

Micro Aerial Vehicle Navigation with Visual-Inertial Integration Aided by Structured Light

Yunshu Wang^{1,2,3}, Jianye Liu^{1,2}, Jinling Wang³, Qinghua Zeng^{1,2},
Xuesong Shen³ and Yueyuan Zhang¹

¹(Navigation Research Center, College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China)

²(Jiangsu Key Laboratory of Internet of Things and Control Technologies (NUAA), Nanjing, 211106, China)

³(School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia)

(E-mail: zengqh@nuaa.edu.cn)

Considering that traditional visual navigation cannot be utilised in low illumination and sparse feature environments, a novel visual-inertial integrated navigation method using a Structured Light Visual (SLV) sensor for Micro Aerial Vehicles (MAVs) is proposed in this paper. First, the measurement model based on an SLV sensor is studied and built. Then, using the state model based on error equations of an Inertial Navigation System (INS), the measurement model based on the error of the relative motion measured by INS and SLV is built. Considering that the measurements in this paper are mainly related to the position and attitude information of the present moment, the state error accumulation in traditional visual-inertial navigation can be avoided. An Adaptive Sage-Husa Kalman Filter (ASHKF) based on multiple weighting factors is proposed and designed to make full use of the SLV measurements. The results of the simulation and the experiment based on real flight data indicate that high accuracy position and attitude estimations can be obtained with the help of the algorithm proposed in this paper.

KEY WORDS

1. Integrated Navigation System (INS).
2. Indoor Navigation.
3. Inertial Navigation System (INS).
4. Optical Navigation.
5. LASER.

Submitted: 18 April 2017. Accepted: 15 May 2019. First published online: 1 July 2019.

1. INTRODUCTION. Considering the small size, light weight, low cost, high manoeuvrability and decreased risk to human life, Micro Aerial Vehicles (MAVs) are becoming widely used in applications such as infrastructure maintenance and management systems (Eschmann et al., 2012), precision farming (Zhang and Kovacs, 2012), aerial photography and cinematography (Santoso et al., 2017) and security applications (Tarchi et al., 2017). Low-cost Inertial Measurement Units (IMU) and Global Navigation

Satellite System (GNSS) receivers are often applied in MAVs to obtain navigation information (Zhang et al., 2017). Although an IMU is able to track sudden motions both in angular velocity and linear acceleration, it is subject to a growing unbounded error caused by the biases of the sensors. The absolute velocity and position of the vehicle can be measured by GNSS directly, which is very helpful in decreasing the accumulated errors of an IMU (Wang et al., 2013; Meng et al., 2017). In a GNSS denied environment (indoor environments for example), other sensors such as pseudolites (Gioia and Borio, 2014) and visual sensors are applied to supply the navigation information for MAVs.

Due to their small size and light weight, visual sensors are widely used in MAVs. Active vision can send laser or structured light to the environment, and receive reflection information to perceive the environment, while passive vision can only receive the natural light of the environment. To obtain detailed information, passive vision requires features from the environment. Artificial marks are often utilised to calculate the pose in a cooperative environment (Cocchioni et al., 2014). In a noncooperative environment, features such as points and lines in two successive images taken by a monocular camera can be matched to obtain depth and position information (Engel et al., 2014; Zeng et al., 2017). In a stereo camera, the image features in right and left cameras can be matched (Liu et al., 2012). However, the former methods will be invalid when the features in the image are sparse, such as in an indoor environment which lacks illumination or texture.

The former problem can be solved in active vision methods by projecting light into the environment. The depth information can be obtained from reflected light, such as using a Light Direction And Ranging (LIDAR), a Red Green Blue-Depth (RGB-D) camera and a structured light visual sensor. LIDAR and an RGB-D camera are usually applied to solve the problem of Simultaneous Localisation And Mapping (SLAM) (Liu et al., 2015; Gao and Zhang, 2015), because these sensors can be used to generate a point cloud of the environment. However, the extremely high costs of hardware and data processing are the major drawbacks for LIDAR and RGB-D cameras.

A Structured Light Visual (SLV) sensor has many advantages, such as simple construction, fast speed and low cost (Wei et al., 2005). Structured light visual sensors have been used in MAV for pose measurement or obstacle avoidance over the past few years. Natraj et al. (2013) presented a pose measurement method using a monocular camera and a circle structured light projector to measure the pitch, roll and altitude of a MAV in dark environments. Sanahuja and Castillo (2013) presented an embedded system composed of a monocular camera and a line structured light projector for indoor MAV navigation. The system detected line structured light on a wall in front of the MAV, to estimate the distance to the wall and yaw angle of the MAV. In Dougherty et al. (2014), multiple dot structured light projectors and monocular vision were used to measure the relative pose between a MAV and an inclined plane for MAV landing. In Wang et al. (2015), a cross-structured light sensor was used to calculate the related pose between MAV and the line on the floor. Harmat and Sharf (2014) combined a dot matrix structured light projector and a stereo camera to estimate the depth of the complex environment. It is similar to the Time of Flight (ToF) camera and RGB-D camera in a certain sense, but it has a lighter weight and can also be used in an outdoor environment. However, the former research works did not take full advantage of the high-frequency IMU information, and the robustness of the system could be improved further.

According to the different combinations, the filtering fusion solutions based on the visual and inertial information can be grouped into loosely coupled and tightly coupled approaches. In the loosely coupled approach, the measurements from IMU and camera can be processed separately, with low complexity algorithms (Huang et al., 2014). In Wang et al. (2013), the motion of a vehicle was estimated by a homography-based approach and the visual measurement was fused with IMU based on an indirect Extended Kalman Filter (EKF). Weiss et al. (2012) fused the camera and IMU data within a real-time EKF framework. In Liu et al. (2016), three separated Kalman Filters (KF) for attitude, orientation and position were assembled for a stereo visual-inertial odometry. Tightly coupled visual-inertial navigation was used to augment the Three-Dimensional (3D) feature positions in the filter state, and to estimate the pose and 3D points concurrently (Kelly and Sukhatme, 2011). Xian et al. (2015) employed two categories of features called 3D points and Inverse Depth (ID) points to integrate with inertial information based on the iterative EKF. Kong et al. (2015) employed the EKF for error state prediction and covariance propagation, and the Sigma Point Kalman filter (SPKF) for measurement updating. The state vectors of these methods can be augmented to obtain the measurements related to the last and current moments. So, the measurement model of a Kalman Filter has to be built, which increases the complexity of the integrated navigation system.

In this paper, a cross-structured light visual sensor, which consists of a cross-structured light projector and a monocular camera, is proposed to integrate with an IMU for MAV navigation. The pose of the MAV can be obtained when the cross-structured light and a segment line on the floor are captured by a monocular camera. Pose information can be obtained even when the MAV is in an insufficient illumination and sparse feature environment. An adaptive Sage-Husa Kalman filter (ASHKF) with multiple weighting factors based on the SLV and Inertial Navigation System (INS) is proposed to solve the problem of the measurement errors in the integrated navigation system.

This paper is organised as follows: The pose measurement model based on cross-structured SLV is introduced in Section 2. The sensor time synchronisation method and adaptive filter are described in Section 3. In Sections 4 and 5, simulation analyses and a real data experiment are carried out, respectively. Finally, the conclusion of this work is summarised in Section 6.

2. POSE ESTIMATION BASED ON CROSS-STRUCTURED LIGHT VISUAL SENSOR. Based on research of the measurement model of the cross-structured light visual sensor, the pose estimation algorithm is proposed in this section.

2.1. Coordinate systems of cross-structured light visual sensor. The model of the cross-structured light visual sensor is shown in Figure 1. $C - X_c Y_c Z_c$ represents the camera frame. $o - uv$ represents the image frame. Axis Z_c is along the forward optical axis of the camera. $L - X_l Y_l Z_l$ represents the cross-structured light frame, which consists of two structured light planes s_1 and s_2 . Point L is the firing point of cross-structured light. Axis Z_l is along the forward intersecting line of the two light planes. Axes X_l and Y_l are perpendicular to axis Z_l in each structured light plane. $W - X_w Y_w Z_w$ represents the world frame. Axis Z_w is perpendicular to the floor which is represented by plane $X_w W Y_w$. Point P is an arbitrary point of the intersections of plane $X_w W Y_w$ and structured light planes. The image point of P on the image plane is p . $O - X_o Y_o Z_o$ represents the segment frame. The origin point O is an arbitrary point on the segment. Axis X_o is along the segment on the floor. Axis Y_o is

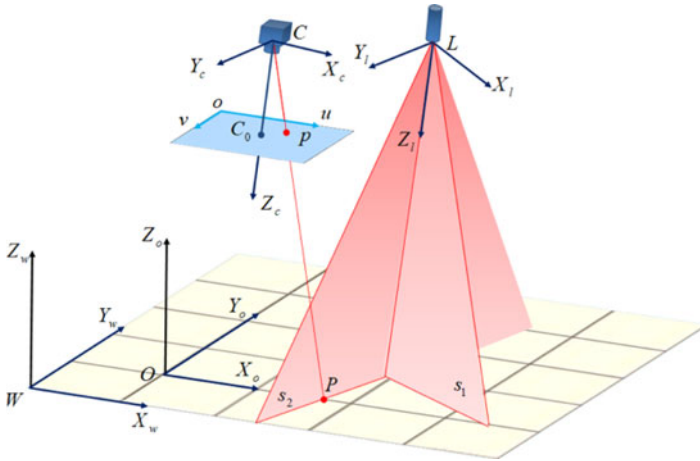


Figure 1. The model of cross-structured light sensor.

perpendicular to axis X_o and is on the floor. Axis Z_o is perpendicular to the floor. All the frames are right-hand Cartesian frames.

