{z}_{k+m} = \begin{bmatrix} z_p \\ z_{\bar{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{p}_2^1 \\ \bar{\boldsymbol{q}}_2^1 \end{bmatrix} + \boldsymbol{\eta}_{k+m} = \begin{bmatrix} \boldsymbol{A}\left(\bar{\boldsymbol{q}}_N^1\right)\left(\boldsymbol{p}_2^N - \boldsymbol{p}_1^N\right) \\ \bar{\boldsymbol{q}}_N^1 \otimes \left(\bar{\boldsymbol{q}}_N^2\right)^{-1} \end{bmatrix} + \boldsymbol{\eta}_{k+m} \quad (6)$$

where z_p and z_q are the outputs of relative motion estimation algorithm, $\boldsymbol{p}_1^N = \boldsymbol{p}^N(t_k)$ and $\bar{\boldsymbol{q}}_N^1 = \bar{\boldsymbol{q}}_N^B(t_k)$ are the position and attitude quaternion of UAV at t_k , $\boldsymbol{p}_2^N = \boldsymbol{p}^N(t_{k+m})$ and $\bar{\boldsymbol{q}}_N^2 = \bar{\boldsymbol{q}}_N^B(t_{k+m})$ are the position and attitude quaternion at t_{k+m}, and $\boldsymbol{\eta}_{k+m}$ is measurement noise, which is assumed as the white Gaussian noise.

III. INTEGRATED NAVIGATION ALGORITHM

A. System model

Although it is beneficial to establish the high-fidelity nonlinear dynamic model for analyzing real motion process of UAV, there will be large uncertainties in the aerodynamic forces and moments. In addition, the complex system dynamic modeling will introduce strong nonlinear and increase algorithm computational complexity, which does not always bring the expected improve in navigation precision. Linear dynamic modeling or random walk modeling will also introduce unmodeled errors. Therefore the IMU measurements are used to construct the state equation in this paper.

Firstly the state variable is defined by

$$\boldsymbol{x} = \left[\left(\bar{\boldsymbol{q}}_N^B\right)^{\mathrm{T}} \quad \boldsymbol{b}_g^{\mathrm{T}} \quad \left(\boldsymbol{p}^N\right)^{\mathrm{T}} \quad \left(\boldsymbol{v}^N\right)^{\mathrm{T}} \quad \boldsymbol{b}_a^{\mathrm{T}} \right]^{\mathrm{T}} \quad (7)$$

where $\boldsymbol{p}^N$ and $\boldsymbol{v}^N$ are the position vector and velocity vector of UAV in inertial frame respectively, $\bar{\boldsymbol{q}}_N^B$ is the attitude quaternion of UAV frame relative to the inertial frame, $\boldsymbol{b}_g$ and $\boldsymbol{b}_a$ are the biases vectors of accelerometers and gyroscopes respectively, and the IMU biases are modeled as random walk processes driven by the white Gaussian noise vectors $\boldsymbol{n}_{wg}$ and $\boldsymbol{n}_{wa}$ respectively, which are augmented to the state vector to estimate for compensation.

It is assumed that the IMU is affixed to the UAV and can measure the angular velocity and acceleration relative to the inertial space. Then the IMU measurement models are respectively given by

$$\tilde{\boldsymbol{\omega}}^B = \boldsymbol{\omega}^B + \boldsymbol{b}_g + \boldsymbol{n}_g \quad (8)$$

$$\tilde{\boldsymbol{a}}^B = \boldsymbol{a}^B + \boldsymbol{b}_a + \boldsymbol{n}_a \quad (9)$$

where $\tilde{\boldsymbol{\omega}}^B$ and $\tilde{\boldsymbol{a}}^B$ are the gyro and accelerometer measurements expressed in body frame coordinate, $\boldsymbol{n}_g$ and $\boldsymbol{n}_a$ are the zero-mean Gaussian white-noise processes, respectively.

