

Article

Real-Time Image Stabilization Method Based on Optical Flow and Binary Point Feature Matching

Zilong Deng ¹, Dongxiao Yang ^{1,*}, Xiaohu Zhang ², Yuguang Dong ³,
Chengbo Liu ¹ and Qiang Shen ¹

¹ School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China; 3120195173@bit.edu.cn (Z.D.); marsike663@126.com (C.L.); bit82shen@bit.edu.cn (Q.S.)

² China North Industries Group Finance Company LTD, Xi'an 710000, China; zhangxiaohu1111@163.com

³ China North Industries Corporation, Beijing 100053, China; ygdong@norinco.cn

* Correspondence: yangdx@bit.edu.cn

Received: 20 December 2019; Accepted: 14 January 2020; Published: 20 January 2020

Abstract: The strap-down missile-borne image guidance system can be easily affected by the unwanted jitters of the motion of the camera, and the subsequent recognition and tracking functions are also influenced, thus severely affecting the navigation accuracy of the image guidance system. So, a real-time image stabilization technology is needed to help improve the image quality of the image guidance system. To satisfy the real-time and accuracy requirements of image stabilization in the strap-down missile-borne image guidance system, an image stabilization method based on optical flow and image matching with binary feature descriptors is proposed. The global motion of consecutive frames is estimated by the pyramid Lucas-Kanade (LK) optical flow algorithm, and the interval frames image matching based on fast retina keypoint (FREAK) algorithm is used to reduce the cumulative trajectory error. A Kalman filter is designed to smooth the trajectory, which is conducive to fitting to the main motion of the guidance system. Simulations have been carried out, and the results show that the proposed algorithm improves the accuracy and real-time performance simultaneously compared to the state-of-art algorithms.

Keywords: strap-down image guidance system; image stabilization; optical flow; Kalman filter

1. Introduction

The strap-down missile-borne image guidance system is a sort of real-time guidance system based on computer vision, which has been applied in practice, such as the image-guided miniature ammunition, Spike [1]. The Spike munition has the advantages of being a light weight, low cost, and no gimbaled system, which makes it simpler to design the strap-down seekers' mechanical configurations [2]. Spike is one of the smallest and the cheapest guided missiles with outstanding maneuverability performance, which can reach a speed of 600 miles per h in 1.5 s after launch, and the onboard guidance system can manipulate the control surfaces to complete the movements of missile on pitch, roll and yaw axes in a prelimited time.

The general working process of strap-down missile-borne image guidance system consists of three steps. First, after launching of the missile, the onboard camera turns on to search, capture, and track the target during the flight. Then, when the target is locked, the onboard computer calculates the angular deviation between the center of the target and the visual central axis. At last, the deviation is sent to the autopilot to manipulate the control surfaces of the vehicle and guide the missile to the target automatically till hitting the target.

During the process, the quality of image sequences is an important factor affecting the navigation accuracy in the image guidance system. Since the strap-down seekers do not require any

moveable gimbal parts to isolate the motion disturbance of the missile and the seeker [3], there are jitters produced by windy condition and attitude adjustment, and the image sequences obtained by the image guidance system may have a large degree of shaking, which will seriously affect the performance of the image guidance system. Therefore, before recognition, matching, and other subsequent processing, image preprocessing is needed to eliminate the influence of jitters. So, it is necessary to stabilize the digital video acquired by the missile-borne guidance system in real time. Through the image stabilization technology, the impact of the moving camera on the image can be eliminated or reduced, which greatly improves the quality of image sequences. In addition, with the development of large area scientific imaging arrays, as well as the high speed processing elements such as DSP (digital signal processing), FPGA (field programmable gate array), GPU (graphics processing unit) and CUDA (compute unified device architecture), real-time image processing technology for high resolution images has become a hot topic in recent years [4,5], while real-time electronic image stabilization technology as one of the significant image processing techniques has also been extensively studied.

In general, video stabilization algorithms are mainly carried out with three steps: global motion estimation, motion compensation, and image compensation. Current motion estimation algorithms include the gray projection algorithm [6,7], block-matching algorithm [8,9], bit-plane-matching algorithm [10,11], optical flow algorithm [12], and representative point matching (RPM) algorithm [13]. These algorithms are easy to implement with low computational cost. However, the accuracy of these algorithms can be insufficient under the circumstances of severe dithering, complicated movements, or image distortions of the video. Further, the cumulative error of motion estimation makes the estimated image trajectory deviate from the true trajectory susceptibly. Except the algorithms above, the binary point feature matching-based video stabilization algorithms could deal with these situations efficiently, but with higher computational cost. For instance, the scale-invariant feature transform (SIFT) algorithm can be used to estimate global motion precisely, even when the scale, rotation, illumination, and angle of view of the images change [14–16]. The speeded up robust features (SURF) algorithm is widely used for video stabilization for it has effective image matching performance with less computational cost [17–19] compared with SIFT. Other point feature matching algorithms such as the binary robust independent elementary features (BRIEF) [20], binary robust invariant scalable keypoints (BRISK) [21], and fast retina keypoint (FREAK) [22] algorithms are also proposed for faster computational speed while trying to retain the robustness to noise, scale invariance, and rotation invariance properties as much as possible.

These algorithms are hard to achieve video stabilization independently due to the limitation of their computational cost or accuracy. Therefore, some research teams have tried to achieve real-time image stabilization by combining or improving the existing image processing algorithms. For example, Dong et al. [23] proposed a motion model based on inter-frame homographic estimation with Kalman filter for the video completion. Lim et al. [24] proposed an algorithm to tackle the problem of real-time video stabilization for unmanned aerial vehicles (UAVs), where they designed a suitable model for the global motion of UAV and employed the optical flow tracking. Hu et al. [25] proposed a real-time video stabilization system for the video sequences captured by a fast-moving in-vehicle camera. The proposed method used feature points to evaluate the global motion and the feature points are checked based on LK (Lucas-Kanade) optical flow method.

However, due to the particular environment where the strap-down missile-borne image guidance system is mounted, the strap-down camera keeps moving when the projectile tracks the target, and the projectile would generate jitters when adjusting its attitude. So, an image stabilization algorithm with low computational cost and steady stabilization performance, which can also distinguish and preserve the low-frequency movements and remove the high-frequency jitters of the video, needs to be found.

Therefore, a real-time image stabilization algorithm for missile-borne strap-down image guidance system is proposed to distinguish the low-frequency movements, and remove the high-frequency jitters of the video. In this paper, an image stabilization algorithm using feature point matching to correct optical flow error with the random sample consensus (RANSAC) algorithm

[26,27] to eliminate false matching is proposed to meet the requirements of image processing accuracy and real-time performance. Moreover, a Kalman filter is designed to smooth the motion trajectory to make the video fit the intentional motion as much as possible.