Coordinates of point P in the camera frame are (x_p^c, y_p^c, z_p^c) . Coordinates of the image point p on the image frame are (u_p, v_p) . The distortion of the camera is tiny so it can be ignored here. The relationship of image coordinates and camera coordinates are as Equation (1):

$$z_p^c \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{dx} & 0 & u_0 \\ 0 & \frac{f}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_p^c \\ y_p^c \\ z_p^c \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_p^c \\ y_p^c \\ z_p^c \end{bmatrix} \tag{1}$$

where \mathbf{K} is called the internal parameter matrix of the camera. It can be obtained through camera calibration (Bouguet, 2015), dx and dy are the effective pixel pitches in the directions of x and y on the image plane, respectively. f is the focal length of the camera and (u_0, v_0) are the optical centre coordinates on the image plane.

According to Equation (1), there is a unique point p in the image frame and there is a unique straight line Cp , which is associated with the camera (C) and the point p . The coordinates of point p in the camera frame are $((u_p - u_0)dx, (v_p - v_0)dy, f)$, so the equation of Cp in the coordinate system of the camera can be represented as Equation (2), where (X, Y, Z) represent the coordinates of the point on the Cp ray.

$$\frac{X}{(u_p - u_0)} \cdot \frac{f}{dx} = \frac{Y}{(v_p - v_0)} \cdot \frac{f}{dy} = Z \tag{2}$$

The structured light planes s_1, s_2 in the camera frame can be assumed as:

$$a_1x_{s_1}^c + b_1y_{s_1}^c + c_1z_{s_1}^c + d_1 = 0 \tag{3}$$

$$a_2x_{s_2}^c + b_2y_{s_2}^c + c_2z_{s_2}^c + d_2 = 0 \tag{4}$$

where, $(x_{s_1}^c, y_{s_1}^c, z_{s_1}^c)$ is an arbitrary point on the plane s_1 , so as $(x_{s_2}^c, y_{s_2}^c, z_{s_2}^c)$.

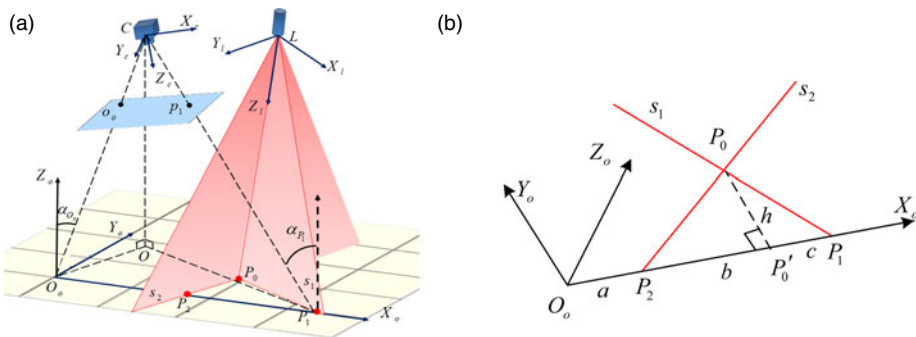


Figure 2. Illustration of attitude and position measurement (a) Relationship between each frame for posture measurement (b) Relationship between three structured light points.

The equations of structured light in the camera frame can be obtained by calibration of the structured light sensor (Kiddee et al., 2016). Then the coordinates of point P in the camera frame can be obtained by calculating the intersection of ray Cp and plane s_1 or s_2 .

2.2. *Pose measurements based on cross-structured light visual sensor.* In Figure 2(a), P_i ($i = 0, 1, 2$) are the points on the cross-structured light. Point P_0 is the intersection of two structured light planes on the floor. Points P_1 and P_2 are the intersections of the cross-structured light and the segment on the floor.

As shown in Figure 2(b), the relationship between P_i ($i = 0, 1, 2$) is:

$$\begin{bmatrix} x_1^o \\ y_1^o \\ z_1^o \end{bmatrix} = \begin{bmatrix} x_0^o + c \\ y_0^o - h \\ z_0^o \end{bmatrix}, \begin{bmatrix} x_2^o \\ y_2^o \\ z_2^o \end{bmatrix} = \begin{bmatrix} x_0^o - b + c \\ y_0^o - h \\ z_0^o \end{bmatrix} \tag{5}$$

where $[x_i^o \ y_i^o \ z_i^o]^T$, ($i = 0, 1, 2$) are the coordinates of P_i ($i = 0, 1, 2$) in the segment frame. The segment $P_0P'_0 = h$ is perpendicular to the segment $P_1P_2 = b$ and $c = P_1P'_0$.

The following equations can be obtained from Equation (5):

$$\begin{bmatrix} x_0^c \\ y_0^c \\ z_0^c \end{bmatrix} = \mathbf{R}_O^C \times \begin{bmatrix} x_0^o \\ y_0^o \\ z_0^o \end{bmatrix} + \mathbf{T}_O^C \tag{6}$$

$$\begin{bmatrix} x_1^c \\ y_1^c \\ z_1^c \end{bmatrix} = \mathbf{R}_O^C \times \begin{bmatrix} x_0^o + c \\ y_0^o - h \\ z_0^o \end{bmatrix} + \mathbf{T}_O^C \tag{7}$$

$$\begin{bmatrix} x_2^c \\ y_2^c \\ z_2^c \end{bmatrix} = \mathbf{R}_O^C \times \begin{bmatrix} x_0^o + c - b \\ y_0^o - h \\ z_0^o \end{bmatrix} + \mathbf{T}_O^C \tag{8}$$

where $\mathbf{R}_O^C = \begin{bmatrix} C_\psi C_\theta & -C_\psi S_\theta S_\gamma & -S_\psi C_\gamma & -C_\psi S_\theta C_\gamma + S_\psi S_\gamma \\ S_\psi C_\theta & -S_\psi S_\theta S_\gamma + C_\psi C_\gamma & -S_\psi S_\theta C_\gamma - C_\psi S_\gamma \\ S_\theta & C_\theta S_\gamma & C_\theta C_\gamma \end{bmatrix}$ is the rotation matrix from the segment frame to the camera frame and \mathbf{T}_O^C is the translation matrix from the segment frame to the camera frame.

With Equations (6) to (8), eliminating \mathbf{T}_O^C and x_0^o, y_0^o, z_0^o , the angles in the rotation matrix \mathbf{R}_O^C can be obtained in Equations (9).

$$\begin{aligned} \theta &= \arcsin \frac{z_1^c - z_2^c}{b} \\ \psi &= \arctan \frac{y_1^c - y_2^c}{x_1^c - x_2^c} \\ \gamma &= \arctan \frac{(z_1^c - z_0^c)C_\theta - (y_1^c - y_0^c)S_\theta C_\psi - (x_1^c - x_0^c)S_\theta S_\psi}{C_\psi(y_1^c - y_0^c) - S_\psi(x_1^c - x_0^c)} \end{aligned} \tag{9}$$

where $b = \sqrt{(x_1^c - x_2^c)^2 + (y_1^c - y_2^c)^2 + (z_1^c - z_2^c)^2}$.

In Figure 2(a), p_1 and o_o are the image points of P_1 and O_o . According to Equation (3), the equations of CP_1 and CO_o in the camera frame can be obtained with Equations (10) and (11), respectively.

$$\frac{X_{P_1}}{(u_{p_1} - u_0)} \cdot \frac{f}{dx} = \frac{Y_{P_1}}{(v_{p_1} - v_0)} \cdot \frac{f}{dy} = Z_{P_1} \tag{10}$$

$$\frac{X_{O_o}}{(u_{o_o} - u_0)} \cdot \frac{f}{dx} = \frac{Y_{O_o}}{(v_{o_o} - v_0)} \cdot \frac{f}{dy} = Z_{O_o} \tag{11}$$

where (u_{p_1}, v_{p_1}) and (u_{o_o}, v_{o_o}) are the coordinates of p_1 and o_o in the image frame.

As the rotation matrix \mathbf{R}_O^C shows, the direction vector of axis Z_o in the camera frame is:

$$\begin{bmatrix} x_{Z_o} \\ y_{Z_o} \\ z_{Z_o} \end{bmatrix} = \begin{bmatrix} -\cos \psi \sin \theta \cos \gamma + \sin \psi \sin \gamma \\ -\sin \psi \sin \theta \cos \gamma - \cos \psi \sin \gamma \\ \cos \theta \cos \gamma \end{bmatrix} \tag{12}$$

Then the included angle α_{P_1} between axis Z_o and CP_1 can be calculated using the included angle formula, as well as the included angle α_{O_o} between the axis Z_o and the axis CO_o .

