

## Unmanned Aerial Vehicle Targeting Based on Image Matching

Zhang Yan\*, Li Jianzeng\*, Li Deliang\*, Zhou Zidong\*, Du Yulong\* and Ji Weidong\*\*

*\*Ordnance Engineering College, Hebei, China*

*\*\*College of Computer Science and Information Engineering Harbin Normal University, China*

*Emails: hillwind@126.com ljz681@sohu.com 271395592@qq.com*

*zzdmichael@163.com 1159316475@qq.com 1209301925@qq.com*

### **Abstract**

*Aiming at the problem of low accuracy of the target for one UAV (Unmanned Aerial Vehicle), an UAV target location method based on image matching is presented according to the data of video and photograph as well as the information of the digital map. Firstly, using the improved artificial matching method, the UAV aerial images are jointed based on the digital map to obtain the coordinates information. Then, the UAV video frame resolution containing the target is improved by the improving method of super-resolution reconstruction. Finally, making use of the improved SIFT algorithm, the video frames are registered with the aerial photo to complete the location of the target. Results prove that this method can achieve the target location without the UAV's position and attitude information, while improving the precision and speed of target location, this method is an effective one for target location based on the image matching in the battlefield.*

**Keywords:** *Image Registration, UAV, Target location, SIFT, Super-resolution reconstruction*

### **1. Introduction**

Technological progress and the need of actual combat promote the development of UAV targeting technology. Nowadays, to position the target quickly and accurately is one of core problems in the UAV research field. There are major principle errors in the real-time target model adopted in some UAV under the condition of complex terrain, and lack of the capability to locate continuously the moving target [1]. In recent years, with the rapid development of the computer visual relevant algorithm performance [2] and UAV technology, it's possible to achieve accurate position using image matching, in which the matching algorithm is the core content.

D. G Lowe and M. Brown [3] put forward an image mosaic algorithm based on SIFT invariant in 2004, which has strong adaptability and high accuracy, but there are questions such as bad robust performance and it can't solve the problem of matching the hidden image. In 2004, Zhao Xiang Yang and Du Limin [4] utilized the Harris algorithm to pick up corner and match them at the same time, then estimated the transformational matrix of image which is going to be matched by robust transformation, the algorithm can not only well adapt to the matching problems of affine transformational image, scene hidden image and uneven gray scale in image caused by illumination variation, but also the robustness is good. There are remaining issues such as the complex calculation, and settled distance ratio threshold that can't adaptively deal with different images. Li Lingling [5] put forward an algorithm which let Harris-Affine and SIFT complementary with same

characteristic, it's free from the influence of illumination, and affine transformation and image sequence scale, and make vector graph's automatic matching possible. But due to the complexity of Harris-Affine's iterative algorithm, the operation rate will be slow.

Aiming at the problems in the SIFT, scholar at home and abroad propose many modified algorithms in different way, most of them major in simplifying the generative process of characteristic vector, by which they solve the matching of noise image and affine image, but the distance ratio threshold in the algorithm can't adjust adaptively, so we can't assure the matching result's accuracy.

In this paper, we can not only provide still image and dynamic video, but also the advantage of digital map's application by using the UAV sensor, then propose a UAV target position method based on image matching technology, under the condition of no UAV location and attitude information, utilizing super-resolution rebuild and image matching method enhance UAV's target position accuracy, which will make sense to UAV's quick and accurate position.

## 2. The Brief Introduction of the Target Position Process

The target position process can be divided into the preparation stage and the application stage.

The preparation stage: launch the UAV to take an aerial image at the target area, utilize the modified artificial source pitching method, then obtain the aerial photo Montage by pitching the aerial photo in digital map for target, at the same time gain the coordinates of any point within the Montage;

The application stage: launch the UAV to camera to the target area, apply aerial photo prepared in the preparation, improve UAV's video resolution by using modified super-resolution reconstruction method, and then enhance the matching accuracy, then use the modified SIFT algorithm to match video frame containing target with the benchmark of Montage, in the way of making the Montage as ink, transfer the coordinate information to the target frame, consequently, achieve the target position.