According to the attitude quaternion kinematical equation

$$\dot{\bar{\boldsymbol{q}}}_N^B = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega}_{B/N}^B)\bar{\boldsymbol{q}}_N^B \quad (10)$$

where

$$\boldsymbol{\Omega}(\boldsymbol{\omega}_{B/N}^B) = \begin{bmatrix} -[\boldsymbol{\omega}\times] & \boldsymbol{\omega}^{\mathrm{T}} \\ -\boldsymbol{\omega}^{\mathrm{T}} & 0 \end{bmatrix}$$

where $[\boldsymbol{\omega}\times]$ is antisymmetric matrix of $\boldsymbol{\omega}$.

Based on Equation (7) and (10), the system dynamics have the following form

$$\begin{cases} \dot{\bar{\boldsymbol{q}}}_N^B = \frac{1}{2}\boldsymbol{\Omega}(\tilde{\boldsymbol{\omega}}^B - \boldsymbol{b}_g - \boldsymbol{n}_g)\bar{\boldsymbol{q}}_N^B \\ \dot{\boldsymbol{b}}_g = \boldsymbol{n}_{wg} \\ \dot{\boldsymbol{p}}^N = \boldsymbol{v}^N \\ \dot{\boldsymbol{v}}^N = \boldsymbol{A}_B^N(\tilde{\boldsymbol{a}}^B - \boldsymbol{b}_a - \boldsymbol{n}_a) + \boldsymbol{g}^N \\ \dot{\boldsymbol{b}}_a = \boldsymbol{n}_{wa} \end{cases} \quad (11)$$

B. State and covariance matrix propagation

The state error vector is defined by

$$\Delta\boldsymbol{x} = \left[\left(\delta\boldsymbol{\theta}_N^B\right)^{\mathrm{T}} \quad \Delta\boldsymbol{b}_g^{\mathrm{T}} \quad \left(\Delta\boldsymbol{p}^N\right)^{\mathrm{T}} \quad \left(\Delta\boldsymbol{v}^N\right)^{\mathrm{T}} \quad \Delta\boldsymbol{b}_a^{\mathrm{T}} \right]^{\mathrm{T}} \quad (12)$$

where the state errors of the position, velocity and IMU biases are defined as the additive errors as follow $\Delta\boldsymbol{x} = \boldsymbol{x} - \hat{\boldsymbol{x}}$ and the attitude error is defined as the multiplicative error which is the attitude rotation from the estimated quaternion to the true value approximately written as

$$\delta\bar{\boldsymbol{q}} = \bar{\boldsymbol{q}} \otimes \hat{\bar{\boldsymbol{q}}}^{-1} \simeq \begin{bmatrix} \frac{1}{2}\delta\boldsymbol{\theta} \\ 1 \end{bmatrix} \quad (13)$$

As the error-quaternion $\delta\overline{\boldsymbol{q}}$ belongs to a small angle attitude rotation, the fourth component δq_4 infinitely approaches 1 and the attitude information is mainly contained in the vector part. The advantage of this three-dimensional attitude error parameterization is that the unit norm constraint of the quaternion can be approximately maintained by first-order and the attitude estimation covariance describes the attitude estimation uncertainty in the body frame, which has intuitive physical meaning.

Based on the definition of state error vector, state error equation can be obtained and be written as matrix vector as follows:

$$\Delta\dot{\boldsymbol{x}} = \boldsymbol{F}(t)\Delta\boldsymbol{x} + \boldsymbol{G}(t)\boldsymbol{n} \tag{14}$$

where the statistics of the system noise $\boldsymbol{n} = \begin{bmatrix} \boldsymbol{n}_g^{\mathrm{T}} & \boldsymbol{n}_{\omega g}^{\mathrm{T}} & \boldsymbol{n}_a^{\mathrm{T}} & \boldsymbol{n}_{\omega a}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ satisfies

$$\begin{cases} E[\boldsymbol{n}(t)] = 0 \\ E\left[\boldsymbol{n}(t)\boldsymbol{n}^{\mathrm{T}}(t')\right] = \boldsymbol{Q}(t)\delta(t-t') \end{cases}$$

Thus the propagation equation of the state estimation error covariance matrix is obtained to be

$$\dot{\boldsymbol{P}}(t) = \boldsymbol{F}(t)\boldsymbol{P}(t) + \boldsymbol{P}(t)\boldsymbol{F}^{\mathrm{T}}(t) + \boldsymbol{G}(t)\boldsymbol{Q}(t)\boldsymbol{G}^{\mathrm{T}}(t) \tag{15}$$

The system model is discretized through the fourth order Runge-Kutta method, and then the state and covariance are propagated using the IMU measurement information.