The structure of this paper is as follows. In Section 2, the overall framework of the proposed algorithm is analyzed and illustrated. Then, in Section 3, the global motion of missile-borne image guidance system is established and estimated by the pyramid LK optical flow algorithm. In Section 4, the computational costs of different feature point matching algorithms are analyzed, and the FREAK algorithm is selected to conduct correction of the trajectory generated by the optical flow algorithm. The equations of Kalman filter are also established based on the global motion of the missile-borne image guidance system. The specific framework of the proposed algorithm is given in the last paragraph. Experiments have been carried out to verify the feasibility and effectiveness of the proposed algorithm in Section 5. Conclusions are presented in Section 6.

2. Proposed Framework

The block diagram of the overall proposed framework is shown in Figure 1.

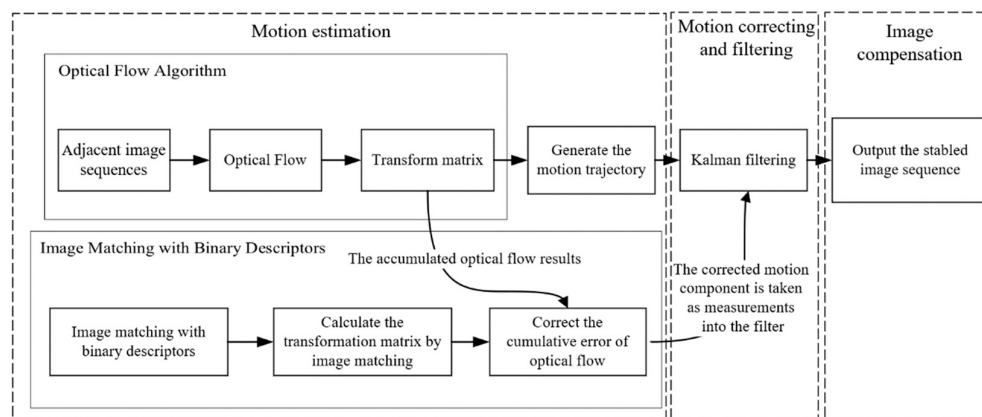


Figure 1. Block diagram of the proposed framework.

As shown in Figure 1, the global motion trajectory between the adjacent frames of the image sequence is obtained by the optical flow algorithm. Then, the image matching algorithm is conducted periodically to correct the cumulative error of the optical flow algorithm. The trajectory generated by optical flow is filtered by a Kalman filter in every frame, and the measurements that corrected by the point feature matching algorithm are taken as input of the Kalman filter to correct the cumulative error of optical flow algorithm as well.

The motion trajectory of the image sequence obtained by the strap-down image guidance system consists of intentional motion and shaky motion components, while the intentional motion components indicate the main motion of the guidance system, and the shaky motion components are the noises produced by jittery control and atmospheric turbulence during the flight of the missile. So, the Kalman filter is introduced to filter out the shaky components of the image sequences. Because the shaky components of the strap-down camera are Gaussian noise like high-frequency jitters, the intentional motion components with low frequency can be well reserved by Kalman filter, where the high-frequency jitters would be removed separately.

3. Global Motion Estimation

In this section, the global motion model of the missile-borne strap-down image guidance system is established based on the four-parameter similarity transformation model. Further, the global motion of video is firstly estimated by the optical flow algorithm.

The main motion of the image sequences captured by the strap-down image guidance system mainly consists of rotation, translation and scale transformation, while jitters of the projectile mainly

contribute to the high-frequency rotations and translations. Therefore, the four-parameter similarity transformation model is chosen to represent inter-frame global motion of the missile-borne image guidance system. The similarity transformation matrix is represented by T , which is given by

$$T = \begin{bmatrix} s \cos \theta & -s \sin \theta & dx \\ s \sin \theta & s \cos \theta & dy \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where θ is the rotation angle, S is the scale factor, dx and dy are translations of horizontal and vertical direction respectively. Thus, the corresponding feature points in two adjacent frames can be represented by

$$\begin{bmatrix} x_i^n \\ y_i^n \\ 1 \end{bmatrix} = T \cdot \begin{bmatrix} x_i^{n-1} \\ y_i^{n-1} \\ 1 \end{bmatrix}, \quad (2)$$

where (x_i^{n-1}, y_i^{n-1}) is the location of feature point i in the previous frame and (x_i^n, y_i^n) is the location of the corresponding feature point in the current frame.

Then, the global motion from the first frame to the current frame can be expressed as

$$X_n = T_n X_{n-1} = T_n (T_{n-1} X_{n-2}) = \dots = \left(\prod_{i=2}^n T_i \right) \cdot X_1, \quad (3)$$

where $X_n = (x_i^n, y_i^n)^T$ represents the location of a feature point in the current frame, $X_{n-1} = (x_i^{n-1}, y_i^{n-1})^T$ is the location of a feature point in the previous frame, and T_n represents the transformation matrix from X_{n-1} to X_n .

There are many different methods to extract the optical flow out of image sequences, among which the Lucas-Kanade (LK) optical flow algorithm [28] is one of the widely applied optical flow algorithms for it has less computational cost with acceptable accuracy [29,30]. Considering the background of this technology, the LK optical flow algorithm could be invalid in the case of fast movement and the error of the LK optical flow algorithm would accumulate rapidly if the objects of the video move too fast. In order to track the target which moves fast and to reduce the cumulative error of the LK optical flow algorithm, the pyramid LK optical flow algorithm [31] is introduced to solve this issue.

With the pyramid LK optical flow algorithm, the key points matching between adjacent frames can be conducted. The RANSAC algorithm is used to eliminate the mismatched keypoints and the matching pairs of foreground objects. In order to verify the tracking accuracy of optical flow method, experiments were carried out and the experimental results are shown in Figures 2 and 3, where there are three levels in the pyramid. The jitters are manually added to the original image sequence, so that the trajectory tracking results of optical flow can be compared with the real trajectory of the image sequence as shown in Figure 3. Three consecutive frames of the image sequence are presented to compare the images before and after stabilization. The image sequence with 450 frames and 360×360 resolution in Figure 2 is filmed by a handheld Daheng Mercury USB3 VISION digital camera. Images in Figure 2a are a sequence with jitters and images in Figure 2b show the stabilized image sequence.