This method can realize target position without UAV's location and attitude information, by the way, avoid ground control points' introduction and orthographical correction, and improve target position's accuracy and rate, which is an effective method for target position.

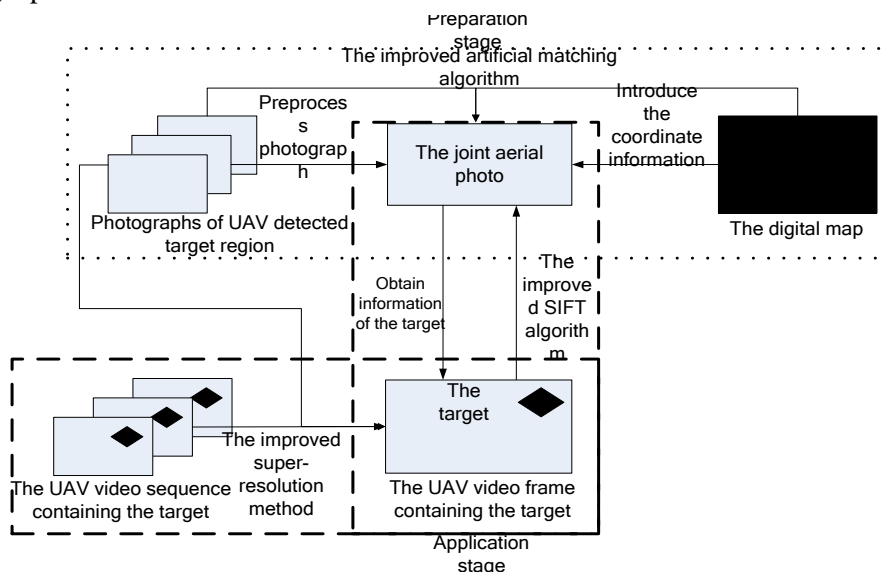


Figure 1. Diagram of the Proposed Algorithm

### 3. The Improved Super-Resolution Reconstruction Algorithm

Target frame resolution can improve the registration accuracy, thus improving the positioning accuracy. Based on the characteristics of the UAV detectors, which can simultaneously video and photograph, and the aerial photo resolution is much higher than the video frame, so the method here is based on reference [6]. However, reference [6] algorithm cannot connect high-frequency with low-frequency continuously, and the motion estimation method has dependencies on matching size and the constrained optimization algorithm for edge-preserving capacity is not high [7-8]. Aiming at these problems above, an improved method is proposed:

Firstly, use the Gaussian pyramid flow algorithm to replace the multiple scales affine block matching method based on least squares to estimate the low resolution video frame and the high resolution images. Secondly, use the wavelet frequency band decomposition method to extract details of the high frequency, then compensate the video according to the motion estimation results. Finally, use the Projection onto Convex Sets method to replace the algorithm of maximum posteriori optimization to iterative the compensated video. This method improves the fidelity of the reconstructed image and is simple and efficient for video super-resolution.

POCS algorithm are some of the closed graphical constraint sets defined in the vector space, these constraint sets are projected by initial estimation, then the intersection is the high-resolution images that has estimated.

$$G^{(1)}(i,j)=\{z^{(k)}(r,s):|r^{(1)}(i,j)|\leq\delta^{(1)}(i,j)\} \quad (1)$$

$$r^{(1)}(i,j)=y^{(1)}(i,j)-\sum z^{(k)}(r,s)A^{(1,k)}(r,s;i,j) \quad (2)$$

$$l=k-\frac{M-1}{2}, \dots, k, k+\frac{M-1}{2}, \forall(i,j) \in \theta^{(1)} \quad (3)$$

$$P(I_j)[x^{(k)}(r,s)]=x^{(k)}(r,s)+\begin{cases} \frac{(r^{(1)}(i,j)-\delta^{(1)}(i,j))A^{(1,k)}(r,s;i,j)}{\sum_{\sigma_1}\sum_{\sigma_2}A^{(1,k)^2}(\sigma_1,\sigma_2;i,j)}, & r^{(1)}(i,j)>-\delta^{(1)}(i,j) \\ 0 & -\delta^{(1)}(i,j)\leq r^{(1)}(i,j)\leq\delta^{(1)}(i,j) \\ \frac{(r^{(1)}(i,j)+\delta^{(1)}(i,j))A^{(1,k)}(r,s;i,j)}{\sum_{\sigma_1}\sum_{\sigma_2}A^{(1,k)^2}(\sigma_1,\sigma_2;i,j)}, & r^{(1)}(i,j)<-\delta^{(1)}(i,j) \end{cases} \quad (4)$$