As Figure 2(a) shows, CO is perpendicular to plane S and the perpendicular foot is O . There are right triangles ΔCOP_1 and ΔCOO_o . The length of CP_1 can be obtained using the coordinates of P_1 in the camera frame, which can be calculated with Equation (7).

$$CP_1 = \sqrt{x_1^{c2} + y_1^{c2} + z_1^{c2}} \tag{13}$$

In the right triangle ΔCOP_1 :

$$CO = CP_1 \cos \alpha_{P_1} \tag{14}$$

In the other right triangle ΔCOO_o :

$$CO_o = \frac{CO}{\cos \alpha_{O_o}} \tag{15}$$

The length of CO_o can be represented as $CO_o = \lambda_{O_o} \cdot \sqrt{x_{O_o}^2 + y_{O_o}^2 + z_{O_o}^2}$. Therefore, $\lambda_{O_o} = CO_o / \sqrt{x_{O_o}^2 + y_{O_o}^2 + z_{O_o}^2}$ can be represented as:

$$\lambda_{O_o} = \frac{\sqrt{x_1^{c2} + y_1^{c2} + z_1^{c2}}}{\sqrt{x_{O_o}^2 + y_{O_o}^2 + z_{O_o}^2}} \cdot \frac{\cos \alpha_{P_1}}{\cos \alpha_{O_o}} \quad (16)$$

where $[x_{o_o} \ y_{o_o} \ z_{o_o}]^T$ is the direction vector of CO_o in the camera frame. According to Equation (1), $x_{O_o} = (u_{o_o} - u_0) \cdot dx/f$, $y_{O_o} = (v_{o_o} - v_0) \cdot dy/f$ and $z_{O_o} = 1$. The coordinates $(\lambda_{O_o}x_{O_o}, \lambda_{O_o}y_{O_o}, \lambda_{O_o}z_{O_o})$ of O_o in the camera frame can be obtained. Therefore, the translation matrix from segment frame to the camera frame is:

$$\mathbf{T}_O^C = \lambda_{O_o} \begin{bmatrix} \frac{(u_{o_o} - u_0)dx}{f} & \frac{(v_{o_o} - v_0)dy}{f} & 1 \end{bmatrix}^T \quad (17)$$

Through the calculation above, the rotation matrix $\mathbf{R}_{O,k}^C$ and translation matrix $\mathbf{T}_{O,k}^C$ from the segment frame to the camera frame at moment k can be obtained. The rotation matrix and translation matrix at moment $k - 1$ is expressed as $\mathbf{R}_{O,k-1}^C$ and $\mathbf{T}_{O,k-1}^C$. So, the rotation matrix and translation matrix of the camera from moment $k - 1$ to moment k can be calculated by Equations (18) and (19), respectively.

$$\mathbf{R}_{k,k-1} = \mathbf{R}_{O,k}^C [\mathbf{R}_{O,k-1}^C]^T \quad (18)$$

$$\mathbf{T}_{k,k-1} = \mathbf{T}_{O,k}^C - \mathbf{T}_{O,k-1}^C \quad (19)$$

3. ADAPTIVE FILTER MODEL FUSING STRUCTURED LIGHT VISUAL SENSOR AND INERTIAL SENSOR. Although the pose of an MAV can be calculated by the cross-structured light visual sensor, it will still be invalid in some situations such as when the segment line is out of sight, or the intersections between structured light and the segment are unable to be detected. Taking these situations into account, the information of the Inertial Navigation System (INS) is used to integrate with the information of the SLV sensor. Although the EKF is still the standard technique for non-linear attitude estimation in practical terms, the EKF suffers from linearization error and requires more system complexities and computational time than a regular Kalman Filter (KF) (Li and Wang, 2013). For the loosely coupled approach, the indirect KF has been applied successfully by many researchers (Liu et al., 2016; Ligorio and Sabatini, 2013).

In this section, the state model based on error equations of an inertial navigation and measurement model based on the error of the relative motion measured by INS and SLV are first built. The augmented state and the measurement model are simplified, because the error analysis of the measurement model in the augmented state method is based on the true values at previous and current moments. While the measurement model is built based on the estimated values of the state at the last moment, and the estimated values of the attitude and position at the last moment can be calculated, the measurement error is only related to the position error and angle error at the present moment. An adaptive Sage-Husa Kalman filter (ASHKF) with forgetting factor matrix is designed to deal with the problem of the varying measurement noise of the SLV sensor.

3.1. *System state model.* The state vector of an integrated navigation system is:

$$\mathbf{X} = [\boldsymbol{\phi}^T \quad \delta\mathbf{v}^T \quad \delta\mathbf{P}^T \quad \boldsymbol{\varepsilon}_b^T \quad \boldsymbol{\varepsilon}_r^T \quad \nabla^T]^T \tag{20}$$

where $\boldsymbol{\phi}^T$ denotes the plane error angle, $\delta\mathbf{v}^T$ denotes the velocity error, $\delta\mathbf{P}^T$ denotes the position error, $\boldsymbol{\varepsilon}_b^T$ denotes the triaxial arbitrary constant, $\boldsymbol{\varepsilon}_r^T$ denotes the triaxial first-order Markov error of the gyroscope and ∇^T denotes the triaxial first-order Markov error of the accelerometer.

Then the system state model can be written as:

$$\dot{\mathbf{X}}_{18 \times 1} = \mathbf{F}_{18 \times 18} \mathbf{X}_{18 \times 1} + \mathbf{G}_{18 \times 9} \mathbf{W}_{9 \times 1} \tag{21}$$

where:

$$\mathbf{F}_{18 \times 18} = \begin{bmatrix} \mathbf{F}_N & \mathbf{F}_S \\ \mathbf{0}_{9 \times 9} & \mathbf{F}_M \end{bmatrix} \tag{22}$$

$$\mathbf{F}_N = \begin{bmatrix} -\boldsymbol{\omega}_{iw}^w & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ -({}_b^w \mathbf{R} \mathbf{f}^b) & -2\boldsymbol{\omega}_{iw}^w & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \tag{23}$$

$$\mathbf{F}_S = \begin{bmatrix} {}_b^w \mathbf{R} & {}_b^w \mathbf{R} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & {}_b^w \mathbf{R} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \tag{24}$$

$$\mathbf{F}_M = \text{Diag} \left[0 \quad 0 \quad 0 \quad -\frac{1}{T_{gx}} \quad -\frac{1}{T_{gy}} \quad -\frac{1}{T_{gz}} \quad -\frac{1}{T_{ax}} \quad -\frac{1}{T_{ay}} \quad -\frac{1}{T_{az}} \right] \tag{25}$$

$$\mathbf{W}_{9 \times 1} = [w_g \quad w_r \quad w_a]^T \tag{26}$$

$$\mathbf{G}_{18 \times 9} = \begin{bmatrix} {}_b^w \mathbf{R} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{9 \times 3} & \mathbf{0}_{9 \times 3} & \mathbf{0}_{9 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \tag{27}$$

where f^b represents the output of the accelerometer. ${}_b^w \mathbf{R}$ represents the transformation matrix from the body frame to the world frame. $\boldsymbol{\omega}_{iw}^w$ represents the angular velocity of the world frame in the inertial frame. T_g and T_a represent the correlation time of triaxial first-order Markov error of gyroscope and accelerometer, respectively. w_g , w_r and w_a are both white noises.

Then, the discrete-time state model can be written as:

$$\mathbf{X}_k = \boldsymbol{\Phi}_{k,k-1} \mathbf{X}_k + \boldsymbol{\Gamma}_{k-1} \mathbf{W}_{k-1} \tag{28}$$

where:

$$\boldsymbol{\Phi}_{k,k-1} = \sum_{n=0}^{\infty} [\mathbf{F}(t_k) T]^n / n! \tag{29}$$

$$\Gamma_{k-1} = \left\{ \sum_{n=0}^{\infty} [\mathbf{F}(t_k)T]^{n-1}/n! \right\} \mathbf{G}(t_k)T \tag{30}$$

and T is the sampling period.

3.2. *Measurement model.* Assume that unit quaternions $\mathbf{q}_{k-1,k}^{SLV}$ and $\mathbf{q}_{k-1,k}^{INS}$ represent the rotation from moment $k - 1$ to moment k , which correspond to the rotation matrix $\mathbf{R}_{k,k-1}^{SLV}$ and $\mathbf{R}_{k,k-1}^{INS}$ measured by SLV and INS respectively. The translation matrices $\mathbf{T}_{k,k-1}^{SLV}$ and $\mathbf{T}_{k,k-1}^{INS}$ represent the translation from moment $k - 1$ to moment k which are measured by SLV and INS, respectively. \otimes denotes the multiplication operation of the quaternion.