C. State augmentation

From measurement equation (6), relative state measurements are related to the state vectors of two consecutive instances. As standard EKF algorithm can only deal with measurements related to current state, state vectors need to be augmented. The state augmentation algorithm copies the state vector of first imaging and maintain it during propagation, thus the measurement update can be performed by EKF based on the state correlation items and relative state measurements.

Let T_i and $T_c = mT_i$ denote the IMU sampling period and the camera measurement period respectively. It is assumed that $t_k = t_{k,0}$ and $t_{k+1} = t_{k,m}$ then we can analyze the state augmentation algorithm in the *k*th update period. It is assumed that the state estimate at time t_k is $\hat{\boldsymbol{x}}_{k,0}^{+}$, then the augmented state vector is present as

$$\tilde{\boldsymbol{x}}_{k,0}^{+} = \begin{bmatrix} \hat{\boldsymbol{x}}_{k,0}^{+} \\ \hat{\boldsymbol{x}}_{sk,0}^{+} \end{bmatrix} \tag{16}$$

where $t_k = t_{k,0}$, $t_{k+1} = t_{k,m}$.

During the interval $\left[t_k, t_{k,m}\right]$, $\hat{\boldsymbol{x}}_{k,0}^{+}$ is invariant, and $\hat{\boldsymbol{x}}_{k,0}^{-}$ is propagated through IMU measurement values.

It is inferred from the equation (16) that the state vector $\hat{\boldsymbol{x}}_{sk,0}^{-}$ and $\hat{\boldsymbol{x}}_{k,0}^{-}$ at the moment t_k are completely correlated. Therefore the augmented covariant matrix of state error is given by

$$\tilde{\boldsymbol{P}}_{k,0}^{+} = \begin{bmatrix} \boldsymbol{P}_{k,0}^{+} & \boldsymbol{P}_{k,0}^{+} \\ \boldsymbol{P}_{k,0}^{+} & \boldsymbol{P}_{k,0}^{+} \end{bmatrix} \tag{17}$$

The one-step state prediction equation can be represented by

$$\tilde{\boldsymbol{P}}_{k,1}^{-} = \begin{bmatrix} \boldsymbol{\Phi}_{k+1}\boldsymbol{P}_{k,0}^{+}\boldsymbol{\Phi}_{k+1}^{\mathrm{T}} + \boldsymbol{N}_{k+1} & \boldsymbol{\Phi}_{k+1}\boldsymbol{P}_{k,0}^{+} \\ \boldsymbol{P}_{k,0}^{+}\boldsymbol{\Phi}_{k+1}^{\mathrm{T}} & \boldsymbol{P}_{k,0}^{+} \end{bmatrix} \tag{18}$$

where $\boldsymbol{\Phi}_{k+1}$ and $\boldsymbol{N}_{k+1}$ are discrete state transition matrix and noise covariance matrix represented by

$$\boldsymbol{\Phi}_{k+1} = \exp\left(\int_{t_{k,0}}^{t_{k,1}} \boldsymbol{F}(\tau)d\tau\right) \quad \boldsymbol{N}_{k+1} = \int_{t_{k,0}}^{t_{k,1}} \boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}d\tau$$

The *m*th step state prediction equation can be obtained by (14) as follows

$$\tilde{\boldsymbol{P}}_{k,m}^{-} = \begin{bmatrix} \boldsymbol{P}_{k,m}^{-} & \boldsymbol{\Lambda}\boldsymbol{P}_{k,0}^{+} \\ \boldsymbol{P}_{k,0}^{+}\boldsymbol{\Lambda}^{\mathrm{T}} & \boldsymbol{P}_{k,0}^{+} \end{bmatrix} \tag{19}$$

where $\boldsymbol{\Lambda} = \prod_{i=0}^{m} \boldsymbol{\Phi}_{k+i}$. As the positions and attitudes are apparently correlated with the relative measurements, in the practice application the reduce-order filter algorithm is used where the dimension of state vector decreases from 32 to 23.