And,  $\theta^{(1)}$  represents the support domain of the first observation image;  $\delta^{(1)}(i,j)$  is the statistical uncertainty of the model, which is ensured by the standard deviation on the process of the noise of the time domain and spatial domain;  $z^{(k)}(r,s)$  indicates the pixel of the pieces of high-resolution images;  $y^{(1)}(i,j)$  denotes the pixel of the first observed image;  $A^{(1,k)}(r,s;i,j)$  is the degradation model from the pieces of high-resolution images to the first low-resolution images;  $r^{(1)}(i,j)$  is deviation which is between the results of

the high resolution image  $z^{(k)}(r, s)$  that is processed by the degradation model and the low-resolution image observed,  $P(i, j)$  is the consistency projection operator,  $G^{(l)}(i, j)$  is the Convex and constraint set [7].

$A^{(l,k)}(r, s; i, j)$  linking  $z^{(k)}(r, s)$  and  $y^{(l)}(i, j)$  becomes the residual item  $r_t^{(l)}(i, j)$ , if the residual term surpass the initial threshold value setting, the residual item is back projected to the current high-resolution image estimated by the projection operator  $P^{(l)}(i, j)$ . Then, becoming the new high-resolution estimated image belongs to  $G^{(l)}(i, j)$  of the Estimate image, iterating until the change between two consecutive estimation is below the initial threshold value.

But due to atmospheric turbulence, the posture change of the UAV and the effect of sensor's error, cause the video and image data acquired exist noise, geometric distortion, and the problems of bad frames when it is serious. So it is necessary to preprocess data. This article focuses on the use of adaptive Wiener filtering, bad frame removed to ensure a better positioning accuracy.

#### 4. The Improved SIFT Algorithm

As the time, position and angle are different between two shot of the UAV, the video frames of the same scene are different, too. The method based on SIFT feature matching [9] is chosen to deal with this kind of non-homologous images. But the traditional SIFT algorithm's threshold is fixed, it cannot accurately handle different images adaptively. To solve this problem, this paper introduced the C-means clustering method [10] to classify the Euclidean distance ratio and find the optimal thresholds to adaptively match different images. Firstly, the Gaussian differential scale space is generated to get the position and orientation of the feature points, thereby generating the feature point descriptor. Then, the optimal threshold of the Euclidean distance ratio is adaptively generated to confirm the matching points. The SIFT algorithm scale-space formula [11] is as follows:

$$D(x, y, \sigma) = \left( \frac{1}{2\pi(k\sigma)^2} e^{-\frac{(x^2+y^2)}{2(k\sigma)^2}} - \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \right) I(x, y) \quad (5)$$

Each pixel of the scale space is compared with the adjacent eight pixels on the same layer as well as the nine pixels on the layers above and below to get the extreme points. Removal of the low contrast and the upper edge extreme points, the rest are the feature points. What follows next is the formula of calculating the modulus value and direction of the feature point.

$$m(x, y, \sigma) =$$

$$\sqrt{\left[\frac{1}{2\pi\sigma^2} e^{-\frac{-(x+1)^2+y^2}{2\sigma^2}} I(x+1,y) - \frac{1}{2\pi\sigma^2} e^{-\frac{-(x-1)^2+y^2}{2\sigma^2}} I(x-1,y)\right]^2 + \left[\frac{1}{2\pi\sigma^2} e^{-\frac{-(x^2+(y+1)^2)}{2\sigma^2}} I(x,y+1) - \frac{1}{2\pi\sigma^2} e^{-\frac{-(x^2+(y-1)^2)}{2\sigma^2}} I(x,y-1)\right]^2} \quad (6)$$