Then the error equations of the SLV measurements are:

$$\mathbf{q}_{k,k-1}^{SLV} = \mathbf{q}_{k,k-1} \otimes \delta \mathbf{q}_k^{SLV} = \mathbf{q}_{k,k-1} \otimes \begin{bmatrix} 1 \\ \frac{1}{2} \boldsymbol{\theta}_k^{SLV} \end{bmatrix} \tag{31}$$

$$\mathbf{T}_{k,k-1}^{SLV} = \mathbf{T}_{k,k-1} + \delta \mathbf{T}_k^{SLV} \tag{32}$$

where $\mathbf{q}_{k,k-1}$ represents the rotation from moment $k - 1$ to moment k , which correspond to the rotation matrix $\mathbf{R}_{k,k-1}$ in Equation (18). $\delta \mathbf{q}_k^{SLV}$ and $\boldsymbol{\theta}_k^{SLV}$ represent the measurement error of $\mathbf{q}_{k,k-1}$ in the form of quaternion and Euler angle respectively. \otimes is the product sign of the quaternion. The translation matrix $\mathbf{T}_{k,k-1}$ represents the translation from moment $k - 1$ to moment k . $\delta \mathbf{T}_k^{SLV}$ represents the error of $\mathbf{T}_{k,k-1}$ measured by the SLV.

For convenience, assume that the IMU frame overlaps the body frame, the quaternion ${}^b_w \mathbf{q}^{INS}$ represents the rotation from the world frame to the body frame which is measured by the INS. It can be represented based on the operation rules of the quaternion, which is shown in Equation (33):

$${}^b_w \mathbf{q}_k^{INS} = {}^b_w \mathbf{q}_k \otimes \delta \mathbf{q}_k^{INS} = {}^b_w \mathbf{q}_k \otimes \begin{bmatrix} 1 \\ \frac{1}{2} \boldsymbol{\Phi} \end{bmatrix} \tag{33}$$

where $\boldsymbol{\Phi}$ is used to describe the rotation vector.

The quaternion $\mathbf{q}_{k,k-1}^{INS}$ can be obtained by Equation (34):

$$\begin{aligned} \mathbf{q}_{k,k-1}^{INS} &= {}^w_b \hat{\mathbf{q}}_{k-1} \otimes {}^b_w \mathbf{q}_k^{INS} \\ &\approx {}^w_b \mathbf{q}_{k-1} \otimes {}^b_w \mathbf{q}_k \otimes \delta \mathbf{q}_k^{INS} \\ &= \mathbf{q}_{k,k-1} \otimes \delta \mathbf{q}_k^{INS} \end{aligned} \tag{34}$$

where ${}^w_b \hat{\mathbf{q}}_{k-1}$ is the estimation of ${}^w_b \mathbf{q}$ at the last moment.

The error equations of $\mathbf{T}_{k,k-1}^{INS}$ are:

$$\begin{aligned} \mathbf{T}_{k,k-1}^{INS} &= {}^b_w \hat{\mathbf{R}}_{k-1} (\mathbf{P}_k^{INS} - \hat{\mathbf{P}}_{k-1}) \\ &= {}^b_w \hat{\mathbf{R}}_{k-1} (\mathbf{P}_k + \delta \mathbf{P}_k - \hat{\mathbf{P}}_{k-1}) \\ &= {}^b_w \hat{\mathbf{R}}_{k-1} (\mathbf{P}_k - \hat{\mathbf{P}}_{k-1}) + {}^b_w \hat{\mathbf{R}}_{k-1} \delta \mathbf{P}_k \\ &\approx \mathbf{T}_{k,k-1} + {}^b_w \hat{\mathbf{R}}_{k-1} \delta \mathbf{P}_k \end{aligned} \tag{35}$$

where, ${}^b_w \hat{\mathbf{R}}_{k-1}$ represents the estimated rotation from the world frame to the body frame at moment $k - 1$. \mathbf{P}_k^{INS} represents the position in the world frame measured by INS at moment k . $\hat{\mathbf{P}}_{k-1}$ represents the estimated position in the world frame at moment $k - 1$.

In Figure 2(a), it can be seen that CO is the height between the MAV and floor, and it can be calculated through Equation (14). Therefore, the height in $\hat{\mathbf{P}}_{k-1}$ can be replaced by the value of CO . Error accumulation of height can be eliminated by this approach.

Then, the measurement error between INS and SLV sensors can be represented as:

$$\begin{aligned}
 (\mathbf{q}_{k,k-1}^{SLV})^T \otimes \mathbf{q}_{k,k-1}^{INS} &= (\delta \mathbf{q}_k^{SLV})^T \otimes (\mathbf{q}_{k,k-1})^T \otimes \mathbf{q}_{k,k-1} \otimes \delta \mathbf{q}_k^{INS} \\
 &= \begin{bmatrix} 1 \\ -\frac{1}{2} \boldsymbol{\theta}_k^{SLV} + \frac{1}{2} \boldsymbol{\Phi} - \frac{1}{4} \boldsymbol{\theta}_k^{SLV} \times \boldsymbol{\Phi} \end{bmatrix} \quad (36)
 \end{aligned}$$

On the other hand, we denote $(\mathbf{q}_{k,k-1}^{SLV})^T \otimes \mathbf{q}_{k,k-1}^{INS} = [z_0 \ z_1 \ z_2 \ z_3]^T$. There is $z_0 \approx 1$ because the angle corresponding to $(\mathbf{q}_{k,k-1}^{SLV})^T \otimes \mathbf{q}_{k,k-1}^{INS}$ is tiny. Ignoring the second order small quantity, the following equations can be obtained:

$$\mathbf{z}_{r,k} = [z_1 \ z_2 \ z_3]^T = \frac{1}{2} \boldsymbol{\Phi} - \frac{1}{2} \boldsymbol{\theta}_k^{SLV} \quad (37)$$

$$\mathbf{z}_{t,k} = \mathbf{T}_{k,k-1}^{INS} - \mathbf{T}_{k,k-1}^{SLV} = \frac{b}{w} \hat{\mathbf{R}}_{k-1} \delta \mathbf{P}_k - \delta \mathbf{T}_k^{SLV} \quad (38)$$

Finally, the measurement model is:

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{z}_{r,k} \\ \mathbf{z}_{t,k} \end{bmatrix} = \mathbf{H}_k \mathbf{X}_k + \mathbf{V}_k \quad (39)$$

where the measurement matrix \mathbf{H}_k and measurement noise matrix $\mathbf{V}_{18 \times 1}$ are:

$$\mathbf{H}_k = \begin{bmatrix} \frac{1}{2} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \frac{b}{w} \hat{\mathbf{R}}_{k-1} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \quad (40)$$

$$\mathbf{V}_k = \begin{bmatrix} -\frac{1}{2} \boldsymbol{\theta}_k^{SLV} & \mathbf{0}_{3 \times 3} & \delta \mathbf{T}_k^{SLV} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix}^T \quad (41)$$

3.3. *Adaptive Sage-Husa Kalman filter based on multiple forgetting factor.* The statistical property of noise is assumed to be known in a regular Kalman Filter. In this paper, the measurement noise $\boldsymbol{\theta}_k^{SLV}$ and $\delta \mathbf{T}_k^{SLV}$ vary with the measurement distance of the SLV sensor (Wang et al., 2015). In this case, the measurement covariance matrix \mathbf{U} is uncertain. \mathbf{U} influences the weight that the filter applies between the existing process information and the latest measurements. Errors in \mathbf{U} may result in the filter being suboptimal or even cause it to diverge (Ding et al., 2007). Thus, the adaptive filter which can adjust \mathbf{U} online is significant in improving the stability and accuracy of the integrated navigation system.

Due to the fact that only the statistical property of measurement noise is varying in the integrated navigation system in this study, a simplified Adaptive Sage-Husa Kalman Filter (ASHKF) is used. The ASHKF has a simple structure and high efficiency so that it is widely used in engineering (Sun et al., 2016). A time-varying noise estimator with a forgetting factor matrix is used in the ASHKF. The ordinary forgetting factor has the same attenuation to each state variable. It is hard to ensure that each estimated accuracy of each state variable is optimal. To solve this problem, a forgetting factor matrix is used here to set different forgetting factors for the different state variables.