D. Measurement update

As the definitions of position error and attitude error are different, measurements residuals of relative position and attitude are supposed to solve respectively to solve measurement sensitivity matrix. From equation (6) the relative position measurement model is given by

$$\boldsymbol{z}_p = \boldsymbol{h}_1\left(\boldsymbol{x}_{k+m}, \boldsymbol{x}_{sk}\right) + \boldsymbol{\eta}_1 = \boldsymbol{A}\left(\overline{\boldsymbol{q}}_M^1\right)\left(\boldsymbol{p}_2^M - \boldsymbol{p}_1^M\right) + \boldsymbol{\eta}_1 \tag{20}$$

Then measurement sensitivity matrix is denoted by

$$\boldsymbol{H}_p = \begin{bmatrix} \dfrac{\partial \boldsymbol{h}_1}{\partial \boldsymbol{x}_{k+m}} & \dfrac{\partial \boldsymbol{h}_1}{\partial \boldsymbol{x}_{sk}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0}_{3\times 6} & \boldsymbol{H}_{p2} & \boldsymbol{0}_{3\times 6} & \boldsymbol{H}_{\overline{q}1} & \boldsymbol{H}_{p1} \end{bmatrix}$$

where

$$\boldsymbol{H}_{p2} = \frac{\partial \boldsymbol{h}_1}{\partial \boldsymbol{p}_{k+m}} = \boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}_M^1) \qquad \boldsymbol{H}_{p1} = \frac{\partial \boldsymbol{h}_1}{\partial \boldsymbol{p}_{sk}} = -\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}_M^1)$$

To avoid solving partial derivative of the attitude directly, it can be resolved by characteristics of attitude matrix as follows

$$\boldsymbol{A}(\overline{\boldsymbol{q}}) = \boldsymbol{A}(\delta\overline{\boldsymbol{q}} \otimes \hat{\overline{\boldsymbol{q}}}) = \boldsymbol{A}(\delta\overline{\boldsymbol{q}})\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}) \approx \boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}) - [\delta\boldsymbol{\theta}\times]\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}})$$

Substituting it into the equation (20) gives

$$\boldsymbol{h}_1 = \left(\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}_M^1) - [\delta\boldsymbol{\theta}\times]\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}_M^1)\right)\left(\boldsymbol{p}_2^M - \boldsymbol{p}_1^M\right) =$$

$$\hat{\boldsymbol{h}}_1 + \left[\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}_M^1)\left(\boldsymbol{p}_2^M - \boldsymbol{p}_1^M\right)\times\right]\delta\boldsymbol{\theta}$$

So $\boldsymbol{H}_{\overline{q}1} = \left[\boldsymbol{A}(\hat{\overline{\boldsymbol{q}}}_M^1)\left(\boldsymbol{p}_2^M - \boldsymbol{p}_1^M\right)\times\right]$

Thus measurements residual equation for the relative position is given by

$$\Delta\boldsymbol{z}_p = \boldsymbol{H}_{p2}\Delta\boldsymbol{p}_2 + \boldsymbol{H}_{p1}\Delta\boldsymbol{p}_1 + \boldsymbol{H}_{\overline{q}1}\delta\boldsymbol{\theta}_1 + \boldsymbol{\eta}_1 \tag{21}$$

From equation (6) the measurement model for the relative attitude is

$$\boldsymbol{z}_{\bar{q}} = \boldsymbol{h}_2\left(\boldsymbol{x}_{k+m}, \boldsymbol{x}_{sk}\right) + \boldsymbol{\eta}_2 = \bar{\boldsymbol{q}}_M^1 \otimes \left(\bar{\boldsymbol{q}}_M^2\right)^{-1} + \boldsymbol{\eta}_2 \tag{22}$$

Then measurement sensitivity matrix is denoted by

$$\boldsymbol{H}_{\bar{q}} = \left[\frac{\partial \boldsymbol{h}_2}{\partial \boldsymbol{x}_{k+m}} \quad \frac{\partial \boldsymbol{h}_2}{\partial \boldsymbol{x}_{sk}}\right] = \left[\boldsymbol{H}'_{\bar{q}2} \quad \boldsymbol{0}_{3\times 12} \quad \boldsymbol{H}'_{\bar{q}1} \quad \boldsymbol{0}_{3\times 3}\right]$$