$$\theta(x,y,\sigma) = \arctan \frac{\frac{1}{2\pi\sigma^2} e^{-\frac{-(x^2+(y+1)^2)}{2\sigma^2}} I(x,y+1) - \frac{1}{2\pi\sigma^2} e^{-\frac{-(x^2+(y-1)^2)}{2\sigma^2}} I(x,y-1)}{\frac{1}{2\pi\sigma^2} e^{-\frac{-(x+1)^2+y^2}{2\sigma^2}} I(x+1,y) - \frac{1}{2\pi\sigma^2} e^{-\frac{-(x-1)^2+y^2}{2\sigma^2}} I(x-1,y)} \quad (7)$$

Rotating the axis to the main direction of the feature point, the  $16 \times 4 \times 4$  sub-windows are built up on the center of the feature point. Statistical the gradient orientations of the pixels of the sub-windows using a histogram, each sub-window contains eight directions. Finally the 128-dimensional feature vector is formed. For each point on the original image, finding the nearest two feature points according to the Euclidean distance. If the ratio of the nearest Euclidean distance feature points and the second nearest Euclidean distance feature points is less than the threshold value, then they are considered as the corresponding points [12].

The Euclidean distance ratio is classified into two parts using the C-means clustering method to find the best threshold of the matching effect. The C-means clustering method is based on the guidelines of the sum of the error square [13-14]. The specific algorithm is as follows:

(1) All the Euclidean distance ratios as samples are randomly divided into two clusters, then calculate the mean of the two clusters  $m_1$ ,  $m_2$  and  $J_e$ .

If  $N_i$  is the number of samples in the  $i$ -th cluster,  $m_i$  is the mean of these samples, where  $i \in [1-2]$ , namely

$$m_i = \frac{1}{N_i} \sum_{y \in \Gamma_i} y \quad (8)$$

Then the sum of the error squares of each sample  $y$  and the mean  $m_i$  is

$$J_e = \sum_{i=1}^c \sum_{y \in \Gamma_i} \|y - m_i\|^2 \quad (9)$$

- (2) Select an alternate sample  $y$ ,  $y$  is now set in.
- (3) If  $N_i=1$ , then turn to (2), otherwise continue.
- (4) Calculate

$$\rho_j = \begin{cases} \frac{N_j}{N_j+1} \|y - m_j\|^2 & j \neq i \\ \frac{N_i}{N_i+1} \|y - m_i\|^2 & j = i \end{cases} \quad (10)$$

(5) For all  $j$ , if  $\rho_k \leq \rho_j$ , put  $y$  from  $\Gamma_i$  to  $\Gamma_k$ .

(6) Recalculate the value of  $m_i, m_k$  and modify  $J_e$ .

(7) If  $J_e$  do not change after  $N$  times of the successive iterations, then stop the

iteration and output the information of classify and the threshold. If  $J_e$  changes, then go to step (2). The information here includes the number of types in the sample and the sample value, as well as the threshold taken near the boundary of the sample mean.

The C-means clustering method divides the Euclidean distance ratio of all the features points into two types. The boundary of the two types is the threshold. This will enhance the matching correct rate, and can adaptively process different images with stability. After getting the transfer matrix according to the corresponding points, the target is transferred to digital maps. Since the digital map is Orthogonal projection, the result is also Orthogonal projection, thus avoiding the complexity of introducing control points, ortho-correction, *etc.* so the localization speed is improved.

UAV aerial photos differences widely with digital maps (satellite photos). Based on the issues of matching directly, this paper presents an improved matching method with artificial choice of points. Firstly select appropriate number of regions artificially. Then find the corresponding points in each selected region using the improved SIFT method, and then obtain the transfer matrix of each region, and finally fit the entire transfer matrix using the method of least squares to obtain the transfer matrix of two images, accomplishing the matching. Compared with the traditional method of artificially selecting points, this method is faster and matches more accurate, thus solving the matching problem of non-homologous image.

## 5. Simulation Analysis

The simulation is tested on the computer. The configuration of the computer:

Programming environment: Matlab R2012a and VC ++ 2008 OPENCV<sup>15</sup>. The operating system: Win8 64-bit systems. Frequency: 2.5GHz; RAM: 8G Processor: Dual-core Intel Core i7 fourth-generation.