The process of the filter is as follows:

(1) Time update:

$$\begin{cases} \hat{\mathbf{X}}_{k,k-1} = \Phi_{k,k-1} \hat{\mathbf{X}}_{k-1} \\ \mathbf{P}_{k,k-1} = \Phi_{k,k-1} \mathbf{P}_{k-1} \Phi_{k,k-1}^T + \Gamma_{k-1} \mathbf{Q}_{k-1} \Gamma_{k-1}^T \end{cases} \quad (42)$$

(2) Measurement update:

$$\begin{cases} \tilde{\mathbf{Z}}_k = \mathbf{Z}_k - \mathbf{H}_k \hat{\mathbf{X}}_{k,k-1} \\ \mathbf{X}_k = \mathbf{X}_{k,k-1} + \mathbf{K}_k \tilde{\mathbf{Z}}_k \\ \mathbf{K}_k = \mathbf{P}_{k,k-1} \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_{k,k-1} \mathbf{H}_k^T + \hat{\mathbf{U}}_k]^{-1} \\ \mathbf{P}_k = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_{k,k-1} [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k]^T + \mathbf{K}_k \hat{\mathbf{U}}_k \mathbf{K}_k^T \end{cases} \quad (43)$$

where $\hat{\mathbf{U}}$ is the estimated value of the systematic measurement noise covariance matrix.

$$\hat{\mathbf{U}}_k = \sqrt{\mathbf{I} - \boldsymbol{\beta}} \hat{\mathbf{U}}_{k-1} \sqrt{\mathbf{I} - \boldsymbol{\beta}} + \sqrt{\boldsymbol{\beta}} (\tilde{\mathbf{Z}}_k \tilde{\mathbf{Z}}_k^T - \mathbf{H}_k \mathbf{P}_{k,k-1} \mathbf{H}_k^T) \sqrt{\boldsymbol{\beta}} \quad (44)$$

$$\boldsymbol{\beta} = \text{diag}(\beta_1, \beta_2, \dots, \beta_{18}); 0 \leq \beta_i \leq 1, i = 1, 2, \dots, 18 \quad (45)$$

where $\boldsymbol{\beta}$ is the forgetting factor matrix, which can be obtained by optimising the selected method. In the integrated navigation system, $\hat{\mathbf{U}}$ can be adaptively estimated based on the varying noise estimator according to Equations (44) and (45).

4. SIMULATIONS AND ANALYSES. Simulation parameters, a flight trajectory and simulation flow are designed in a simulation environment. Then, five navigation methods are compared to verify the proposed method.

4.1. *Simulation environment setup.* A 1 m interval grid was set on the floor to simulate the floor in the real environment. The intersections of structured light and the grid were captured to calculate the attitude and position by SLV. The flight path of an MAV is shown in Figure 3. The original longitude, latitude and height are 110° , 20° and 10 m respectively. The time span of the flight is 300 s. The accuracy of the gyroscopes is $1^\circ/\text{h}$. The accuracy of the accelerometers is 10^{-3} g. The distance between the structured light sensor and the visual sensor is 200 mm. The image resolution is 640×480 pixels. The coordinate deviation of the feature points is $[-0.5, 0.5]$. The sampling rates of INS and SLV are 50 Hz and 10 Hz respectively. The R matrix in KF is set as $(\text{diag}[0.3 \ 1.5 \ 1.5 \ 0.003 \ 0.003 \ 0.007])^2$, which has been analysed in our previous work (Wang et al., 2015). The sequence of elements in R matrix is roll/ $^\circ$, pitch/ $^\circ$, head/ $^\circ$, longitude/m latitude/m, height/m. For simplicity, the calibration deviations of structured light planes are ignored.

The simulation flow is shown in Figure 4(a). The true trajectory was generated by the trajectory generator and the true navigation data was sent to the IMU simulator and cross-structured light simulator. The angular speed and acceleration with their errors were sent to the INS to calculate the state of the MAV. The cross-structured light was generated according the navigation data by the cross-structured light simulator. The image coordinates of the intersections between the structured light and grid were calculated by visual sensor imaging simulation. The simulation details of the cross-structured light and images are shown in

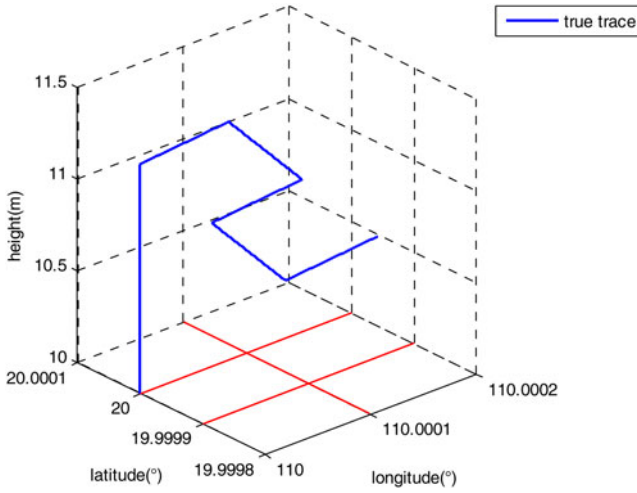


Figure 3. The simulated MAV trajectory.

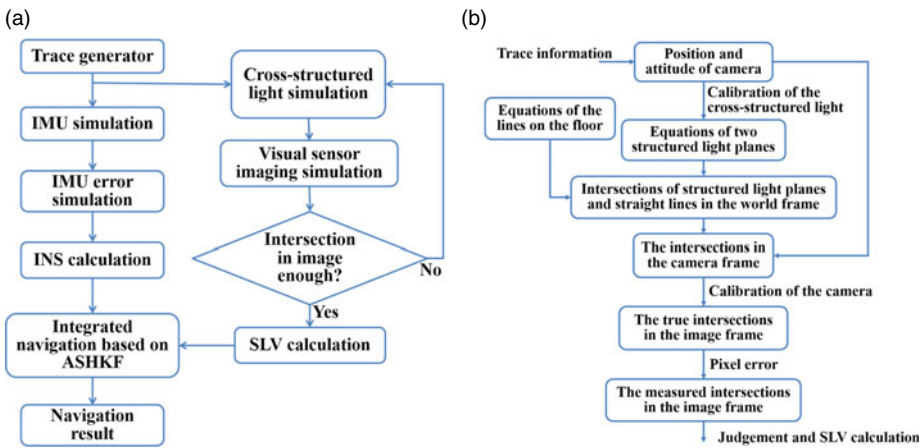


Figure 4. The simulation flow. (a) The total simulation flow. (b) The simulation flow of SLV sensor.

Figure 4(b). The equations of the lines on the floor are fixed. The parameters of the structured light planes change with the camera. Then the intersections of the structured light planes and the lines on the floor can be calculated. According to the position and attitude of the camera in the world frame, the coordinates of the intersections in the camera frame can be obtained. According to the geometry principle, we can check these points and obtain the non-collinear intersections. If there are three non-collinear intersections, the attitude and position will be calculated according to the measurement data. If not, the image of the next moment will be waited for. Then the state of MAV and measurement data are integrated by ASHKF, and the navigation results are supplied as the initial value for the next moment.

4.2. Comparison of different navigation methods. To obtain the values of the forgetting factor matrix, the navigation results with different forgetting factors are compared in

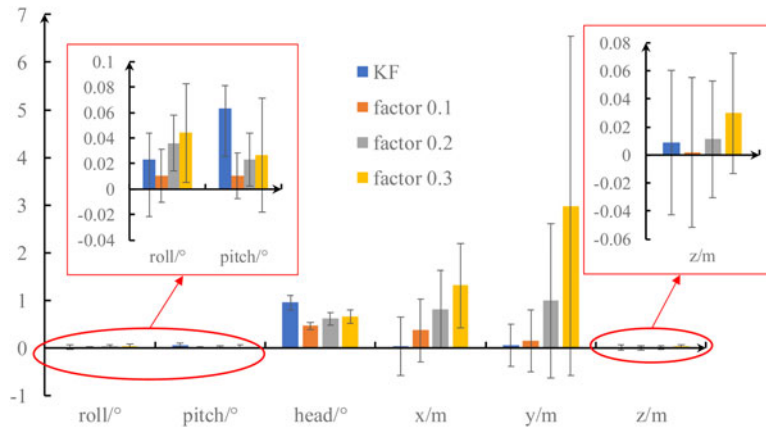


Figure 5. Mean error and RMSE of KF and ASHKF with different forgetting factors.

this part. Figure 5 shows the mean error and Root Mean Square Error (RMSE) of KF and ASHKF with different forgetting factors.

In Figure 5, if the forgetting factor is 0, the ASHKF is changed to a KF. The results of forgetting factors bigger than 0.3 are hidden because of their huge errors. It can be seen from Figure 5 that ASHKF with forgetting factor 0.1 has lower errors than the KF in the roll, pitch, heading and height channels. However, KF has the least error in the horizontal position. It can also be seen that ASHKF with forgetting factor 0.1 has better performance than ASHKF with forgetting factor 0.2 and 0.3. In this case, the values of the forgetting factor matrix β can be obtained by optimising the parameters, and the forgetting factor matrix in Equation (45) can be set as: $\beta_1 = \beta_2 = \beta_3 = \beta_6 = 0.1$; the others are zero.

Five types of navigation results based on different navigation methods are compared in Figures 6 and 7. “INS only” means that only the inertial navigation system is used in the navigation solution. “SLV only” means that only the structured light visual system is used in the navigation solution. “KF” means that the Kalman Filter is used in integrated navigation. “Forgetting factor 0.1” means that the ASHKF with forgetting factor 0.1 is used in the integrated navigation. The “Forgetting factor matrix” means that the ASHKF with the former forgetting factor matrix was used in the integrated navigation. The time synchronised method was used in the KF and the ASHKF. Due to the errors of different cases having huge differences, the results of the single navigation systems with large errors are shown in Figure 6, and the results of different filters with small errors are shown in Figure 7.