Similarly

$$\boldsymbol{H}'_{\bar{q}2} = -\frac{1}{2}\boldsymbol{A}(\hat{\bar{\boldsymbol{q}}}_2^1) \qquad \boldsymbol{H}'_{\bar{q}1} = \frac{1}{2}\boldsymbol{I}_{3\times 3}$$

In summary, the measurements residual equation can be present as:

$$\boldsymbol{\vartheta}_{k+m} \simeq \boldsymbol{z}_{k+m} - \hat{\boldsymbol{z}}_{k+m} = \boldsymbol{H}_{sk}\Delta\boldsymbol{x}_{sk} + \boldsymbol{H}_{k+m}\Delta\boldsymbol{x}_{k+m} + \boldsymbol{\eta}_{k+m} \tag{23}$$

where

$$\boldsymbol{H}_{k+m} = \begin{bmatrix} 0 & 0 & \boldsymbol{H}_{p2} & 0 & 0 \\ \boldsymbol{\Xi}\boldsymbol{H}'_{\bar{q}2} & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{H}_{sk} = \begin{bmatrix} \boldsymbol{H}_{\bar{q}1} & \boldsymbol{H}_{p1} \\ \boldsymbol{\Xi}\boldsymbol{H}'_{\bar{q}1} & 0 \end{bmatrix}$$

Then the correction to the state estimation can be expressed by the filter gain matrix $\boldsymbol{K}_{k+m}$ as follow

$$\Delta\hat{\boldsymbol{x}}_{k+m} = \boldsymbol{K}_{k+m}\ \boldsymbol{\vartheta}_{k+m} \tag{24}$$

The attitude quaternion is updated using

$$\hat{\bar{\boldsymbol{q}}}_{k+m}^{+} = \delta\hat{\bar{\boldsymbol{q}}}_{k+m} \otimes \hat{\bar{\boldsymbol{q}}}_{k+m}^{-} \simeq \begin{bmatrix} \frac{1}{2}\delta\hat{\boldsymbol{\theta}}_{k+m} \\ 1 \end{bmatrix} \otimes \hat{\bar{\boldsymbol{q}}}_{k+m}^{-} \tag{25}$$

Other state vectors are updated as follows

$$\begin{bmatrix} \hat{\boldsymbol{b}}_{gk+m}^{+} \\ \hat{\boldsymbol{p}}_{k+m}^{+} \\ \hat{\boldsymbol{v}}_{k+m}^{+} \\ \hat{\boldsymbol{b}}_{ak+m}^{+} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{b}}_{gk+m}^{-} \\ \hat{\boldsymbol{p}}_{k+m}^{-} \\ \hat{\boldsymbol{v}}_{k+m}^{-} \\ \hat{\boldsymbol{b}}_{ak+m}^{-} \end{bmatrix} + \begin{bmatrix} \Delta\hat{\boldsymbol{b}}_{gk+m} \\ \Delta\hat{\boldsymbol{p}}_{k+m} \\ \Delta\hat{\boldsymbol{v}}_{k+m} \\ \Delta\hat{\boldsymbol{b}}_{ak+m} \end{bmatrix} \tag{26}$$

As the measurement update process has been finished at t_{k+m}, the state estimates and line-of-sight measurements are augmented to the state vectors. The process above will be repeated at each filtering period of the navigation algorithm.

IV. Simulation Results

In order to validate the proposed navigation algorithm, the mathematical simulation is performed. The block diagram is illustrated in Fig.2. As many features tracking algorithms have been developed in image processing field, the real tracking process is not studied in this simulation and the stochastic errors are added to the feature tracking. The position of the feature in the focal plane is computed by a pinhole camera model.

