


Article

Laser Ranging-Assisted Binocular Visual Sensor Tracking System

Qilong Wang ^{1,*} , Yu Zhang ², Weichao Shi ¹ and Meng Nie ³

¹ School of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an 710048, China; shiweichao@xaut.edu.cn

² Beijing Aerospace Times Optical-electronic Technology CO, Ltd., China Aerospace Science and Technology Corp, Beijing 100094, China; zhangyu2006101@126.com

³ School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China; nmbc@bjtu.edu.cn

* Correspondence: wangql@xaut.edu.cn or 13116329@bjtu.edu.cn; Tel.: +86-1336-668-8696

Received: 25 December 2019; Accepted: 22 January 2020; Published: 27 January 2020



Abstract: Aimed at improving the low measurement accuracy of the binocular vision sensor along the optical axis in the process of target tracking, we proposed a method for auxiliary correction using a laser-ranging sensor in this paper. In the process of system measurement, limited to the mechanical performance of the two-dimensional turntable, the measurement value of a laser-ranging sensor is lagged. In this paper, the lag information is updated directly to solve the time delay. Moreover, in order to give full play to the advantages of binocular vision sensors and laser-ranging sensors in target tracking, federated filtering is used to improve the information utilization and measurement accuracy and to solve the estimated correlation. The experimental results show that the real-time and measurement accuracy of the laser ranging-assisted binocular visual-tracking system is improved by the direct update algorithm and the federal filtering algorithm. The results of this paper are significant for binocular vision sensors and laser-ranging sensors in engineering applications involving target tracking systems.

Keywords: binocular vision; laser ranging sensor; target tracking; information fusion; time delay

1. Introduction

Visual measurement has many advantages, such as high accuracy and a non-contact nature. It is widely used in industrial and military applications and daily life [1–3]. According to the different number of cameras used in the measurement process, visual measurement techniques can be generally divided into monocular vision, binocular vision, and multieye vision [4–6].

Monocular vision has the problem of scale ambiguity. In the process of solving the corresponding points in the monocular visual image, the fundamental matrix or the homography matrix lacks the depth constraint in the decomposition and cannot determine the proportional coefficient [7]. Compared with monocular vision, stereo vision consists of two or more fixed vision sensors that collect target data simultaneously. By establishing the correspondence between views, the scale information can be estimated quickly, and then the depth recovery and the reconstruction and pose estimation of the target can be realized in the Euclidean space. In reference [8], the three-dimensional coordinates of the moving object are obtained by two cameras that measure the spatial constraints of the same spatial point between different image planes. After the 3D a reconstruction of single acquisition, the absolute orientation of the two acquisition point clouds is solved to achieve the target pose estimation and point cloud fusion. Terui. F et al. [9] used binocular vision to estimate the pose of a known semi-cooperative target and verified its effectiveness in a ground test. However, this method is computationally complex

and needs to be processed offline. Du. X et al. [10] explored the feasibility of binocular vision to measure non-cooperative targets with the support of European Space Agency, and then Segal et al. [11] extended the method to discrete multi-objective feature recognition and matching. The Bayer and Kalman filtering algorithms are used to improve the pose estimation accuracy of moving targets. Engel et al. [12] extended binocular vision to simultaneous localization and mapping (SLAM) and implemented a fast initialization of the direct method. In addition, the calibration accuracy determines the measurement accuracy for a binocular vision sensor. J. Apolinar Muñoz Rodríguez [13] proposed a technique to perform microscope self-calibration via micro laser line and soft computing algorithms. In this technique, the microscope vision parameters are computed by means of soft computing algorithms based on laser line projection.

In engineering practice, due to the limitations of a single sensor, target measurement technology based on multisensor information fusion with high-precision and high-robustness has become the main method to make up for the shortcoming of a single sensor [14]. Multisensor information fusion (MSIF), which appeared in the late 1970s, is widely used in target tracking and image and signal processing [15–17]. MSIF is a method for improving the adaptability and measurement accuracy of the system by integrating the correlation of each sensor in the same scene.

Based on the fusion of vision and inertial measurement unit (IMU), Leutenegger. S et al. [18] integrated the IMU error into the reprojection error cost function to construct a nonlinear error function, and maintained a fixed optimization window through marginalization to ensure the real-time performance of the process. Forster C et al. [19] proposed a preintegration theory that aggregates the inertial measurements of adjacent keyframes into independent motion constraints and fuses them into a factor graph framework. For the fusion of vision and Lidar, Tomic T et al. [20] used a multihead camera and point cloud registration to jointly estimate the motion state of the drone, and they used the map pyramid to improve the efficiency of the algorithm to achieve real-time performance. Zhang J et al. [21] proposed the depth-enhanced monocular odometry (DEMO) method based on positioning and reconstruction of the camera and 3D lidar. Vision sensors can also be used with sonic sensors for automatic parking, and with computed tomography (CT) images for medical diagnosis, with infrared sensors for target tracking, among other applications. For the fusion of vision and laser. J. Apolinar Muñoz Rodríguez and Francisco Carlos Mejía Alanís [22] proposed an accurate technique to perform binocular self-calibration by means of an adaptive genetic algorithm based on a laser line. In this calibration, the genetic algorithm computes the vision parameters through simulated binary crossover (SBX).