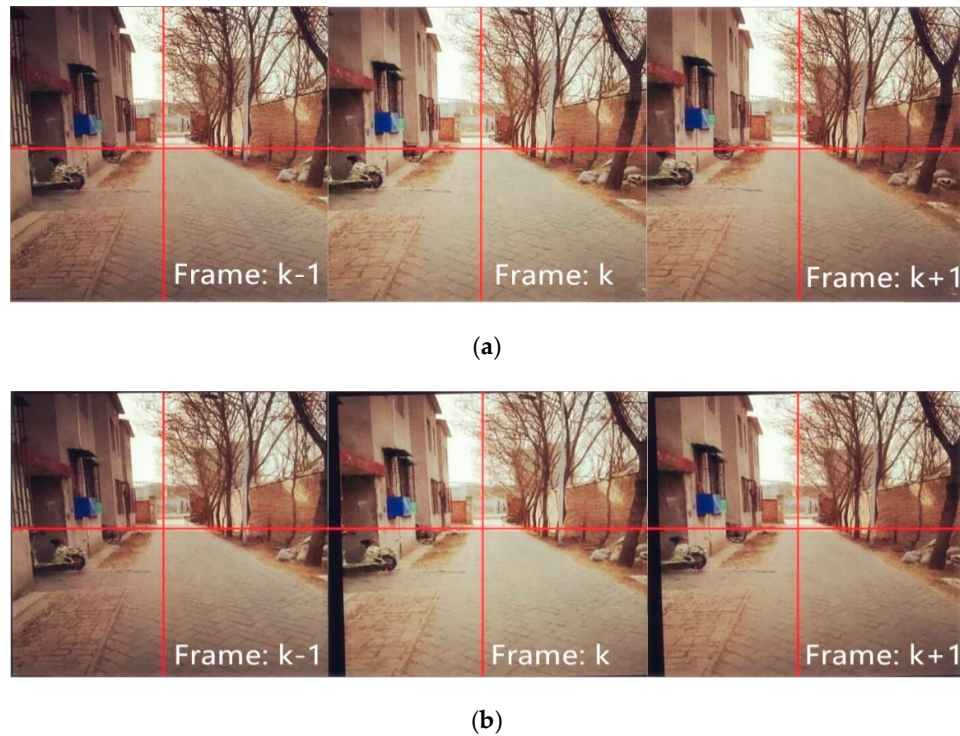


Figure 2. The image sequence that stabilized by optical flow. (a) Before stabilization; (b) After stabilization.

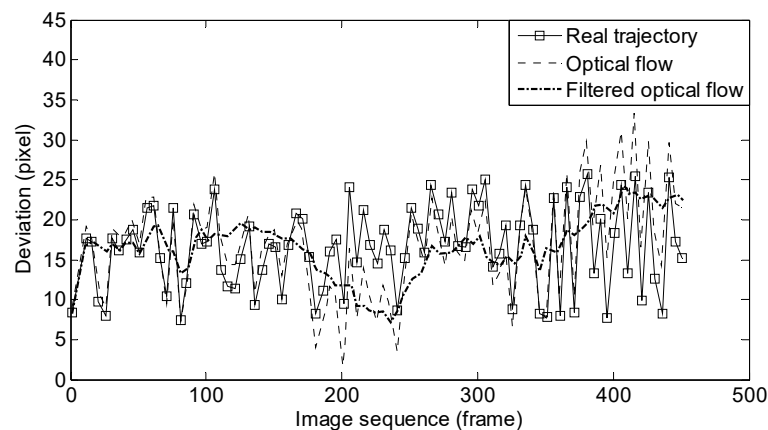


Figure 3. Simulation results of pyramid LK (Lucas-Kanade) optical flow algorithm.

Figure 3 shows the real trajectory, trajectory obtained by optical flow, and the filtered optical flow trajectory that conducted by Kalman filter. As shown in Figure 3, the trajectory obtained by optical flow gradually deviates from the real trajectory, and the tracking error of optical flow becomes larger when it is not compensated or corrected. In order to reduce the accumulative error of optical flow, an image matching algorithm with higher tracking accuracy is introduced to correct it.

4. Motion Trajectory Correction and Filtering Based on Binary Feature Descriptors Matching

To achieve accurate image matching with minimum computational cost, multiple image matching algorithms are compared and the FREAK algorithm is applied to obtain the transformation matrix between two images with a constant number of frames apart. The Kalman filter is designed based on the transformation matrix of the missile-borne image guidance system to filter out the high-frequency jitters of the intentional motion trajectory.

4.1. Trajectory Correction Based on FREAK Feature Descriptor

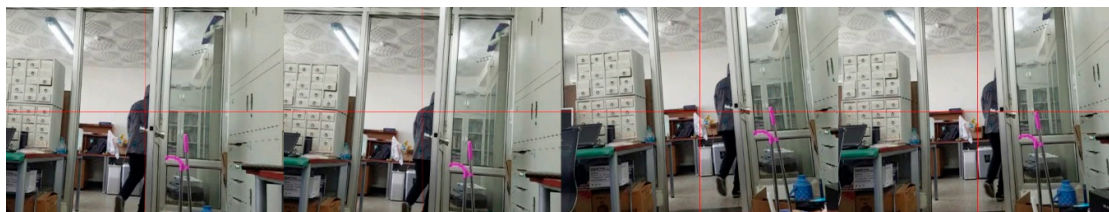
The point feature matching algorithms are state-of-art image matching algorithms for video stabilization, where the computational costs are also comparably higher. Schaeffer [32], Jared [33], and Bekele [34] compared and evaluated different binary key point descriptors such as BRIEF, BRISK, SURF, and FREAK, etc., and conclusions have been drawn that the accuracy of the FREAK algorithm is comparably better with less computational cost among these algorithms. A number of videos were tested with different feature matching algorithms, and the time consumptions are compared in milliseconds. The tests were carried out on a computer with Core i5 of 2.4 GHz, and RAM of 8 GB, with Visual Studio 2015 C++ and Open CV. The results are shown in Table 1, where the selection of feature points of the FREAK algorithm is conducted by SURF algorithm.

It can be seen from Table 1 that the efficiency of the FREAK algorithm is better than other binary feature point matching algorithms, which is consistent with other studies performed. The computational cost of the SIFT algorithm is the highest, and it is more than 300 ms per frame even at the resolution of 360×360 , which is not suitable for real-time implementation. As for the SURF algorithm, it is a little bit higher in computational cost compared to the FREAK algorithm. In addition, considering the FREAK algorithm has better scale invariance [32] performance, while the scale variation is the main factor as the missile approaching the target. So, the FREAK algorithm is selected to correct the tracking error of optical flow.

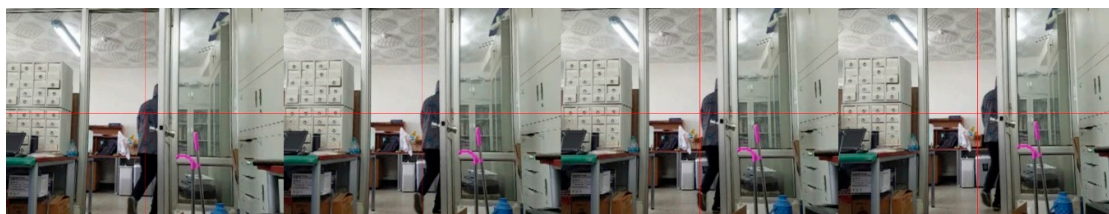
Table 1. The average time-consuming of each feature matching algorithm (ms).