The roll maneuver process is assumed as follow: first rolling the UAV to the right of 30 deg, then rolling the UAV to the left of 60 deg, finally rolling the UAV to the horizontal position. The IMU and the camera are sampled every 0.01 sec and 0.2 sec respectively in this simulation. The measurement noise intensities of the gyroscope and the accelerometer are set to $0.01(^\circ)/\sqrt{\text{h}}$ and $10\mu\text{g}\sqrt{\text{s}}$ respectively, the corresponding bias noise intensities are set to $0.05(^\circ)/\text{h}$ and 0.1mg respectively. The camera has a 30 deg field-of-view (FOV) and the measurement accuracy is assumed to be 1 pixel. The number of the frames is 375, and the number of the features varies from 10 to 40. To obtain optimal filter performance, the process noise covariance matrix is set up to be the noise intensities of IMU, which is then adjusted to compensate the process model nonlinearity.

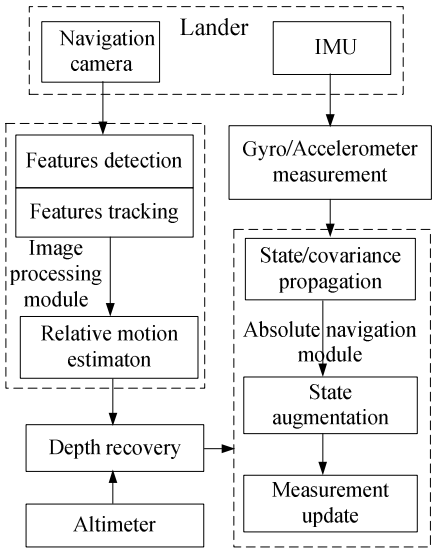

Fig. 2 navigation algorithm block diagram

The EKF algorithm processes the measurements of the camera and IMU simultaneously after the navigation initialization. Figures 3 through 5 present the navigation outputs for the triaxial position errors, attitude errors and biases errors. The simulation results show that The navigation algorithm can recover from great initial errors and converge to satisfactory steady state value, which means that relative pose measurement information can correct the inertial navigation biases and efficiently improve the navigation performance.