In this paper, we propose a method for correcting the distance of binocular vision along the optical axis with a one-dimensional point laser ranging sensor that is installed on a two-dimensional turntable firstly. Limited by the mechanical structure of the two-dimensional turntable in the actual measurement process, the measurement frequency of the laser ranging sensor is lower than that of the binocular vision sensor, resulting in time-delay and redundancy of the measurement information of the entire system. Second, we propose a direct update algorithm based on one-step prediction, which transforms the time-delay information into real-time information, and combines the federal Kalman filtering algorithm to complete the positioning and tracking of spatial targets. Finally, the experimental results show that the system structure and algorithm that we proposed can effectively improve the accuracy and real-time performance of the multisensor system. It should be noted that the parameters of the binocular vision sensor used in this paper are obtained after image processing. Image processing such as camera calibration, distortion correction, etc. is not the research content of this article.

2. System Construction

In this paper, the experimental platform of the target tracking system is mainly composed of two parts: the measurement system and the target system. The measurement system is a typical heterogeneous sensor system that is composed of a binocular vision sensor system and a laser-ranging sensor system. The laser-ranging sensor is mounted on a two-dimensional (2D) turntable. The target

system consists of a model that moves via a two-dimensional moving slide-table. The space schematic of the target tracking system is shown in Figure 1; the physical experimental system is shown in Figure 2.

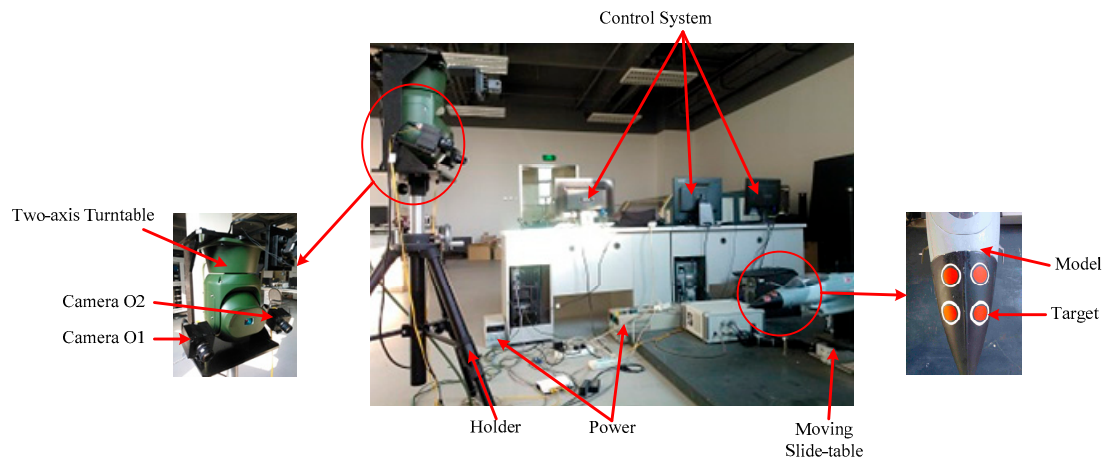


Figure 1. The physical experimental system. The target to be tracked is a red round sticker attached to the model.

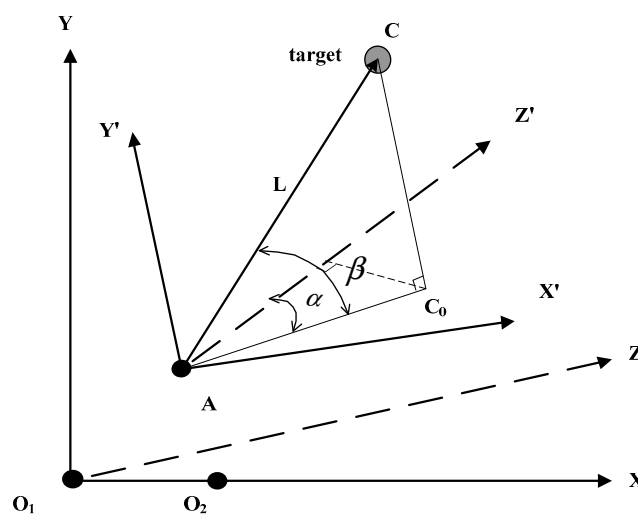


Figure 2. The space geometry relation of a multisensor system.

As shown in Figure 2, O_1 and O_2 are two cameras of the binocular vision sensor. A is the point of the laser ranging sensor that is fixed in the center of the two-axis turntable. C is the point of the target. According to the spatial location relation of each sensor and target, the coordinate systems O_1XYZ and $AX'Y'Z'$ are established.

The coordinate system O_1XYZ : camera O_1 is the ordinate origin. The extension to the other camera O_2 is in the X-axis direction and the Y-axis is vertical to O_1O_2 . The right-hand rule is used to determine that the direction along the optical axis is the Z-axis. The positive direction of the coordinate axes is shown in Figure 2. The coordinates of target C are (x_c, y_c, z_c) in coordinate system O_1XYZ .

The coordinate system $AX'Y'Z'$: the point A is the ordinate origin. The directions of the X-axis, Y-axis and Z-axis are the same as in the coordinate system O_1XYZ . The coordinates of target C and camera O_1 are (x'_c, y'_c, z'_c) and (x_1, y_1, z_1) in the coordinate system $AX'Y'Z'$, respectively.

The projection of point C on surface $X'AY'$ is C_0 . The angle between AC_0 and the Y' axis is α . The angle between AC and surface $X'AZ'$ is β .