Video Case Algorithms	1 (720 × 720 pixel)	2 (640 × 480 pixel)	3 (480 × 480 pixel)	4 (360 × 360 pixel)
SIFT	667	424	336	313
SURF	127	80	64	40
FREAK	105	60	51	32
Pyramid LK	27	18	13	9
Optical Flow				

The accuracy of the three-level pyramid LK optical flow algorithm and image matching based on the FREAK descriptors are tested and compared using an image sequence where jitters are manually added as well. The image sequence with 120 frames and 1280×720 resolution in Figure 4 is filmed by the Daheng Mercury USB3 VISION digital camera, where the false matching is detected by RANSAC algorithm. The stabilized image sequences are shown in Figure 4, and the tracking errors are given in Figure 5.



(a)



(b)



(c)

Figure 4. Stabilized image sequences. (a) Image sequence before stabilization; (b) Stabilized by optical flow; (c) Stabilized by image matching with FREAK (Fast Retina Keypoint) descriptor.

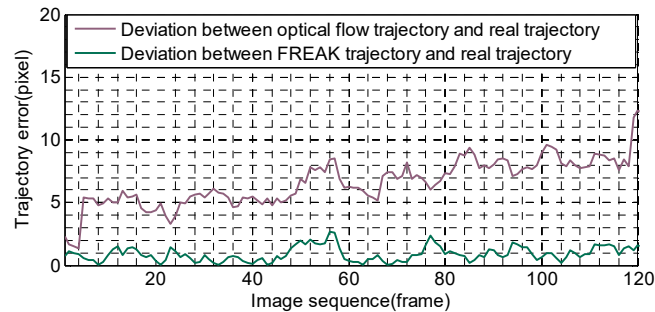
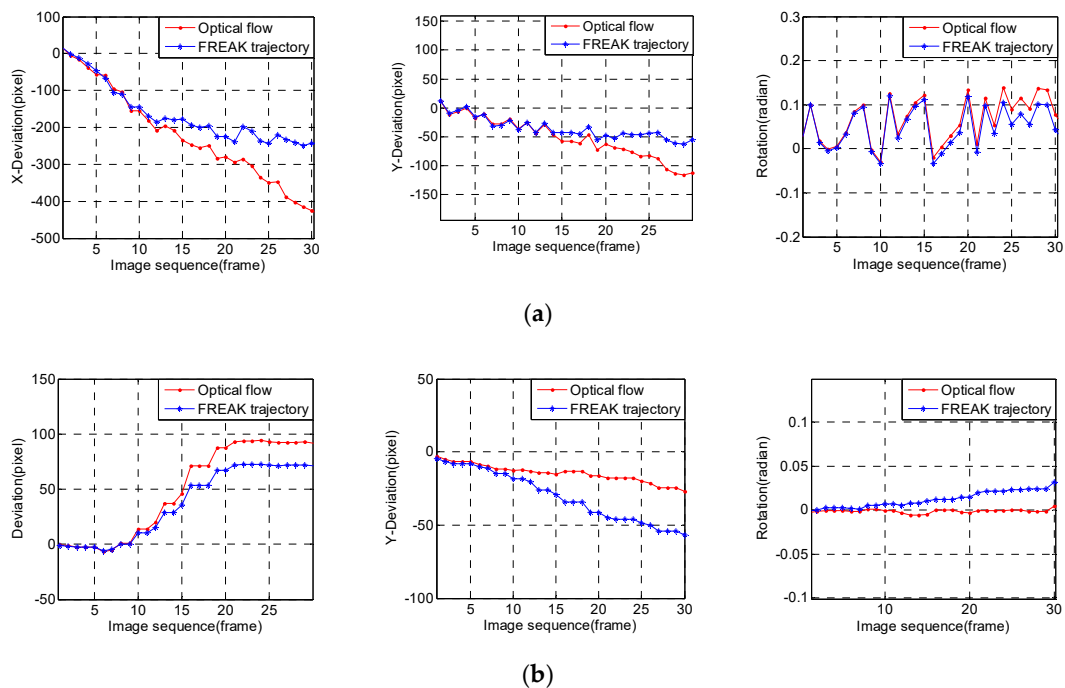


Figure 5. Trajectory tracking error of optical flow and image matching with FREAK (Fast Retina Keypoint) descriptor.

As shown in Figure 5, the image matching with FREAK descriptors has a much better performance of image stabilization accuracy than that of optical flow, which indicates that it can be used to correct the trajectory generated by the optical flow algorithm. In order to determine the period of point feature matching to correct the trajectory of optical flow, simulations were carried out to test the optical flow and FREAK algorithm. For the missile-borne image guidance system moves fast in the air, image sequences with moving objects were tested. The motion trajectories generated by optical flow and the FREAK algorithm are compared, as shown in Figure 6.



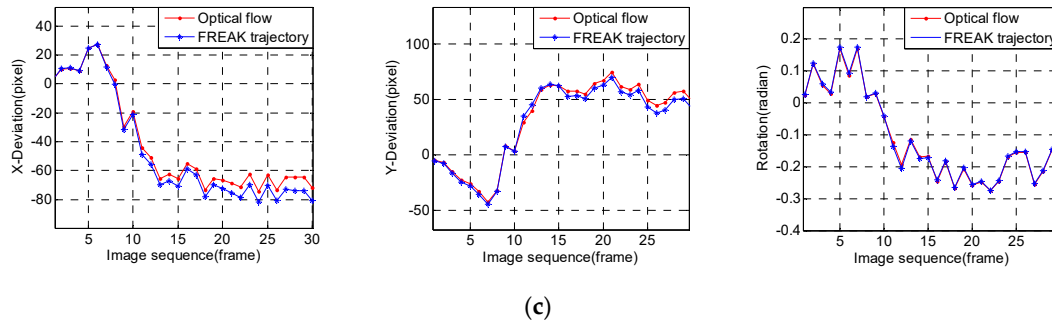


Figure 6. Comparison of the trajectories obtained by optical flow and FREAK (Fast Retina Keypoint) algorithm with different image sequences. (a) Image sequence of moving vehicle; (b) Image sequence of unmanned aircraft; (c) Image sequence of missile.

It can be seen from Figure 6 that the deviation between the trajectory obtained by optical flow and FREAK is getting larger from about the 10th frame, and there is no obvious difference between them before that point. To ensure high trajectory tracking accuracy with as low computational cost as possible, the trajectory tracking results of the optical flow method are corrected by FREAK once every 10 frames, and the trajectory of each frame is then filtered by a Kalman filter.