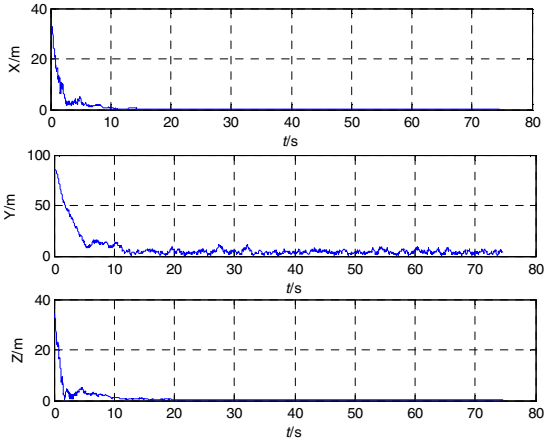

Fig. 3 Position estimation errors

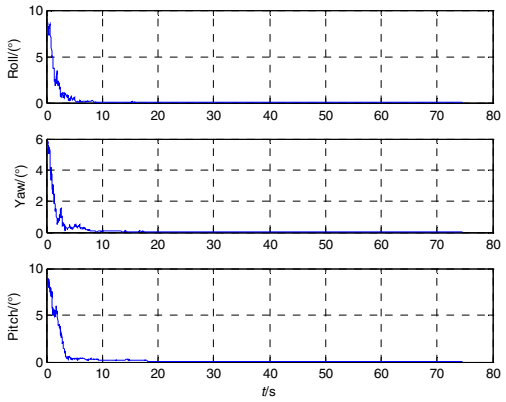

Fig. 4 Attitude estimation errors

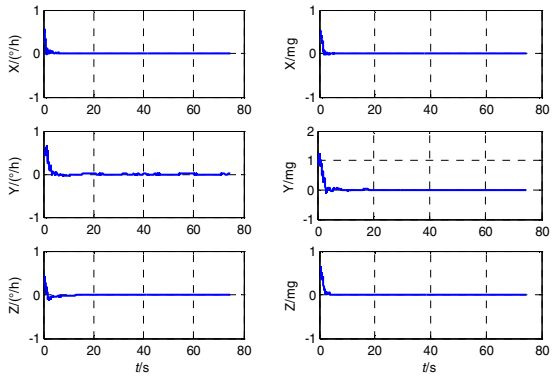

Fig. 5 IMU biases estimation errors

V. CONCLUSION

Aiming at high-precision UAV navigation without the known ground cooperative target, we propose a vision-aided inertial navigation scheme. To avoid the homography decomposition, the inter-frame relative motion parameters are explicitly related to the navigation state and then fed into the EKF framework. The simulation results show that the algorithm is able to overcome the flaw of SLAM-based navigation algorithm and inertial navigation algorithm, and assures precision, autonomy and security of UAV missions. The performance analysis and observability analysis of the integrated navigation algorithm under different influencing factor are the next steps for further work.

References

[1] C. Liu, S. Prior, W. Teacy, and M. Warner, "Computationally efficient visual– inertial sensor fusion for Global Positioning System–denied navigation on a small quadrotor," Advances in Mechanical Engineering, vol. 8, no. 3, pp. 1-11, 2016.

[2] F. Santoso, M. A. Garratt and S. G. Anavatti, "Visual–Inertial Navigation Systems for Aerial Robotics: Sensor Fusion and Technology," IEEE Transactions on Automation Science and Engineering, vol. 14, no. 1, pp. 260-275, Jan. 2017.

[3] Bloesch, Michael, et al. "Robust visual inertial odometry using a direct EKF-based approach." Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on. IEEE, 2015.

[4] A. I. Mourikis, N. Trawny, S. I. Roumeliotis, A. E. Johnson, A. Ansar, and L. H. Matthies, "Vision-aided inertial navigation for spacecraft entry, descent, and landing," IEEE Transactions on Robotics, vol. 25, no. 2, pp. 264-280, 2009.

[5] D. D. Diel, "Stochastic constraints for vision-aided inertial navigation," Master's thesis, MIT, January 2005.

[6] E. Jones and S. Soatto, "Visual-inertial navigation, mapping and localization: A scalable real-time causal approach," International Journal of Robotics Research, vol. 30, no. 4, pp. 407–430, Apr. 2011.

[7] J. Kelly and G. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," International Journal of Robotics Research, vol. 30, no. 1, pp. 56–79, Jan. 2011.

[8] M. Li and A. I. Mourikis, "High-precision, consistent EKF-based visual-inertial odometry," International Journal of Robotics Research, vol. 32, no. 6, pp. 690–711, May 2013.

[9] B D Hoffman, E T Baumgartner, T L Huntsberger, et al. "Improved state estimation in challenging terrain". Autonomous Robots, vol. 6, no. 2, pp. 113-130, 1999.

[10] T P Webb, R J Prazenica, A J Kurdila, et al. "Vision-based state estimation for autonomous micro air vehicles". Journal of Guidance Control and Dynamics, vol. 30, no. 3, pp. 816-826, 2007.

[11] P Gurfil, H Rotstein. "Partial aircraft state estimation from visual motion using the subspace constraints approach". Journal of Guidance Control and Dynamics, vol. 24, no. 5, pp. 1016-1028, 2001.

[12] V Indelman, P Gurfil, E Rivlin, et al. "Real-time vision-aided localization and navigation based on three-view geometry". IEEE Transactions on Aerospace and Electronic Systems, vol. 48, no. 3, pp. 2239-2259, 2012.

[13] S Baker, I Matthews. "Lucas-Kanade 20 years on: A unifying framework". International Journal of Computer Vision, vol. 56, no. 3, pp. 221-255, 2004.

[14] D Lowe. "Distinctive image features from scale-invariant keypoints". International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[15] M Fischler, R Bolles. "Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography". Communication of the Association for Computing Machinery, vol. 24, pp. 381-395, 1981.

[16] S Y Zhao, F Lin, K M Peng, et al. "Homography-based vision-aided inertial navigation of UAVs in unknown environments". AIAA Guidance, Navigation and Control Conference, Minneapolis, Minnesota, August, pp. 13-16, 2012.

[17] J L Crassidis, R Alonso, J L Junkins. "Optimal attitude and position determination from line-of-sight measurements". The Journal of the Astronautical Sciences, vol. 48, no. 2/3, pp. 391-408, 2000.