The steps for obtaining the spatial location of the target are as follows:

(1) Coordinate system O_1XYZ is set as the reference coordinate system. The binocular vision sensor system first acquires the spatial location of the target, and then transmits the measurement data to the central control system through the visual control computer.

(2) The target space information acquired by the binocular vision sensor is the coordinates in the coordinate system O_1XYZ . Coordinate transformation to coordinate system $AX'Y'Z'$ is required to obtain the pitch and yaw angles by which the turntable must be rotated.

The coordinates (x'_c, y'_c, z'_c) are given as follows:

$$\begin{aligned} [x'_c, y'_c, z'_c] &= [x_c, y_c, z_c] \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos m & -\sin m \\ 0 & \sin m & \cos m \end{bmatrix} \times \begin{bmatrix} \cos n & 0 & \sin n \\ 0 & 1 & 0 \\ -\sin n & 0 & \cos n \end{bmatrix} \times \begin{bmatrix} \cos k & -\sin k & 0 \\ \sin k & \cos k & 0 \\ 0 & 0 & 1 \end{bmatrix} + [x_1, y_1, z_1] \\ &= [x_c, y_c, z_c] \times \text{Rot}(x, m) \times \text{Rot}(y, n) \times \text{Rot}(z, k) + [x_1, y_1, z_1] \end{aligned} \quad (1)$$

where $\text{Rot}(x, m)$ is the rotation matrix around the X' axis. m is the angle of rotation which is a constant value and is obtained from the initial calibration. $\text{Rot}(y, n)$ and $\text{Rot}(z, k)$ are the same.

It is obtained from the spatial geometry of Figure 2:

$$\alpha = \arctan\left(\frac{x'_c}{z'_c}\right), \quad (2)$$

$$\beta = \arctan\left(\frac{|y'_c|}{\sqrt{x'^2_c + z'^2_c}}\right), \quad (3)$$

(3) The yaw angle α and pitch angle β of the 2D turntable are transferred to the laser control computer. Then, after the 2D turntable is adjusted to the designated position, the laser ranging sensor is controlled to shoot towards target C through the laser control computer and the distance value $AC = l$ is returned to the central control system.

(4) The measurement values l of the laser ranging sensor are used to correct the measured value of the binocular vision sensor along the optical axis direction (Z axis). Because of the high accuracy of the binocular vision sensor along the vertical optical axis, the coordinates on the X axis and Y axis can be used as the measured values.

The Z coordinate of point C in the coordinate system $AX'Y'Z'$ is obtained as follows:

$$z'_c = l \times \cos \alpha \times \cos \beta, \quad (4)$$

By the coordinate transformation, the Z' coordinate of point C is obtained:

$$z'_c = z_c \times \text{Rot}(x, m) \times \text{Rot}(y, n) \times \text{Rot}(z, k) + z_1, \quad (5)$$

Therefore, the new Z coordinate value of the C point z_{nc} in the coordinate system O_1XYZ after the correction by the laser ranging sensor can be solved:

$$z_{nc} = (l \times \cos \alpha \times \cos \beta) / [\text{Rot}(x, m) \times \text{Rot}(y, n) \times \text{Rot}(z, k)], \quad (6)$$

Through error analysis and the error transfer formula, the error of z_{nc} can be obtained:

$$\Delta z_{nc} = \frac{\cos \alpha \cos \beta |\Delta l| + l \sin \alpha \cos \beta |\Delta \alpha| + l \cos \alpha \sin \beta |\Delta \beta|}{\text{Rot}(x, m) \times \text{Rot}(y, n) \times \text{Rot}(z, k)}, \quad (7)$$

where, Δl is the measurement error of laser-ranging sensor, $\Delta \alpha$ is the error of the yaw angle, $\Delta \beta$ is the error of the pitch angle.

In practice, because the position calibration error between the binocular vision sensor and the laser-ranging sensor is quite small, the primary error is concentrated in the above three errors. The

measurement error of the laser-ranging sensor is determined by its own performance. The angle error includes the calculation error caused by binocular vision sensor and the rotation error of the 2D turntable. Therefore, the spatial position error of the target corrected by the laser-ranging sensor is affected by the measurement error of the binocular vision sensor.

3. Problem Description

According to the calculation process of the coordinate z_c , the information flow of the target tracking is shown in Figure 3.

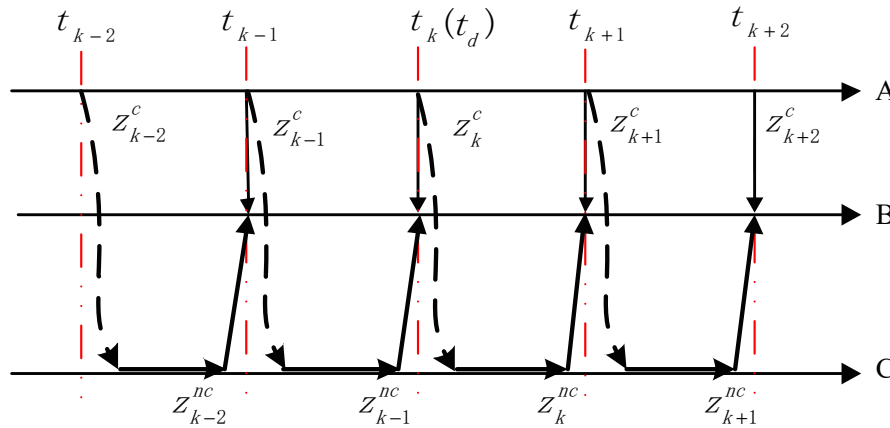


Figure 3. The measurement message passing of system: (A): the measurement time of binocular vision; (B): the measurement time of fusion; (C): the measurement of laser ranging sensor.