4.2. Motion Trajectory Filtering Based on Kalman Filter

To filter out the high-frequency noises of the motion trajectory, a Kalman filter is designed to obtain the intentional motion of the trajectory, for the motion of missile-borne image guidance system is predictable, which is suitable for the application of Kalman filter. According to the global motion model in Section 3, the state model of the trajectory can be expressed as

$$X(k) = [\theta(k), dx(k), dy(k), v_\theta(k), dv_x(k), dv_y(k)]^T, \quad (4)$$

where $v_\theta(k)$ is the velocity of rotation angle, $dv_x(k)$ and $dv_y(k)$ are velocities of pixels in x- and y-direction.

The measurements of the Kalman filter obtained by optical flow and the image matching algorithm are represented by

$$Z(k) = [\theta(k), dx(k), dy(k)]^T. \quad (5)$$

Assuming that the motion of adjacent frames is uniformly varying, according to Equations (4) and (5), the dynamic description model A and the observation model H are given by

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \quad (6)$$

With the equations above, the Kalman filter can be achieved through the following procedures, where the equations of the prediction phase of the Kalman filter from $k-1$ to k are given by

$$X(k|k-1) = A \cdot X(k-1|k-1), \quad (7)$$

$$P(k|k-1) = A \cdot P(k-1|k-1)A^T + Q, \quad (8)$$

where Q is the covariance matrix of processing noises and R is the covariance matrix of measurements noises of the Kalman filter.

The updated equations of the Kalman filter from $k-1$ to k are given as follows:

$$K_g(k) = P(k|k-1)H^T(HP(k|k-1)H^T + R)^{-1}, \quad (9)$$

$$X(k|k) = X(k|k-1) + K_g(k)(Z(k) - HX(k|k-1)), \quad (10)$$

$$P(k|k) = (I - K_g(k)H)P(k|k-1). \quad (11)$$

The motion trajectories obtained by FREAK, which has higher accuracy, are brought into the Kalman filter as the measurements to update the state matrix for every 10 frames, while the optical flow trajectory is taken as input of the Kalman filter at other times.

From the analysis above and the overall framework in Figure 1, the specific framework of the proposed algorithm is shown in Figure 7. By using this method, fast image stabilization can be realized. Experiments were carried out to verify the proposed method.

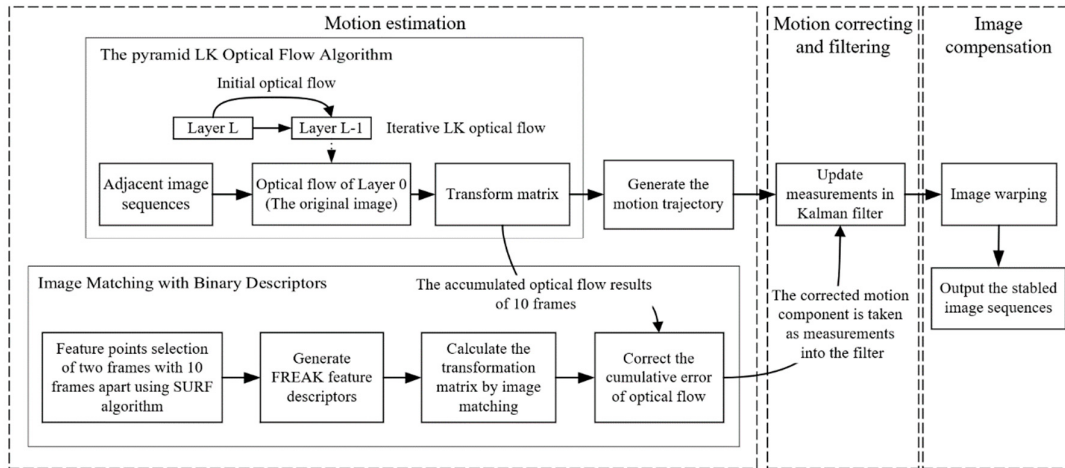


Figure 7. Block diagram of the specific framework of proposed algorithm.

5. Experimental Results

Experiments were conducted on a computer with an i5-8265U CPU, and 8 GB of RAM. The algorithms were implemented in Visual Studio 2015 C++ and OpenCV. To examine the accuracy of the proposed algorithm, the inter-frame transformation fidelity (ITF) which is based on the peak signal-to-noise ratio (PSNR) is used to evaluate the image stabilization results. The PSNR value between two consecutive frames represents how similar an image is to another by measuring the similarity between them, which is defined as follows,

$$\text{PSNR}(I_{k-1}, I_k) = 10 \log \frac{255^2}{\text{MSE}(I_{k-1}, I_k)}, \quad (12)$$

where I_{k-1} and I_k refer to the two adjacent frames, $\text{MSE}(I_{k-1}, I_k)$ refers to the mean square error of the two frames. $\text{MSE}(I_{k-1}, I_k)$ is defined as,

$$\text{MSE}(I_{k-1}, I_k) = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M (I_{k-1}(i, j) - I_k(i, j))^2, \quad (13)$$

where M and N refer to the number of rows and columns.

$$\text{ITF} = \frac{1}{N_{\text{frame}} - 1} \sum_{i=1}^{N_{\text{frame}} - 1} \text{PSNR}(I_{k-1}, I_k), \quad (14)$$

where N_{frame} is the total number of video frames.

By comparing the ITF values of video sequences before and after stabilization, the performance of the video stabilization algorithm can be evaluated, and the higher the ITF value, the better the image stabilization is.

In order to evaluate the performance of the proposed method objectively and show the advantages intuitively, the proposed method is compared with two state-of-the-art methods, the Deshaker [35] and the real-time video stabilization system presented by Hu et al. [27], on the accuracy performance and computational costs of image stabilization. The two publicly available video sequences named 0004TU and 2WL are used for tests, which can be downloaded on the website [36]. The frame numbers of 0004TU and 2WL are 450 and 500 respectively. The resolution of them is 1280×720 pixels. Figures 8 and 9 show the video stabilization results of different methods. The figures in Figures 8 and 9 were trimmed to remove the black peripheries after image stabilization. Table 2 shows the ITF results of the stabilized image sequences. Table 3 shows the computational cost of video stabilization for various methods.