As shown in Figure 6, the attitude error is very large if only the SLV system is used in the navigation solution. The position error has serious divergence if pure INS is used in the navigation solution. These two types of systems are difficult to use alone for navigation. The results of three types of filters are shown in Figure 7. Better accuracy can be obtained by integrating the two sensors, especially in the height channel. Due to the fact that the height between the MAV and the floor can be calculated directly by the SLV, there is no error accumulation in the height channel. Compared with KF, ASHKF with forgetting factor matrix has better accuracy in the roll, pitch, heading and height channels. Comparing with ASHKF with forgetting factor 0.1, ASHKF with forgetting factor matrix has better

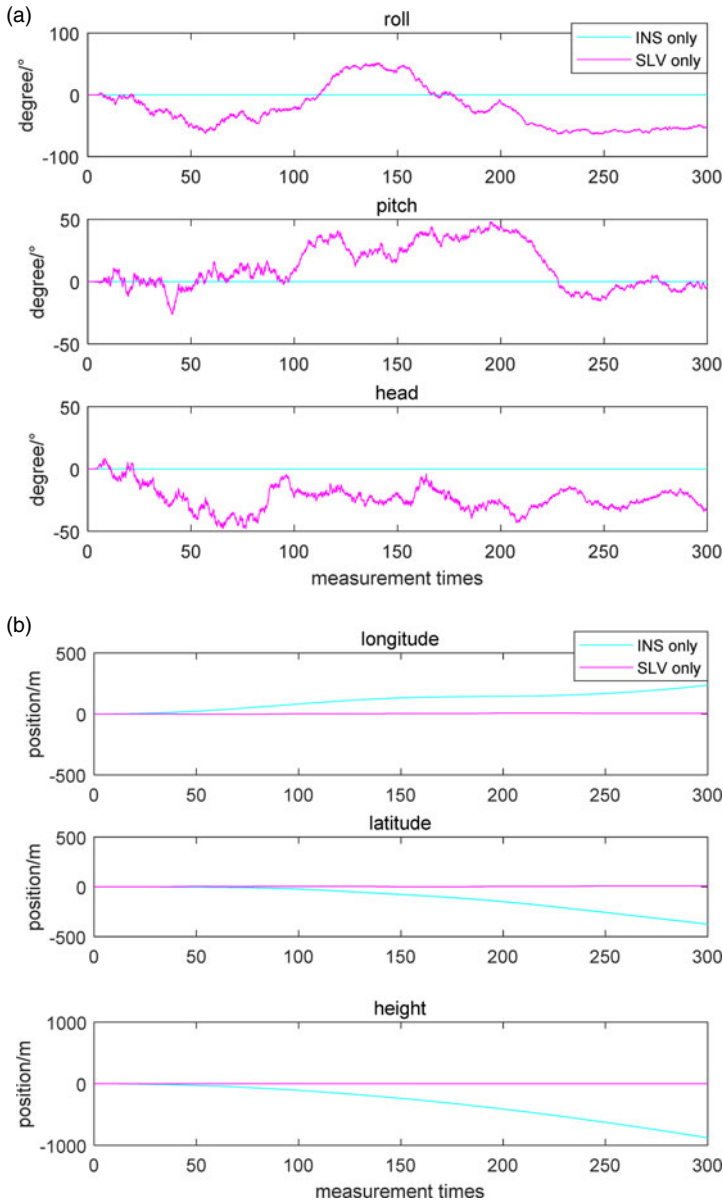


Figure 6. The navigation results of two cases. (a) The comparison of attitude for single navigation systems. (b) The comparison of position for single navigation systems.

accuracy in the longitude and latitude channels. The errors of ASHKF with forgetting factor matrix in three attitude channels can be limited in the range of $(-0.2^\circ, 0.2^\circ)$, the errors of ASHKF with forgetting factor matrix in two horizontal position channels can be limited in the range of $(-0.5\text{ m}, 0.5\text{ m})$. The errors of ASHKF with forgetting factor matrix in height channel can be limited in the range of $(-0.2\text{ m}, 0.2\text{ m})$.

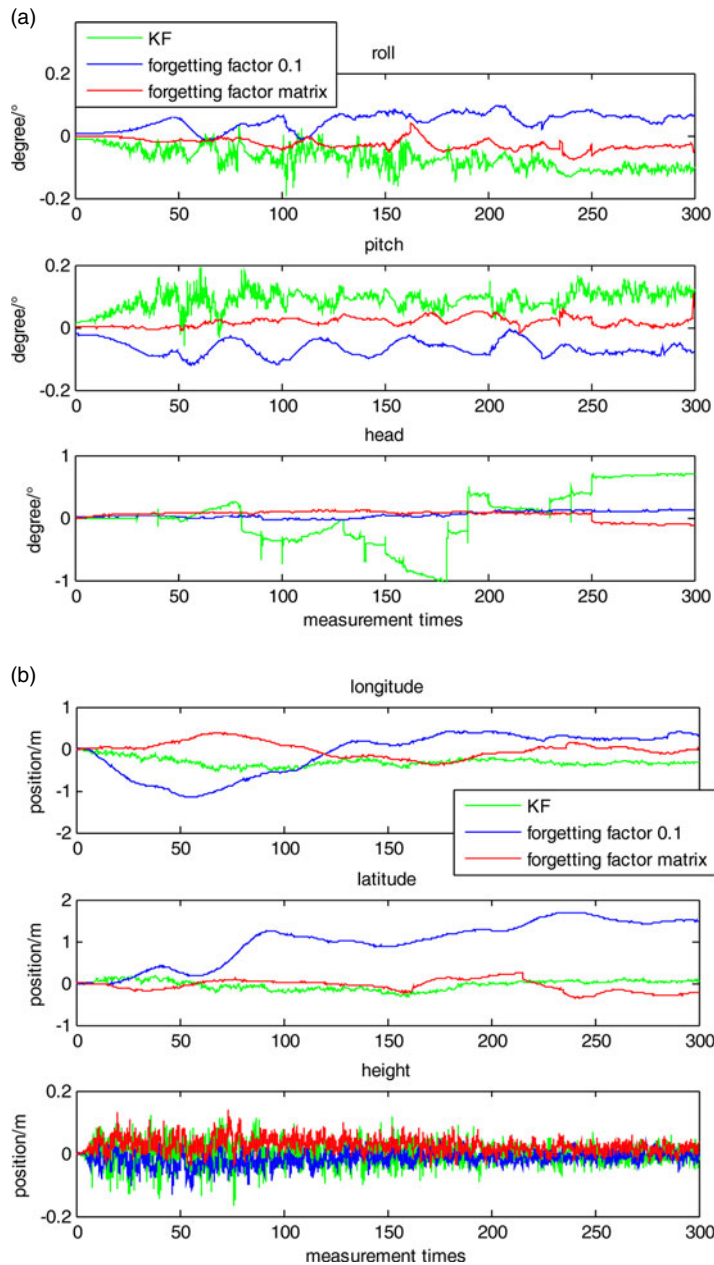


Figure 7. The navigation results of three cases. (a) The comparison of attitude for three cases. (b) The comparison of position for three cases.

5. EXPERIMENT AND ANALYSIS BASED ON REAL DATA. An experiment based on real data was carried out on an independently designed quad rotor test bed. The quad rotor is shown in Figure 8(a). The precision of the gyroscope is $0.01^\circ/\text{s}$ and the precision



Figure 8. The environment of verification. (a) The quad rotor test bed designed independently. (b) The *easy track* motion capture system. (c) The screenshot of motion capture system software.

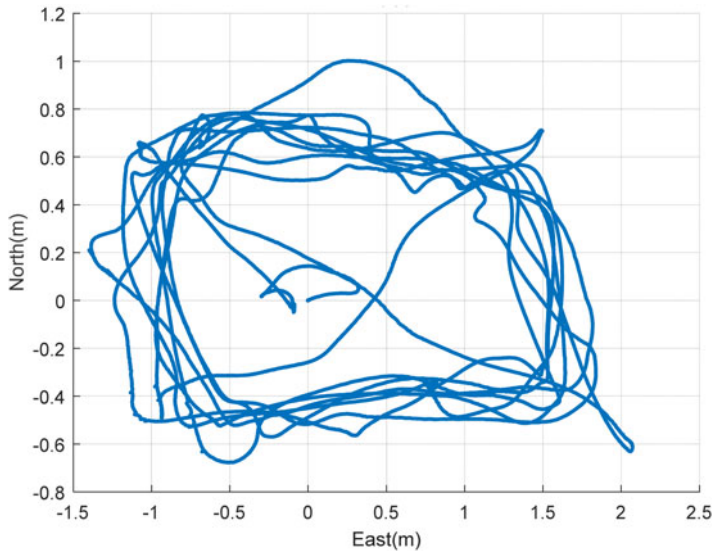


Figure 9. The trajectory of dynamic flight from top view.

of the accelerometer is 0.05 g . The downward camera (resolution: 640×480 pixels) and cross-structured light sensor (power: 5 mW , wavelength: 650 nm) were mounted underneath the quad rotor. The standard navigation information was provided by the *easy track* motion capture system, which is shown in Figure 8(b) and Figure 8(c). The markers on the quad rotor were captured by *easy track* and the attitude and position of the quad rotor in the motion capture system frame can be calculated. The errors of attitude and position measured by *easy track* are less than 0.05° and 0.03 mm , respectively. The sampling rates of INS and the motion capture system are both 50 Hz , and the sampling rate of the SLV is about 10 Hz . The period of filter interval is 0.1 s . The attitude and position in the motion capture system frame can be transformed to the navigation frame by calibration.