The measured value z_{k-1}^c is obtained by the binocular vision sensor at time t_{k-1} . Then, the yaw angle α and pitch angle β are calculated by coordinate transformation and transferred to the 2D turntable controller. After the turntable is rotated to the corresponding angle, the laser-ranging sensor is tested and the corrected coordinate value z_{k-1}^{nc} at time t_k is calculated. Since the frequency of mechanical rotation of the 2D turntable is much lower than the measurement frequency of the binocular vision sensor, the measured value z_k^c is obtained by the binocular vision sensor at time t_k .

From the process of constructing the system and the acquisition of the target space position, the measurement information of the binocular vision sensor and that of the laser-ranging sensor are related, and the laser-ranging sensor system has constant time delay.

Consider the following multiple sensors system with observing time-delay:

$$x_k = Fx_{k-1} + Gw_{k,k-1}, \quad (8)$$

$$z_k^i = Hx_{k-d_i} + v_k^i, \quad (9)$$

where, $1 \leq d_1 \leq d_2 \leq \dots \leq d_N \leq d_i$ ($i = 1, 2, \dots, N$), in which d_i is the observation delay of the i th local sensor; $x_k \in R^{n \times 1}$ is the state vector of system at time t_k , $F \in R^{n \times n}$ is the $n \times n$ dimensional state-transitional matrix. $G \in R^{n \times h}$ is the $n \times h$ dimensional noise input matrix. $z_k^i \in R^{m \times 1}$, $i = 1 \dots L$ is the m -dimensional measured vector of the i th sensor, and L is the number of sensors. $H^i \in R^{m \times n}$, $i = 1 \dots L$ is the measured matrix of the i th sensor. $w_{k,k-1} \in R^{h \times 1}$ is h -dimensional process noise vector of i th sensor. $v_k^i \in R^{m \times 1}$, $i = 1 \dots L$ is the observation noise of the i th sensor.

In the real-time sensor system, at time $t = t_k$, the following can be obtained:

$$\begin{cases} \hat{x}_{k|k} = E^*[x_k|Z^k] \\ P_{k|k} = \text{cov}[x_k|Z^k] \end{cases}, \quad (10)$$

where Z^k is the measurement set $Z^k = \{z_i^N\}_{i=1}^k$ of the N th sensor at time t_k .

Assume that the lag time of the time-delay sensor (laser ranging sensor) is t_{k-d} . There are real-time measurement z_k^i (binocular vision sensor) and time-delay measurement z_{k-d}^j (laser ranging sensor) in the fusion center at time t_k . We need to use earlier measurements of the time-delay sensor to update the estimation $\hat{x}_{k|k}$:

$$\begin{cases} \hat{x}_{k|k,d} = E^*[x_k|Z^k, z_d] \\ P_{k|k,d} = \text{cov}[x_k|Z^k, z_d] \end{cases} \quad (11)$$

Moreover, the estimated value $\hat{x}_{k|k-d}$ of the time-delay sensor system at time t_k is considered to be the real-time measured value z_k^j at time t_k . The measured value of all sensors at time t_k are $\{z_k^1, z_k^2, \dots, z_k^i, \dots, z_k^j\} (1 \leq i \leq j \leq N)$, where the estimates of z_k^i and z_k^j are relevant. To obtain more accurate space coordinates of the target, we need to resolve the above correlation problem and the optimal fusion estimation of the target motion state x_k in the fusion center.

4. Time-Delay Information Update and Fusion

4.1. Processing of Measurement Constant Time Lag

Based on the stochastic linear time-invariant discrete system and the iterated state equation, we can obtain the following:

$$x_{k+1} = F^{d_i} x_{k+1-d_i} + \sum_{j=1}^{d_i} F^{j-1} G w_{k+1-j}, \quad (12)$$

Namely,

$$x_{k-d_i} = F^{-d_i} x_k - \sum_{j=1}^{d_i} F^{j-1-d_i} G w_{k-j}, \quad (13)$$

Inserting the above formula into the observation equation,

$$z_k^i = H^i F^{-d_i} x_k - \sum_{j=1}^{d_i} H^i F^{j-1-d_i} G w_{k-j} + v_k^i, \quad (14)$$

Assuming $\bar{H}^i = H^i F^{-d_i}$, $\eta_k^i = v_k^i - \sum_{j=1}^{d_i} H^i F^{j-1-d_i} G w_{k-j} = v_k^i - \sum_{j=1}^{d_i} \bar{H}^i F^{j-1} G w_{k-j}$,