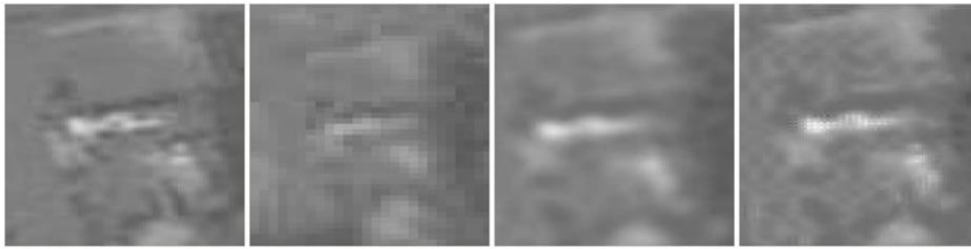
### 5.1 The Improved Super-Resolution Method

Applying the 75-frame ideal high-resolution video (810 pixels  $\times$  612 pixels per frame), 75-frame low-resolution video is generated through the method of the Gaussian blur and down-sampling (405 pixels  $\times$  306 pixels per frame). As is shown in Figure 2, Figure 3, the low-resolution video sequences are treated through super-resolution reconstruction using the bilinear interpolation method, the reference 6 algorithm and the method put forward in this article. And algorithm in the reference 6 are reconstructed using 75-frame image which closes to the target frame image. But, in the algorithm proposed in this article, thinking of the first frame as the target frame, the first frame is recovered making use of other frames. The quality of recovery image enhances obviously, when compared to other algorithms.



(a) Image with High-Resolution (b) Image Using the Bilinear Interpolation Method (c) Image Using Method in Ref 6 (d) Image Using the Improved Method

**Figure 2. Restoration Result Comparison of the First Frame of UAV Video**



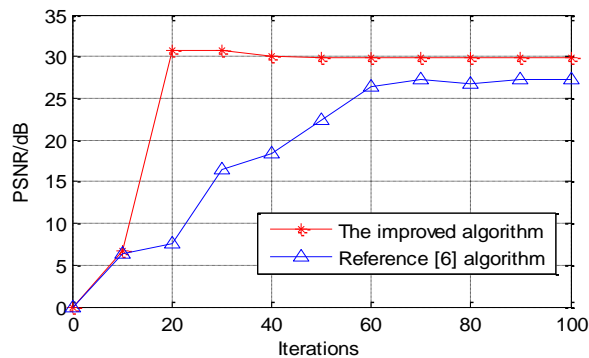
(a) Image Details of High-Resolution (b) Image Details of Using the Bilinear Interpolation Method (c) Image Details Using Method in Ref 6 (d) Image Details Using the Improved Method

**Figure 3. Restoration Result Comparison of the First Frame of UAV Video**

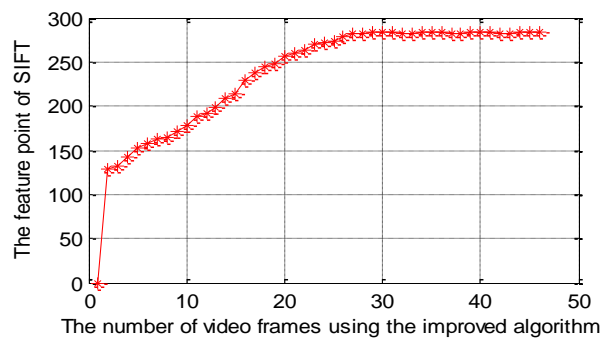
**Table 1. Reconstruction Quality Assessment Data**

	MSE	PSNR	Time/s
Bilinear interpolation	186.07	22.43	<0.05
Reference[6] algorithm	151.03	26.22	611.41
The improved algorithm	63.65	30.13	225.58

The comparison results are shown in Table 1, the reconstruction image is lower in error rate and faster in speed than the original image. (The time distribution is shown in Figure 4 and the number of POCS iterations is 20), Comparing to Figure 3, the results can be achieved: High-resolution images' edge portions have no obvious jagged streaks using this article's method, while retaining the details of the image.