The quad rotor was flying in the motion capture system. The flight times for the hovering test and the dynamic flight test were 60 s and 120 s respectively. The trajectory of dynamic flight is shown in Figure 9. The frame in Figure 8 is East-North-Up (ENU) and the centre of the frame is at coordinate $(0, 0, 0)$. One of the images captured by the onboard camera is shown in Figure 10(a). Then the line is extracted by edge detection and Hough transform

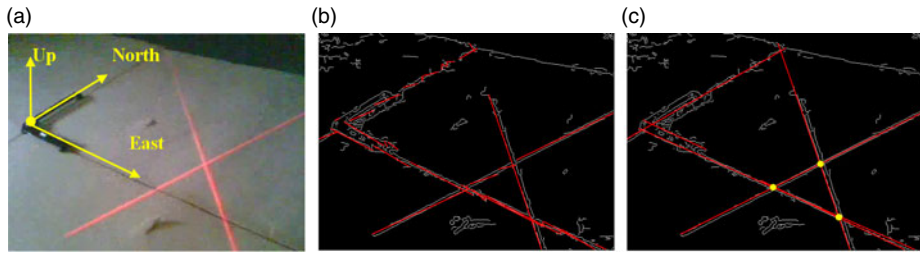


Figure 10. Lines extraction and intersections detection. (a) The image captured by onboard camera. (b) The original lines extraction. (c) The lines enhancing and intersections detection.

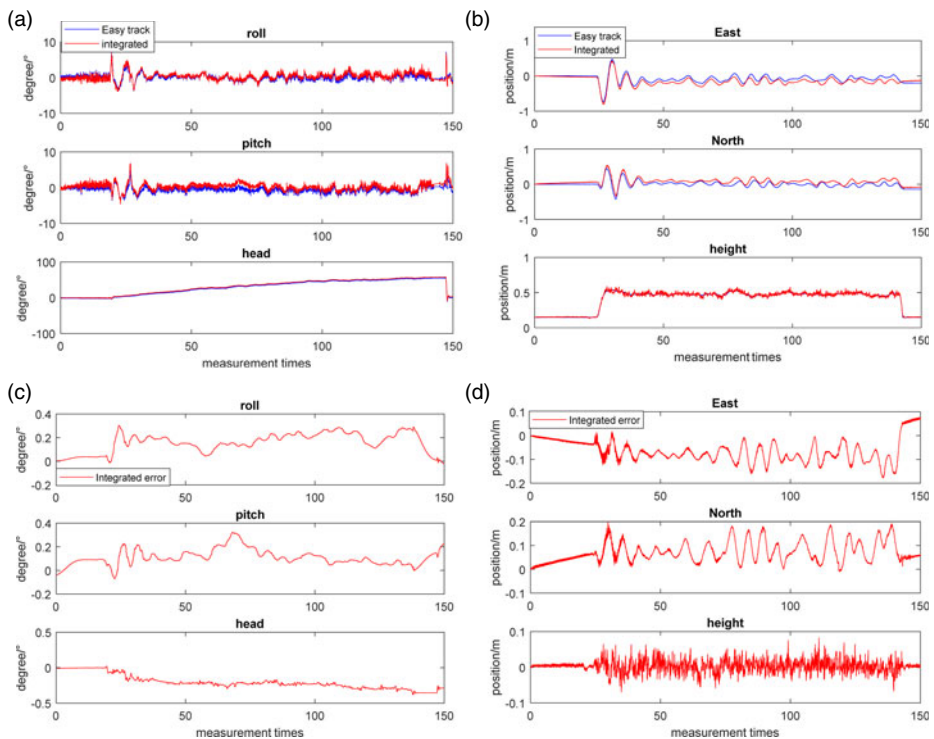


Figure 11. Comparison of hover test between easy track and the proposed method. (a) Attitude comparison of hover test. (b) Position comparison of hover test. (c) Attitude error of hover test. (d) Position error of hover test.

(Fernandes and Oliveira, 2008). As Figure 10(b) shows, the cross-structured light and segment lines are disconnected in many places. Thus, the intersections of structured light and segment lines cannot be detected. If the lines have the same slope and the distances are within a certain threshold, they will be combined as one line. The result of improved line detection is shown in Figure 10(c).

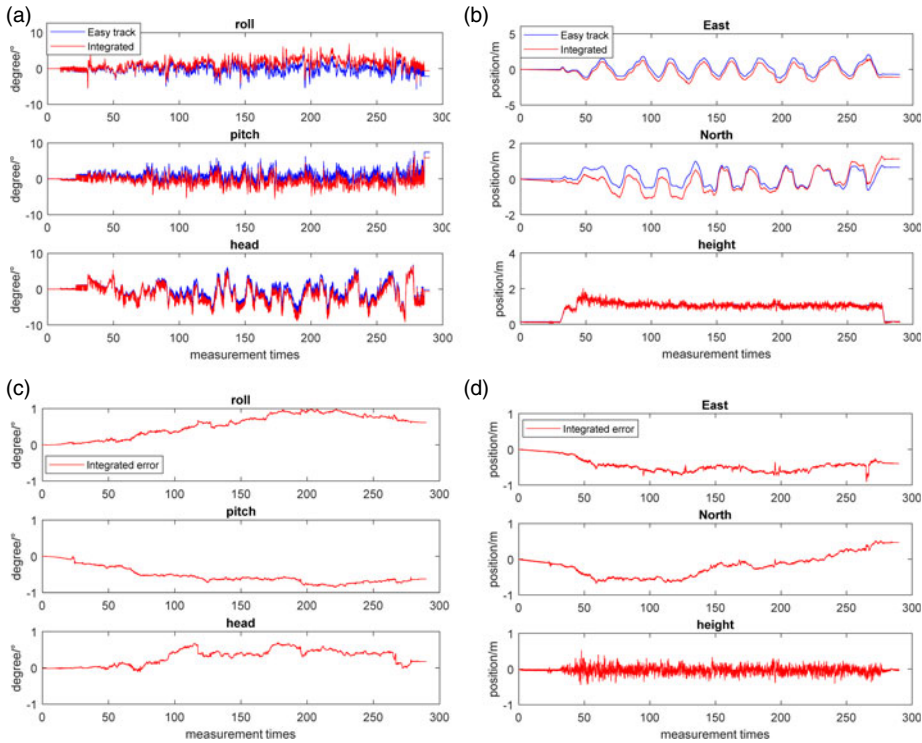


Figure 12. Comparison of dynamic flight test between easy track and the proposed method. (a) Attitude comparison of dynamic flight test (b) Position comparison of dynamic flight test (c) Attitude error of dynamic flight test (d) Position error of dynamic flight test

The coordinates of the intersections in the image frame can be calculated as follows:

- (1) The brightness of each line is analysed, and the two brightest intersecting lines are recognised as cross-structured light;
- (2) The coordinates of intersection points of cross-structured light in the image frame are calculated;
- (3) The closest line to the intersection point of cross structured light is chosen and the intersection points of the line and cross-structured light are found;
- (4) If there are two intersection points in step (3), the relative attitude can be calculated, and the endpoint of the line will be recognised as the origin of the object frame to calculate the relative position;
- (5) If there are less than two intersection points in step (3), the next closest line will be chosen. If there are no lines that have two intersection points with cross-structured light in the image, the pose information of the last measurement will be recognised as the output of this measurement.

Navigation results of the hover test are shown in Figure 11. The curve “Easy track” shows the navigation solutions calculated by *easy track*. The curve “integrated” shows the data calculated by the proposed method. It can be seen that the errors of attitude are less than 0.5° and the errors of position are less than 0.2 m . In addition, the error of height

Table 1. The Root Mean Squared Error (RMSE) of the navigation data.

	Roll/ $^{\circ}$	Pitch/ $^{\circ}$	Head/ $^{\circ}$	East/m	North/m	Height/m
Hover test	0.2231	0.1315	0.3884	0.1587	0.1566	0.0481
Dynamic flight test	0.2376	0.1857	0.4794	0.4310	0.3423	0.1044

is not accumulated while the others are both accumulated. This is because the height is measured by the SLV sensor directly in each measurement period. The advantage will be more obvious in the dynamic flight test (Figure 12).

Figure 12 shows the results of the dynamic flight test. Images captured in the dynamic flight are more complex than the images captured in the hover test, such as the overlap between structured light and the segment on the floor or the endpoint of the segment which is unable to be detected. These situations will generate saltation or lag on the measurement such as at about 200 s and 270 s in Figure 12(d). The attitude is influenced little because the INS can continuously provide accurate attitude data. In this case, the errors of attitude are less than 1° while the errors of position are increased, and the error of height is much less than the errors of east and north. A more detailed comparison is shown in Table 1.