Then the system observation equation is transformed into:

$$z_k^i = \bar{H}^i x_k + \eta_k^i, \quad (15)$$

Combined with Equation (15), the time-delay subsensor system has the following optimal Kalman filter and one-step predictor,

$$\hat{x}_{k+1|k+1}^i = \hat{x}_{k+1|k}^i + K_{k+1}^i \varepsilon_{k+1}^i, \quad (16)$$

$$\hat{x}_{k+1|k}^i = F \hat{x}_{k|k}^i, \quad (17)$$

$$\varepsilon_{k+1}^i = z_{k+1}^i - \hat{z}_{k+1|k}^i = z_{k+1}^i - \bar{H}^i \hat{x}_{k+1|k}^i, \quad (18)$$

$$K_{k+1}^i = (P_{k+1|k}^i (\bar{H}^i)^T - \sum_{j=1}^{d_i} F^{j-1} G Q G^T (F^{j-1})^T (\bar{H}^i)^T) P_{\varepsilon}^i, \quad (19)$$

$$P_{\varepsilon}^i = \bar{H}^i P_{k+1|k}^i (\bar{H}^i)^T + R^{ii} - \sum_{j=1}^{d_i} \bar{H}^i F^{j-1} G Q G^T (F^{j-1})^T (\bar{H}^i)^T, \quad (20)$$

$$P_{k+1|k}^i = F P_{k|k}^i F^T + G Q G^T, \quad (21)$$

$$P_{k+1|k+1}^i = P_{k+1|k}^i - K_{k+1}^i P_{\varepsilon}^i (K_{k+1}^i)^T, \quad (22)$$

where, ε_{k+1}^i is the innovation. $P_{\varepsilon}^i = E[\varepsilon_{k+1}^i (\varepsilon_{k+1}^i)^T]$ is the innovation variance. K_{k+1}^i is the filter gain. $P_{k+1|k+1}^i$ is the filtering error variance matrix. $P_{k+1|k}^i$ is the one step prediction error variance matrix.

4.2. Information Fusion

To solve the estimation correlation between two sensor systems and further improve the accuracy of measurement, the federal Kalman filtering algorithm is used for subsequent processing, which includes information distribution, time updating, measurement updating, and estimation fusion.

(1) Information distribution

The main filter only updates the timing and dose not measure. The process information of the system is shared among the subfilters and the main filters according to the principle of information distribution.

$$\begin{cases} P_{k-1}^i = \beta_i^{-1} P_{k-1} \\ Q_{k-1}^i = \beta_i^{-1} Q_{k-1} \\ x_{k-1}^i = X_{k-1} \end{cases}. \quad (23)$$

According to the Law of Information Conservation, $\sum_{i=0}^n \beta_i = 1$.

(2) Time updating

The covariance of the system state and estimation error is transferred according to the system transfer matrix, which is performed independently for the sub-filter and the main filter.

$$\hat{x}_{k|k-1}^i = F_{k|k-1}^i \hat{x}_{k-1}^i, \quad (24)$$

$$P_{k|k-1}^i = F_{k|k-1}^i P_{k-1}^i (F_{k|k-1}^i)^T + G_{k-1}^i Q_{k-1}^i (G_{k-1}^i)^T. \quad (25)$$

(3) Measurement updating

The system state and estimated error covariance are updated using the new measurement information. Since the main filter performs no measurements, the measurement updating is only performed in the subfilter.

$$K_k^i = P_{k|k-1}^i (H_k^i)^T (R_k^i)^{-1}, \quad (26)$$

$$\hat{x}_k^i = \hat{x}_{k|k-1}^i + K_k^i (z_k^i - H_k^i \hat{x}_{k|k-1}^i), \quad (27)$$

$$(P_k^i)^{-1} = (P_{k|k-1}^i)^{-1} + (H_k^i)^T (R_k^i)^{-1} H_k^i, \quad (28)$$

(4) Estimation fusion

$$\hat{X}_g = P_g \sum (P_k^i)^{-1} \hat{x}_k^i, \quad (29)$$

$$P_g = \left(\sum (P_k^i)^{-1} \right)^{-1}. \quad (30)$$

In above steps, information distribution is a key part of federated filtering, which is an important feature that distinguishes it from other decentralized filtering methods. The coefficient of the information distribution determines the accuracy of the final fusion result.

According to the estimated error covariance,

$$P_i = E[(X_i - \hat{X}_g)(X_i - \hat{X}_g)^T]. \quad (31)$$

It can be seen that P describes the estimation accuracy of X , and the smaller the P , the higher the estimation accuracy of X .

Considering the use of a globally optimal solution to reset the filter values and error variance matrices in the next filtering step, the influence of the information distribution coefficients on global estimates is discussed:

By inserting Equation (30) into Equation (23), the following is obtained:

$$P_{k+1}^i = \beta_i^{-1} P_g. \quad (32)$$

Equations (30) and (32) are substituted into Equation (25) to obtain:

$$\hat{X}_{k+1|k}^i = F_{k+1|k}^i \hat{X}_{g,k} = F_{k+1|k}^i P_g \sum (\beta_i^{-1} P_k^i)^{-1} \hat{x}_k^i \quad (33)$$

where $\hat{x}_k^i = F_{k|k-1}^i \hat{x}_{k-1}^i + K_k^i (z_k^i - H_k^i \hat{x}_{k|k-1}^i)$.

P and Q have the same dimension in general. It is given by the following:

$$P_i = E[(X_i - \hat{X}_g)(X_i - \hat{X}_g)^T] \quad (34)$$

$$\begin{aligned} P_{k+1|k}^i &= F_{k+1|k}^i P_k^i (F_{k+1|k}^i)^T + G_k^i Q_k^i (G_k^i)^T = F_{k+1|k}^i \beta_i^{-1} P_{g,k} (F_{k+1|k}^i)^T + G_k^i \beta_i^{-1} Q_{g,k} (G_k^i)^T \\ &= \beta_i^{-1} (F_{k+1|k}^i P_{g,k} (F_{k+1|k}^i)^T + G_k^i Q_{g,k} (G_k^i)^T) = \beta_i^{-1} P_{k+1|k}^i \end{aligned} \quad (35)$$

where $P_{k+1|k}^i \triangleq F_{k+1|k}^i P_{g,k} (F_{k+1|k}^i)^T + G_k^i Q_{g,k} (G_k^i)^T$.