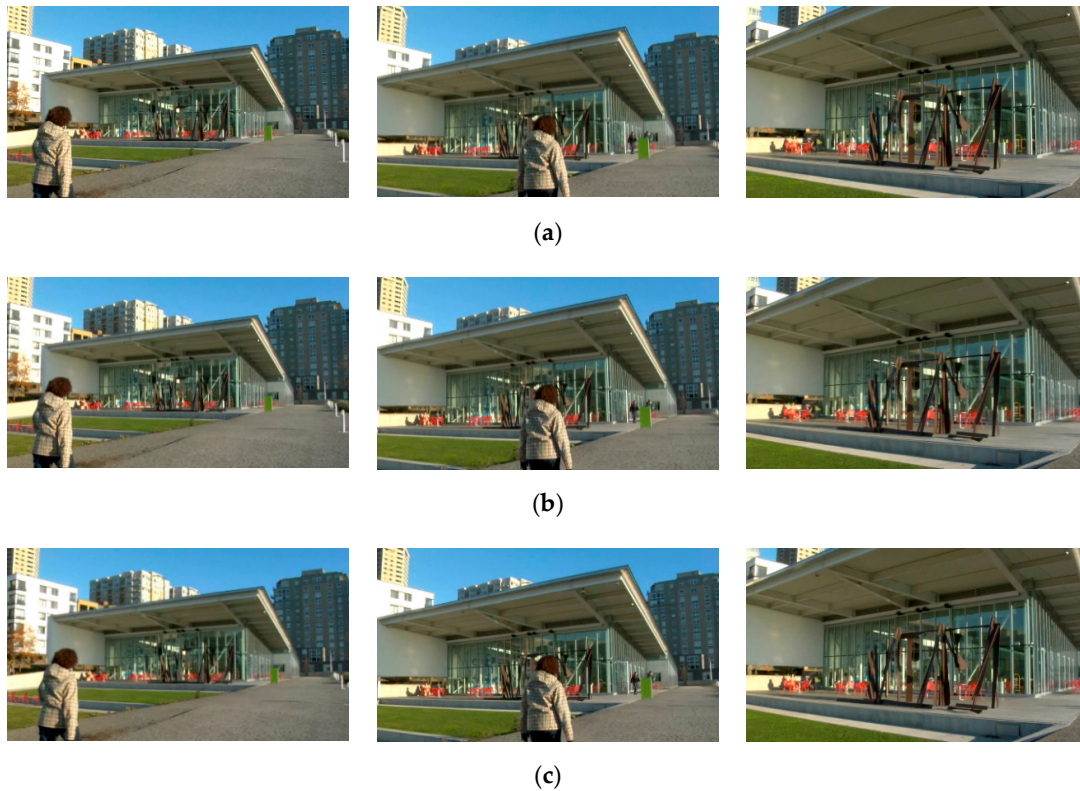


Figure 8. Video stabilization of 0004TU video sequence. (a) Original frames (30th, 200th, and 430th frames); (b) Stabilized by Deshaker; (c) Stabilized by proposed method.

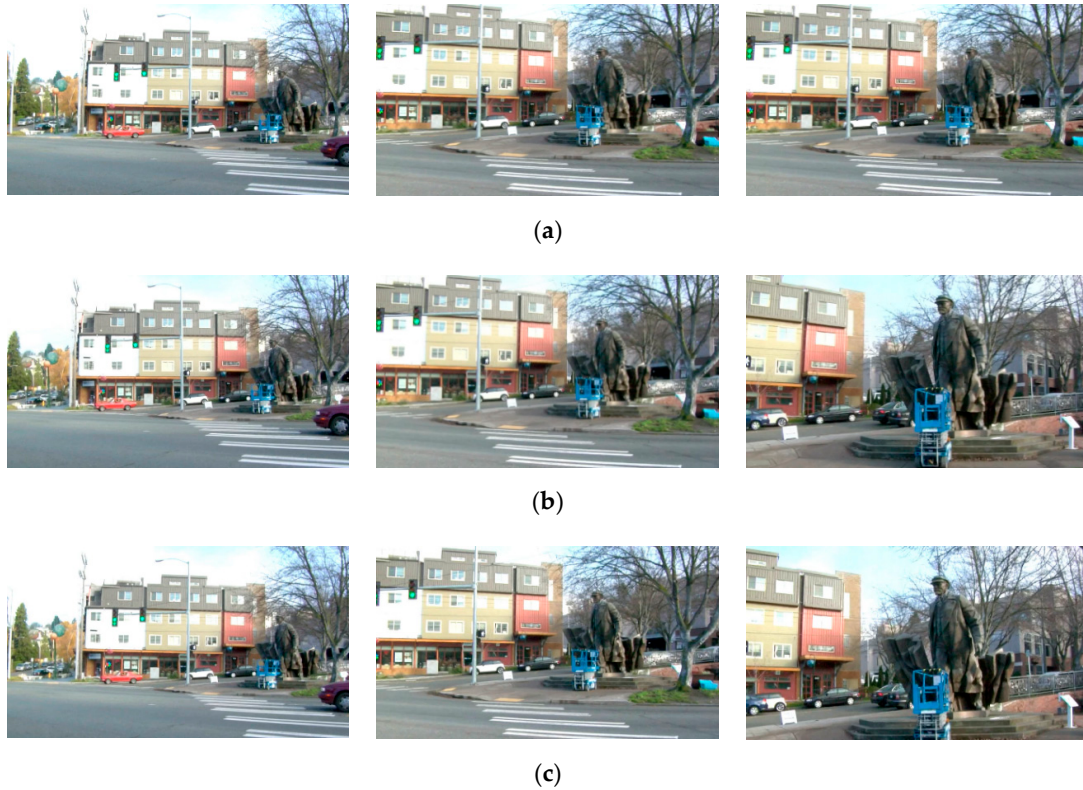


Figure 9. Video stabilization of 2WL video sequence. (a) Original frames (30th, 200th, and 430th frames); (b) Stabilized by Deshaker; (c) Stabilized by proposed method.

Table 2. ITF values of image sequences (dB).

	0004TU	2WL
Original	19.5356	17.6181
Deshaker	22.0803	21.1841
Hu et al.	20.43	19.30
Proposed method	21.9681	19.97

Table 3. Computational cost of video stabilization for various methods (ms/frame).

		0004TU	2WL
In our condition ¹	Deshaker	55.5	54.25
	Proposed method	46.3	48.0
In Hu's condition ²	Deshaker	34.51	34.20
	Hu et al.	31.01	30.07

¹ Computer with an Intel i5-8265U CPU and 8GB of RAM. ² Computer with an Intel i7-4790 CPU and 16GB of RAM.

As shown in Table 2, the Deshaker has the best performance of the video stabilization in both 0004TU and 2WL image sequences, where the ITF values are improved with 2.54 dB and 3.56 dB, respectively. The ITF values of the image sequences stabilized by Hu et al.'s method improved with 0.89 dB in 0004TU and 1.68 dB in 2WL compared to the original image sequences, while the proposed method improved the ITF values with 2.43 dB in 0004TU and 2.35 dB in 2WL. As a result, the proposed method has a better accuracy performance than Hu et al.'s method in both cases, but a bit worse than the Deshaker. And it can be noticed from Table 3 that the proposed method has

approximate performance in 0004TU compared to Deshaker, which is about 0.1 dB less improved. The ITF value of proposed method in 2WL is 1.2 dB less improved than the Deshaker.