**Figure 4. The Influence Results of the Number of Iterations**



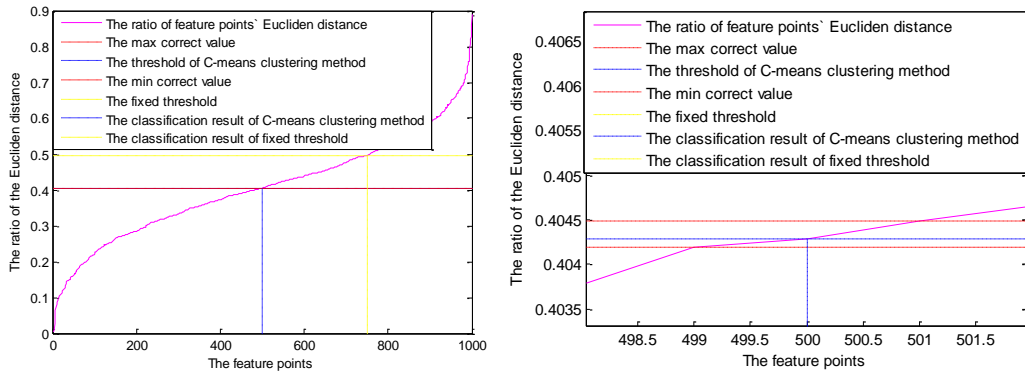
**Figure 5. Relationship between the Number of Video Frames Using the Improved Algorithm and the Number of SIFT Feature Points**

What is shown in Figure 4 is the test simulation of the number of iterations optimize for the results of reconstruction aiming at Coastguard sequence. With the enhancing of the iteration number of the two algorithms, reconstruction effects become better. The number of iteration in this article's algorithm is about 20 to achieve the optimal reconstruction effects, while the number of the algorithm in article 6 is 78. And the optimal effect of this article is 5dB higher than the algorithm in reference 6, which increases the speed of convergence and decreases the computational complexity.

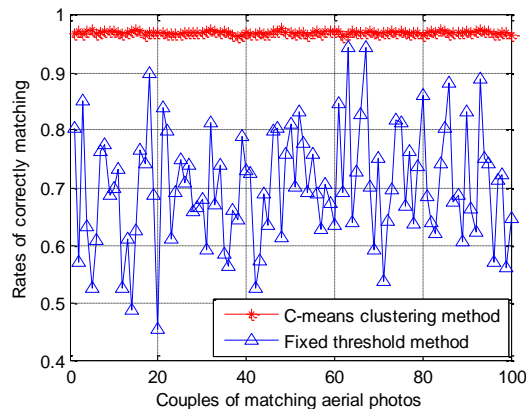
Figure 5 is the test results of the relationship between the frames using the improved super-resolution reconstruction method and the matching points using the method of SIFT. And the conformity resolution doubles, with the addition of the frames number, the matching points of SIFT algorithm gradually increase, whose entirety is doubling. However, when the distance between the target frame and the frame ready to be processed becomes farther, the frame ready to be processed can't achieve the useful data from the target frame. At the same time, the error of image registration stacks, which can't increase the number of the matching points by using of more frames.

### 5.2 The Improved SIFT Algorithm

The simulation images are 100 aerial photos of one air band taken by the UAV (2000 × 3000) and 100 homologous matching aerial photos in Tang County, the city of Shijiazhuang, Hebei Province in August. Selecting randomly 1000 feature points in each aerial photo, the C-means clustering method and the fixed threshold method (threshold set at 0.5) are respectively used to divide the 1000 Continental distance ratio into two categories.  $N_1$  is the number of correct matching point judged by human eye,  $N_2$  is the number of points resulting from this methods.  $N_1/N_2$  is namely the rate of correctly matching. The experimental results are shown in Figure 6.