Taking the inverse of the Equation (34) on both sides, it is the one-step predictive state information matrix of local filter and global filter.

$$(P_{k+1|k}^i)^{-1} = \beta_i (P_{k+1|k}^i)^{-1} \quad (36)$$

Taking the trace of both sides, we can obtain:

$$\beta_i = \frac{\text{tr}[(P_{k+1|k}^i)^{-1}]}{\text{tr}[(P_{k+1|k}^i)^{-1}]} \quad (37)$$

β_i is inversely proportional to the estimated error covariance. When the estimation covariance is larger, the estimation quality is poorer, the subfilter accuracy is lower, and the information distribution coefficient is smaller.

5. System Experiment and Analysis

The target-tracking system of binocular vision laser-ranging sensor is shown in Figure 1. The target is fixed on a sliding platform and performs linear motion in space. At present, only the movement of the target in the direction of the optical axis (Z axis) is studied, and the measured values in the experimental results only represent the measured values in the Z axis direction.

Through error analysis of the laser-ranging sensor, the parameters of the system are brought into the error transfer function. The measurement error of the laser-ranging sensor is $\pm 1.5\text{mm}$, the rotation error of the two-dimensional turntable is $\pm 0.02^\circ$, the measurement error of the binocular vision X and Y directions is $\pm 3\text{mm}$, and the Z-axis measurement error is $\pm 40\text{mm}$. After several calculations, the final average measurement error of the laser-ranging sensor is $\pm 22.3\text{mm}$, which proves that the laser-ranging sensor improves the measurement accuracy of binocular vision along the optical axis.

As shown in Figure 4, the measurement value of the laser-ranging sensor is directly predicted, and the estimated value of the lag information is used as real-time information for subsequent calculations. It can be seen from the simulation results at the time of $t(25)$ to $t(40)$ that the lag information is improved after the direct update algorithm, and the estimated value has errors due to the influence of noise, but the change trend is basically consistent with the original data. At the same time, when

observing the whole curve, the target position changes slowly at the beginning stage, and the slope of the curve gradually increases with the passage of time, and then basically remains unchanged from $t = 38$. This is because there is acceleration at the beginning of the sliding platform where the target is located. After reaching the specified speed, the target enters the stage of uniform motion, and the slope does not change. The target moves for a long time, and the curve of the subsequent deceleration phase is not drawn in Figure 4.

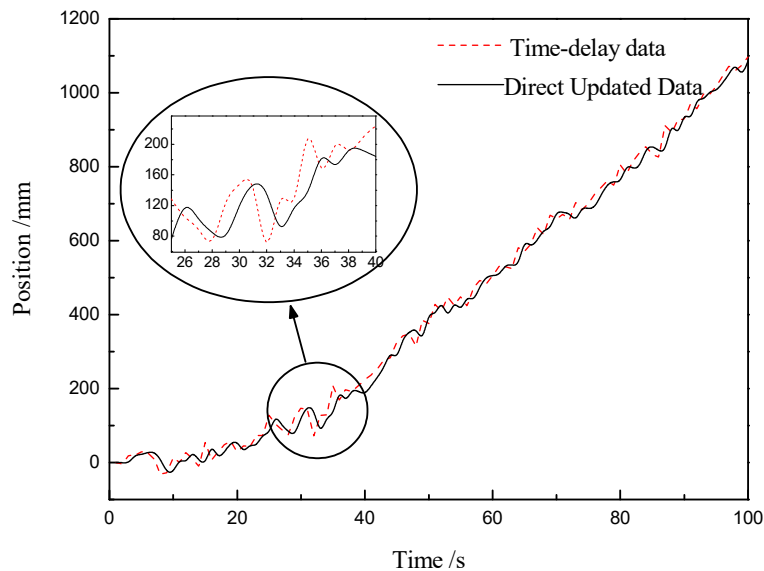


Figure 4. Direct update of time-delay data.

Then, the information fusion algorithm based on the federated Kalman filter is verified by experiments. A comparison between the measurement results of a single sensor and the fusion results is shown in Figure 5.