For the difference of the experimental environment between the proposed method and Hu et al., it is not fair to compare the computational cost of proposed method directly. So, the Deshaker has been taken as the reference to examine the computational cost of different methods, for it has the same characteristics in different computers. Table 3 shows that the computational costs of Deshaker are 55.5 ms/frame and 54.25 ms/frame in our hardware environment, which is 1.2 times higher than our proposed method. Meanwhile, the experimental results of the method of Hu et al. show that the computational costs of Deshaker are 34.51 ms/frame and 34.20 ms/frame, which is 1.1 times higher than Hu et al.'s method in their hardware environment. Thus, the proposed method has the best performance on computational cost among these methods.

To ensure that the algorithm can be applied to various of harsh application environments, and validate the robustness of it, the algorithm on another 15 image sequences are tested as well, which can be obtained publicly on the website [37].

The first frames of these stabilized image sequences are shown in Figure 10. And their features and tracking results are given in Table 4.



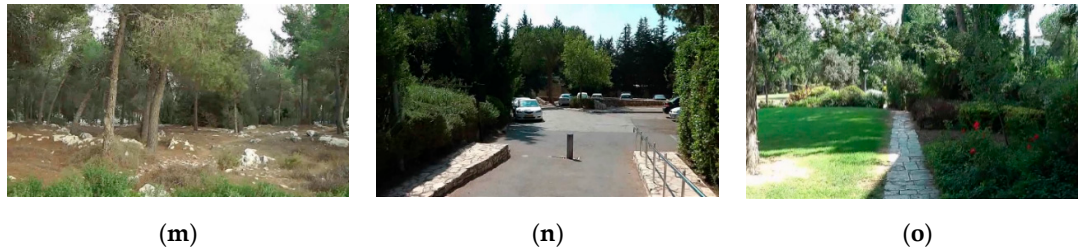


Figure 10. The tested image sequences. (a) Image sequence 2; (b) Image sequence 3; (c) Image sequence 4; (d) Image sequence 5; (e) Image sequence 6; (f) Image sequence 7; (g) Image sequence 8; (h) Image sequence 9; (i) Image sequence 10; (j) Image sequence 11; (k) Image sequence 12; (l) Image sequence 13 (m) Image sequence 14; (n) Image sequence 15; (o) Image sequence 16;.

Table 4. Features and ITF values of image sequences (dB).

Video Name	Resolution	Total Number of Frames	Original Image Sequence (ITF)	Stabilized by Proposed Method (ITF)
2	640 × 360	449	19.366389	23.760345
3	640 × 360	574	24.587738	27.230761
4	640 × 360	401	26.471917	26.996680
5	640 × 360	599	21.075880	22.299292
6	640 × 360	434	21.181671	23.129672
7	640 × 360	389	23.744771	24.069733
8	640 × 360	434	21.506194	25.739252
9	640 × 360	999	15.904668	17.523298
10	640 × 360	404	14.737232	16.787729
11	640 × 360	434	16.129186	21.106888
12	640 × 360	494	17.427601	21.500074
13	640 × 360	509	18.040385	20.447358
14	640 × 360	299	16.936780	21.077630
15	640 × 360	479	15.899816	20.115547
16	640 × 360	449	14.282314	18.031555

By analyzing the ITF values of these stabilized videos in Table 4, the ITF values have been improved by 2–5 dB compared to the original videos in the great majority of cases. However, the ITF value of Video 4 has only been improved by about 0.52 dB. This is because the original image sequence already has the comparably higher ITF value of 26.47 dB. In addition, the proposed algorithm is used to distinguish and maintain the low-frequency motion in the image sequence, and eliminate the high-frequency jitters in the video which are caused by the shaky movements of missile. So, the low-frequency movements in Video 4 are preserved after processing, while a small number of high-frequency jitters are removed, and the ITF value is not improved significantly.

Generally, the proposed method is verified with multiple video cases and compared with the existing methods. The experimental results in Tables 2 and 3 show that the accuracy performance of proposed method is close to that of Deshaker and it performs best in computational cost. Considering the background of this technology, the Deshaker method is the most difficult to apply in practice due to its poor real-time performance, while the method of Hu et al. has a lower accuracy with higher time cost than the proposed method. In addition, the simulations carried out in Figure 4 show that the proposed method has a good performance on the stability and robustness in various conditions. As a result, the proposed method has the advantages of real-time performance compared to previous methods with preferable accuracy, and it has been proven as well to be steady and robust, which is applicable in the strap-down image guidance system.

6. Conclusions

The results of comparative simulations show that the proposed method can be competitive to the existing state-of-art image stabilization methods, whether in terms of accuracy or real-time performance, making it suitable to meet the real-time and high accuracy requirements of a strap-down image guidance system.

Aiming at the problem of imaging dithering and real-time requirements of the strap-down missile-borne image guidance system, a real-time electronic image stabilization algorithm which combines the optical flow and binary feature matching algorithm is proposed. The overall framework of this algorithm is proposed at first. Then, the global motion of the image sequences is established and estimated by the pyramid LK optical flow algorithm. The efficiency of different image matching algorithms is analyzed, and the FREAK binary feature matching algorithm is selected to correct the cumulative error caused by optical flow with an interval of 10 frames. The Kalman filter is introduced to filter out the high-frequency jitters of the motion trajectory. The experimental results show that the image sequences can be stabilized by the proposed algorithm with comparably less computational cost and better accuracy. The proposed algorithm has not been implemented on an image guidance system for the immature test conditions. The implementation of this algorithm on board will be investigated in the future.

Author Contributions: Conceptualization, Q.S. and C.L.; Data Curation, Z.D.; Formal Analysis, Z.D.; Funding Acquisition, Q.S.; Investigation, C.L. and Y.D.; Methodology, C.L.; Project Administration, D.Y.; Resources, D.Y.; Software, Z.D. and Y.D.; Supervision, D.Y.; Validation, Z.D. and Y.D.; Visualization, Z.D.; Writing—Original Draft Preparation, Z.D. and X.Z.; Writing—Review and Editing, X.Z.

Funding: This research received no external funding. And the APC was funded by Qiang Shen.

Acknowledgments: We would like to thank the National Key Laboratory of Science and Technology on electromechanical dynamic characteristics at the Beijing Institute of Technology for their support during the performing in this paper.