The RMSE of the navigation solutions are shown in Table 1. The accuracy of attitude in both the hover test and dynamic flight test is higher than 1° . The accuracy of east and north in the hover test is about 0.15 m while it is about 0.4 m in the dynamic flight test. The accuracy of height is about 0.05 m in the hover test while it is about 0.1 m in the dynamic flight test. It is difficult to achieve such an accuracy in height measurement by the traditional integrated navigation algorithm.

6. CONCLUDING REMARKS. To solve the problem that traditional visual navigation methods are difficult to utilise in insufficient illumination and sparse feature environments, an active visual inertial navigation method based on a structured light vision sensor for MAVs is proposed in this paper. Based on the analysis of the characteristics of the structured light vision sensor, a measurement model for the sensor is built. With the help of the error equation of an inertial navigation system, the relative motion error measurement model of inertial navigation system and SLV is built. This model is mainly related to the position and attitude information at the current moment, which is helpful to decrease the error accumulation of a traditional visual inertial navigation system. According to the statistical properties of the SLV, an Adaptive Sage-Husa Kalman Filter (ASHKF) with forgetting factor matrix is designed. The forgetting factor matrix is used to set different forgetting factors for different stated variables, which is helpful for solving the problem of measuring noise changing with aircraft height.

Simulation results show that this method has better performance than KF and ASHKF. The errors of ASHKF with forgetting factor matrix in three attitude channels, two horizontal position channels and height channel can be limited in the range of $(-0.2^{\circ}, 0.2^{\circ})$, $(-0.5 \text{ m}, 0.5 \text{ m})$ and $(-0.2 \text{ m}, 0.2 \text{ m})$, respectively. The results of the experiments indicate that the attitude accuracy is higher than 1° , location accuracy in north or east is about 0.4 m, while the accuracy of height is about 0.1 m. The requirements of MAV navigation in an indoor environment can be satisfied by the work presented in this paper.

ACKNOWLEDGEMENT

This paper is supported by the Fundamental Research Funds for the Central Universities (No. NS2018021).

REFERENCES

- Bouguet, J. Y. (2015). Camera calibration toolbox for MATLAB, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.
- Cocchioni, F., Mancini, A. and Longhi, S. (2014). Autonomous navigation, landing and recharge of a quadrotor using artificial vision. Proceedings of the 2014 IEEE International Conference on Unmanned Aircraft Systems (ICUAS), 418–429, Orlando, FL, USA.
- Ding, W., Wang, J., Rizos, C. and Kinlyside, D. (2007). Improving adaptive Kalman estimation in GPS/INS integration. *The Journal of Navigation*, **60**(3), 517–529.
- Dougherty, J., Lee, D. and Lee, T. (2014). Laser-based guidance of a quadrotor UAV for precise landing on an inclined surface. *American Control Conference (ACC)*, 1210–1215, Portland, OR, USA.
- Engel, J., Sturm, J. and Cremers, D. (2014). Scale-aware navigation of a low-cost quadcopter with a monocular camera. *Robotics and Autonomous Systems*, **62**(11), 1646–1656.
- Eschmann, C., Kuo, C. M., Kuo, C. H. and Boller, C. (2012). Unmanned aircraft systems for remote building inspection and monitoring. *6th European workshop on structural health monitoring*, 1–8, Dresden, Germany.
- Fernandes, L. A., and Oliveira, M. M. (2008). Real-time line detection through an improved Hough transform voting scheme. *Pattern Recognition*, **41**(1), 299–314.
- Gao, X. and Zhang, T. (2015). Robust RGB-D simultaneous localization and mapping using planar point features. *Robotics and Autonomous Systems*, **72**, 1–14.
- Gioia, C. and Borio, D. (2014). Stand-alone and hybrid positioning using asynchronous pseudolites. *Sensors*, **15**(1), 166–193.
- Harmat, A. and Sharf, I. (2014). Towards full omnidirectional depth sensing using active vision for small unmanned aerial vehicles. *Proceedings of the 2014 IEEE Canadian Conference on Computer and Robot Vision (CRV)*, 24–31, Montreal, QC, Canada.
- Huang, G., Kaess, M., and Leonard, J. J. (2014). Towards consistent visual-inertial navigation. *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA)*, 4926–4933, Hong Kong, China.
- Kelly, J. and Sukhatme, G. S. (2011). Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. *The International Journal of Robotics Research*, **30**(1), 56–79.
- Kiddee, P., Fang, Z. and Tan, M. (2016). A practical and intuitive calibration technique for cross-line structured light. *Optik-International Journal for Light and Electron Optics*, **127**(20), 9582–9602.
- Kong, X., Wu, W., Zhang, L. and Wang, Y. (2015). Tightly-coupled stereo visual-inertial navigation using point and line features. *Sensors*, **15**(6), 12816–12833.
- Li, W. and Wang, J. (2013). Effective adaptive Kalman filter for MEMS-IMU/magnetometers integrated attitude and heading reference systems. *The Journal of Navigation*, **66**(1), 99–113.
- Ligorio, G. and Sabatini, A. M. (2013). Extended Kalman filter-based methods for pose estimation using visual, inertial and magnetic sensors: comparative analysis and performance evaluation. *Sensors*, **13**(2), 1919–1941.
- Liu, C., Zhou, F., Sun, Y., Di, K. and Liu, Z. (2012). Stereo-image matching using a speeded up robust feature algorithm in an integrated vision navigation system. *The Journal of Navigation*, **65**(4), 671–692.
- Liu, S., Atia, M. M., Karamat, T. B. and Noureldin, A. (2015). A LiDAR-Aided Indoor Navigation System for UGVs. *The Journal of Navigation*, **68**(02), 253–273.
- Liu, Y., Xiong, R., Wang, Y., Huang, H., Xie, X., Liu, X. and Zhang, G. (2016). Stereo Visual-Inertial Odometry with Multiple Kalman Filters Ensemble. *IEEE Transactions on Industrial Electronics*, **63**(10), 6205–6216.
- Meng Q., Liu J., Zeng Q., Feng, S. and Chen R. (2017). Neumann-Hoffman Code Evasion and Stripping Method For BeiDou Software-defined Receiver. *The Journal of Navigation*, **70**(1), 101–119.
- Natraj, A., Ly, D. S., Eynard, D., Demonceaux, C. and Vasseur, P. (2013). Omnidirectional vision for UAV: Applications to attitude, motion and altitude estimation for day and night conditions. *Journal of Intelligent & Robotic Systems*, **69**(1–4), 459–473.
- Sanahuja, G. and Castillo, P. (2013). Embedded Laser Vision System for Indoor Aerial Autonomous Navigation. *Journal of Intelligent & Robotic Systems*, **69**(1–4), 447–457.
- Santoso, F., Garratt, M. A. and Anavatti, S. G. (2017). Visual-Inertial Navigation Systems for Aerial Robotics: Sensor Fusion and Technology. *IEEE Transactions on Automation Science and Engineering*, **14**(1), 260–275.

- Sun, J., Xu, X., Liu, Y., Zhang, T. and Li, Y. (2016). FOG Random Drift Signal Denoising Based on the Improved AR Model and Modified Sage-Husa Adaptive Kalman Filter. *Sensors*, **16**(7), 1073.
- Tarchi, D., Vespe, M., Gioia, C., Sermi, F., Kyovtorov, V. and Guglieri, G. (2017). Low-cost mini radar: design prototyping and tests. *Journal of Sensors*, Volume 2017, Article ID 8029364, 1–15.
- Wang, T., Wang, C., Liang, J., Chen, Y. and Zhang, Y. (2013). Vision-aided inertial navigation for small unmanned aerial vehicles in GPS-denied environments. *International Journal of Advanced Robotic Systems*, **10**(6), 276.
- Wang, Y. S., Liu, J. Y., Zeng, Q. H. and Liu, S. (2015). Visual pose measurement based on structured light for MAVs in non-cooperative environments. *Optik-International Journal for Light and Electron Optics*, **126**(24), 5444–5451.
- Wei, Z., Zhou, F. and Zhang, G. (2005). 3D coordinates measurement based on structured light sensor. *Sensors and Actuators A: Physical*, **120**(2), 527–535.
- Weiss, S., Achtelik, M. W., Lynen, S., Chli, M. and Siegwart, R. (2012). Real-time onboard visual-inertial state estimation and self-calibration of MAVs in unknown environments. *Proceedings of 2012 IEEE International Conference on Robotics and Automation (ICRA)*, 957–964, Saint Paul, MN, USA.
- Xian, Z., Hu, X. and Lian, J. (2015). Fusing stereo camera and low-cost inertial measurement unit for autonomous navigation in a tightly-coupled approach. *The Journal of Navigation*, **68**(3), 434–452.
- Zeng, Q. H., Wang, Y. S., Liu, J. Y. and Liu, S. (2017) A matching algorithm for large viewpoint changes images[J]. *Optik - International Journal for Light and Electron Optics*, **137** (2017) 268–278.
- Zhang, C. and Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, **13**(6), 693–712.
- Zhang, X., Zhan, X. Q., Feng, S. J. and Ochieng, W. (2017). An Analytical Model for BDS B1 Spreading Code Self-Interference Evaluation Considering NH Code Effects. *Sensors*, **17**(4), 663–684.