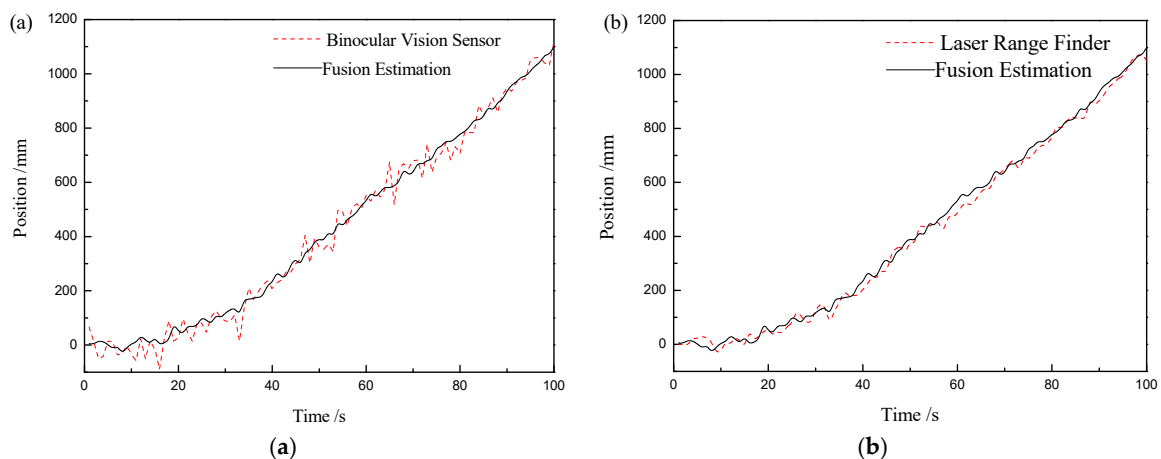


Figure 5. The tracking results for two sensors and fusion. (a) Tracking via the binocular vision sensor; (b) tracking via the laser range sensor.

It can be clearly observed in Figure 5 that compared with the binocular vision sensor with larger error, the curve of the fusion result is smoother, and the accuracy is improved.

Compared with the measured value of the laser-ranging sensor, the fusion result can not be seen directly, so the mean squared error (MSE) between each measurement result and the calibrated true value of the target is calculated. The MSE results in Figure 6 indicate that the accuracy of the target

position after fusion is improved compared to a single sensor, and the error of the result after fusion is the smallest.

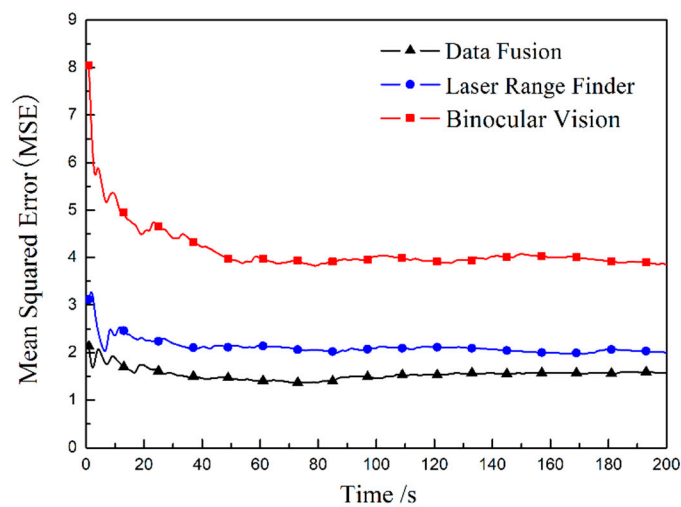


Figure 6. The mean squared error (MSE) of sensors and fusion results.

Figure 7 shows the change curve of the binocular vision sensor's information distribution coefficient. At the beginning of the measurement, the information distribution coefficient changes rapidly from 0.5 in 5 S. After 15 s, it tends to be stable and its value is 0.37. Because the measurement error of the binocular vision sensor along the optical axis is large, its information distribution coefficient is small, while the laser-ranging sensor has a large information distribution coefficient, which is consistent with the conclusions obtained in the paper.

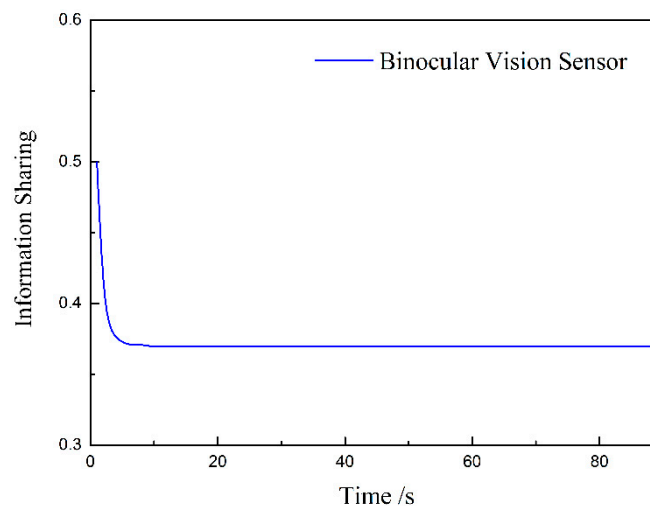


Figure 7. The binocular vision information distribution.

6. Conclusions

Through theoretical calculations and experimental verifications, the accuracy of binocular vision along the optical axis is improved in this paper. First, regarding the system structure, we propose to use a one-dimensional point laser-ranging sensor to correct the measurement value of the binocular vision sensor along the optical axis. Second, regarding the measurement process, we found that it is limited by the performance of the two-dimensional turntable, the system has time delay. To improve the utilization and real-time performance of the information of the multisensor measurement system, an optimal information fusion algorithm for the multisensor target tracking system characterized by estimation correlation and constant time delay is studied. We propose a method to separate the

complex multisensor environment, which first uses the one-time prediction of the constant delay information as the real-time information at the current moment to solve the delay problem, and then uses the federal Kalman filter to address the estimation correlation. Finally, the experimental results verify the validity and accuracy of this method.

Although the research in this article provides a basis for the study of the experimental system in actual environments, the current experimental environment is relatively simple, and the experimental parameters and errors are under control. When the distance between the target and sensor system increases, the visual error will increase, and the laser spot will become larger, which will cause the errors of the laser-ranging sensor to increase. There are still many improvements to be studied.