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

References

1. Felix, S. *US Navy Spike Missile System: A new Generation of Miniature Precision Guided Weapons*; Naval Air Warfare Center Weapons DIV China Lake CA, California, USA: 2006.
2. Rudin, R.T. Strapdown Stabilization for Imaging Seekers. In Proceedings of the 2nd Annual AIAA and SDIO Interceptor Technology Conference, China Lake, CA, USA, 6–9 June 1993.
3. Özkan, B.; Uçar, A. Comparison of the strapdown and gimbaled seekers utilized in aerial applications. In Proceedings of the Conference on Infrared Technology and Applications XXXVIII, Baltimore, MD, USA, 23–27 April 2012; Volume 8353, id. 83530K.
4. Yuan, Y.; Yang, X.; Wu, W.; Li, H.; Liu, Y. A fast single-image super-resolution method implemented with CUDA. *J. Real Time Image Process.* **2018**, *16*, 81–97.
5. Georgis, G.; Lentaris, G.; Reisis, D.I. Acceleration techniques and evaluation on multi-core CPU, GPU and FPGA for image processing and super-resolution. *J. Real Time Image Process.* **2016**, *16*, 1207–1234.
6. Yang, W.; Zhang, Z.; Zhang, Y.; Lu, X.; Li, J.; Shi, Z. A real-time gray projection algorithm for electronic image stabilization. In Proceedings of the SPIE 8193, International Symposium on Photoelectronic Detection and Imaging 2011: Advances in Infrared Imaging and Applications, 81932G, Beijing, China, 8 September 2011.
7. Tian, S.; Zhao, P.; Wang, N.; Wang, C. Aims at moving objects' improvement based on gray projection of algorithm of the electronic image stabilization. In Proceedings of the 2010 3rd International Congress on Image and Signal Processing, Yantai, China, 16–18 October 2010; Volume 5, pp. 2483–2487.
8. Vella, F.; Castorina, A.; Mancuso, M.; Messina, G. Digital image stabilization by adaptive block motion vectors filtering. *IEEE Trans. Consum. Electron.* **2002**, *48*, 796–801.
9. Zhu, S.; Ma, K. A new diamond search algorithm for fast block-matching motion estimation. *IEEE Trans. Image Process.: A Publ. IEEE Signal Process. Soc.* **2000**, *9*, 287–290.

10. Lee, S.; Lee, K.; Ko, S.C. Digital Image Stabilizing Algorithms Based on Bit-plane Matching. In Proceedings of the International 1998 Conference on Consumer Electronics, Los Angeles, CA, USA, 2–4 June 1998; pp. 126–127.
11. Ko, S.; Lee, S.; Jeon, S. Fast digital image stabilizer based on Gray-coded bit-plane matching. 1999 Digest of Technical Papers. In Proceedings of the International Conference on Consumer Electronics, Los Angeles, CA, USA, 22–24 June 1999; pp. 90–91.
12. Beauchemin, S.S.; Barron, J.L. The Computation of Optical Flow. *ACM Comput. Surv. (CSUR)* **1995**, *27*, 433–467.
13. Chui, H.; Rangarajan, A. A new point matching algorithm for non-rigid registration. *Comput. Vision Image Underst.* **2003**, *89*, 114–141.
14. Battiato, S.; Gallo, G.; Puglisi, G.; Scellato, S. SIFT Features Tracking for Video Stabilization. In Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP 2007), Modena, Italy, 10–14 September 2007; pp. 825–830.
15. Lowe, D.G. Object Recognition from Local Scale-Invariant Features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–25 September 1999; pp. 1150–1157.
16. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
17. Bay, H.; Ess, A.; Tuytelaars, T.; Gool, L.V. Speeded-up robust features (surf). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359.
18. Cheng, X.; Hao, Q.; Xie, M. A Comprehensive Motion Estimation Technique for the Improvement of EIS Methods Based on the SURF Algorithm and Kalman Filter. *Sensors* **2016**, *16*, 486.
19. Pinto, B.; Anurenjan, P.R. Video stabilization using Speeded Up Robust Features. In Proceedings of the 2011 International Conference on Communications and Signal Processing, Calicut, India, 10–12 February 2011; pp. 527–531.
20. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. BRIEF: Binary Robust Independent Elementary Features. In *Lecture Notes in Computer Science, Proceedings of the Computer Vision—ECCV 2010, 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5–11 September 2010*; Springer: Berlin/Heidelberg, Germany, 2010.
21. Leutenegger, S.; Chli, M.; Siegwart, R.Y. BRISK: Binary Robust invariant scalable keypoints. In Proceedings of the International Conference on Computer Vision. IEEE Computer Society, Barcelona, Spain, 6–13 November 2011; pp. 2548–2555.
22. Alahi, A.; Ortiz, R.; Vanderghyeyns, P. FREAK: Fast Retina Keypoint. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, Providence, Rhode Island, 16–21 June 2012; pp. 510–517.
23. Dong, J.; Liu, H. Video Stabilization for Strict Real-Time Applications. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *27*, 716–724.
24. Lim, A.; Ramesh, B.; Yue, Y.; Cheng, X.; Zhi, G.; Feng, L. Real-time optical flow-based video stabilization for unmanned aerial vehicles. *J. Real Time Image Process.* **2019**, *16*, 1975–1985.
25. Hu, W.; Chen, C.; Chen, T.; Peng, M.; Su, Y. Real-time video stabilization for fast-moving vehicle cameras. *Multimed. Tools Appl.* **2016**, *77*, 1237–1260.
26. Derpanis, K.G. Overview of the RANSAC algorithm. Department of Computer Science and Engineering, York University, **2010**, *1*, 2.
27. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395.
28. Lucas, B.D.; Kanade, T. An iterative image registration technique with an application to stereo vision. *Int. Joint Conf. Artif. Intell. (IJCAI)* **1981**, *81*, 674–679.
29. Thota, S.D.; Vemulapalli, K.S.; Chintalapati, K.; Gudipudi, P.S.S. Comparison Between the Optical Flow Computational Techniques. *Int. J. Eng. Trends Technol. (IJETT)* **2013**, *4*, 4507–4511.
30. Liu, H.; Hong, T.H.; Herman, M.; Camus, T.; Chellappa, R. Accuracy vs efficiency trade-offs in optical flow algorithms. *Comput. Vis. Image Underst.* **1998**, *72*, 271–286.
31. Bouguet, J.Y. Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm. *Intell. Corp.* **2001**, *5*, 4.
32. Chatoux, H.; Lecellier, F.; Fernandez-Maloigne, C. Comparative study of descriptors with dense key points. In Proceedings of the 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016.

33. Heinly, J.; Dunn, E.; Frahm, J.M. Comparative Evaluation of Binary Features. In *Lecture Notes in Computer Science, Proceedings of the 2012 European Conference on Computer Vision (ECCV), Florence, Italy, 7–13 October 2012*; Springer: Berlin/Heidelberg, Germany, 2012.
34. Bekele, D.; Teutsch, M.; Schuchert, T. Evaluation of binary keypoint descriptors. In *Proceedings of the 2013 IEEE International Conference on Image Processing, Melbourne, VIC, Australia, 15–18 September 2013*; pp. 3652–3656.
35. Deshaker Method. Available online: <http://www.guthspot.se/video/deshaker.html> (accessed on 14 September 2014).
36. Video Figures. Available online: <http://web.cecs.pdx.edu/~fliu/project/3dstab.htm> (accessed on 15 June 2019)
37. Simple. Available online: <http://liushuaicheng.org/SIGGRAPH2013/database.html> (accessed on 18 June 2019)



© 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/>).