Author Contributions: Concept design, Q.W., Y.Z. and M.N.; Experiment, Q.W. and W.S.; Data analysis, Q.W.; Manuscript writing, Q.W., W.S. and M.N.; Software, Y.Z. and M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Basic Research Program of Shaanxi (Program No. 2019JM-468).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Q.; Chu, B.; Peng, J. A Visual Measurement of Water Content of Crude Oil Based on Image Grayscale Accumulated Value Difference. *Sensors* **2019**, *19*, 2963. [[CrossRef](#)] [[PubMed](#)]
2. Sejong, H.; Jaehyuck, C.; Gook, P.C. EKF-Based Visual Inertial Navigation Using Sliding Window Nonlinear Optimization. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 2470–2479.
3. Zhang, J.; Yang, Z.; Deng, H. Dynamic Visual Measurement of Driver Eye Movements. *Sensors* **2019**, *19*, 2217. [[CrossRef](#)] [[PubMed](#)]
4. Liu, G.; Xu, C.; Zhu, Y. Monocular Vision-based Pose Determination in Close Proximity for Low Impact docking. *Sensors* **2019**, *19*, 3261. [[CrossRef](#)] [[PubMed](#)]
5. Yu, X.; Fan, Z.; Wan, H. Positioning, Navigation, and Book Accessing/Returning in an Autonomous Library Robot using Integrated Binocular Vision and QR Code Identification Systems. *Sensors* **2019**, *19*, 783. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, T.; Liu, J.; Liu, S. A 3D Reconstruction Method for Pipeline Inspection Based on Multi-vision. *Measurement* **2017**, *98*, 35–48. [[CrossRef](#)]
7. Hilsenbeck, S.; Möller, A.; Huitl, R.; Schroth, G.; Kranz, M.; Steinbach, E. Scale-preserving long-term visual odometry for indoor navigation. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sydney, Australia, 13–15 November 2012; pp. 1–10.
8. Jiang, G.; Gong, H.; Jiang, T. Close-form solution of absolute orientation based on inverse problem of orthogonal matrices. In Proceedings of the 2008 Congress on Image and Signal Processing, Sanya, China, 27–30 May 2008; pp. 329–333.
9. Terui, F.; Kamimura, H.; Nishida, S. Motion Estimation to A Failed Satellite on Orbit Using Stereo Vision and 3D Model Matching. In Proceedings of the 2006 9th International Conference on Control, Automation, Robotics and Vision, Singapore, 5–8 December 2006; pp. 1–8.
10. Du, X.; Liang, B.; Xu, W. Pose Measurement of Large Non-cooperative Satellite Based on Collaborative Cameras. *Acta Astronautica* **2011**, *68*, 2047–2065. [[CrossRef](#)]
11. Segal, S.; Carmi, A.; Gurfil, P. Vision-based Relative State Estimation of Non-cooperative Spacecraft under Modeling Uncertainty. In Proceedings of the 2011 Aerospace Conference, Washington, DC, USA, 5–12 March 2011; pp. 1–8.
12. Engel, J.; Stücker, J.; Cremers, D. Large-Scale Direct SLAM with Stereo Cameras. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015; pp. 1935–1942.
13. Rodríguez, J.A.M.; Mejía Alanís, F.C. Binocular self-calibration performed via adaptive genetic algorithm based on laser line imaging. *J. Mod. Opt.* **2016**, *63*, 1219–1232. [[CrossRef](#)]
14. Qian, C.; Liu, H.; Tang, J. An Integrated GNSS/INS/LiDAR-SLAM Positioning Method for Highly Accurate Forest Stem Mapping. *Remote Sens.* **2017**, *9*, 3. [[CrossRef](#)]

15. Aguilera, A.A.; Brena, R.F.; Mayora, O. Multisensor Fusion for Activity Recognition—A Survey. *Sensors* **2019**, *19*, 3808. [[CrossRef](#)] [[PubMed](#)]
16. Ji, S.Q.; Huang, M.B.; Huang, H.P. Robot Intelligent Grasp of Unknown Objects Based on Multisensor Information. *Sensors* **2019**, *19*, 1595. [[CrossRef](#)] [[PubMed](#)]
17. Li, X.; Chen, W.; Chan, C. Multisensor Fusion Methodology for Enhanced Land Vehicle Positioning. *Inf. Fusion* **2019**, *46*, 51–62. [[CrossRef](#)]
18. Leutenegger, S.; Lynen, S.; Bosse, M. Keyframe-based Visual-inertial Odometry Using Nonlinear Optimization. *Int. J. Rob. Res.* **2015**, *34*, 314–334. [[CrossRef](#)]
19. Forster, C.; Carlone, L.; Dellaert, F. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry. *IEEE Trans. Rob.* **2016**, *33*, 1–21. [[CrossRef](#)]
20. Tomic, T.; Schmid, K.; Lutz, P. Toward A Fully Autonomous UAV: Research Platform for Indoor and Outdoor Urban Search and Rescue. *IEEE Rob. Autom. Mag.* **2012**, *19*, 46–56. [[CrossRef](#)]
21. Zhang, J.; Kaess, M.; Singh, S. Real-time Depth Enhanced Monocular Odometry. In Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, USA, 14–18 September 2014; pp. 4973–4980.
22. Rodríguez, J.A.M. Microscope self-calibration based on micro laser line imaging and soft computing algorithms. *Opt. Lasers Eng.* **2018**, *105*, 75–85.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).