(a) Results of the Two Classify Methods (b) Details of the Classify Results

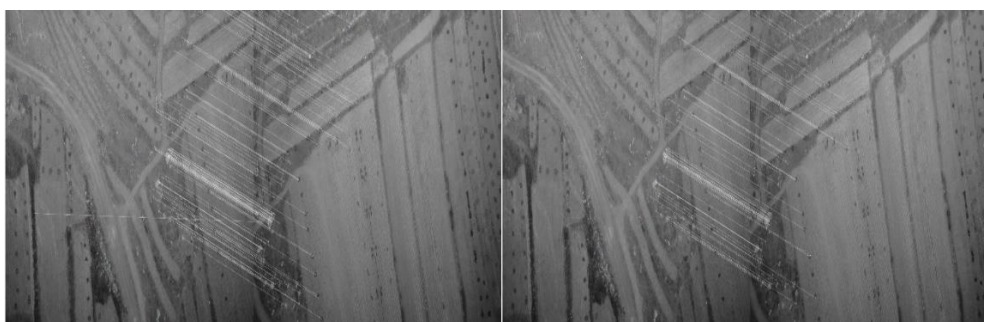




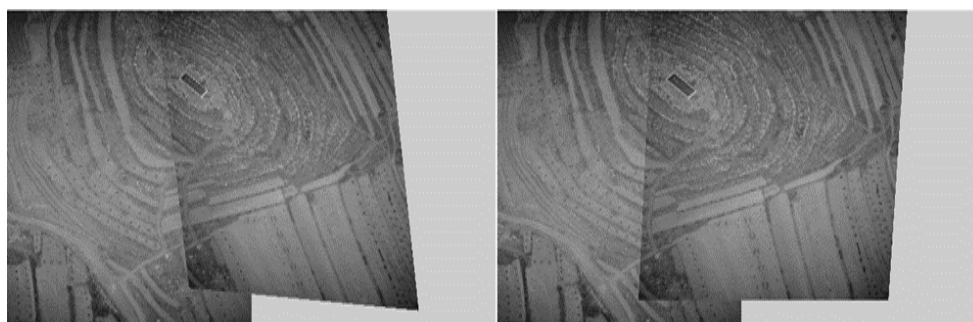
(c) The Comparison of Rates of Correctly Matching of the Two Methods

**Figure 6. The Comparison Image of Classification Effect**

Figure (a) is the classification result of 1000 feature points in one aerial getting through the C- means clustering method and the fixed threshold method, which is intuitively compared. All the points have been arranged from small to large, and the correct range is between the two red lines (Correct maximum value, correct minimum value shown in Figure (b)). The threshold value within the range can get all the correct matching points. Figure (c) is the statistical result of 100 pairs of matching aerial processed by the two methods. The correctly matching rate of the fixed threshold method varies with the changes of images, while the C- means clustering method can adaptively deal with different images to obtain stable and accurate results.



(a) Matching Points in Traditional SIFT Algorithm (b) Matching Points in Improved SIFT Algorithm



(c) The Joints Image in Traditional SIFT Algorithm (d) The Joints Image in Improved SIFT Algorithm

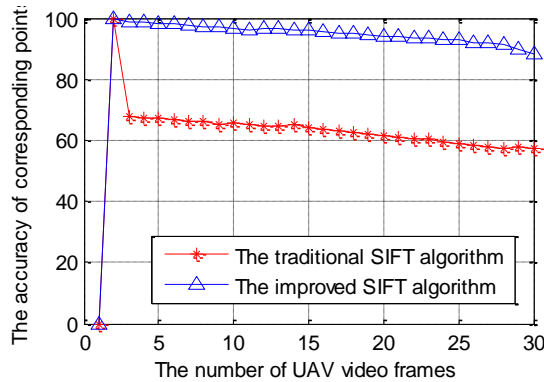
**Figure 7. The Contrast Details of the Results of UAV Stitching**

**Table 2. The Evaluation Table of Stitching Quality**

	The number of corresponding points	The accuracy of corresponding points	Time/s
The traditional SIFT algorithm	448	67.73%	8.236
The improved SIFT algorithm	287	98.67%	8.745

As is shown in Figure 7, the simulation images are two aerials photos taken by UAV in

Tang County, the city of Shijiazhuang, Hebei Province in August ( $2000 \times 3000$ ). The video frames are spliced respectively using the improved SIFT method (regarding the left video frame as the reference) and the traditional SIFT method (Threshold is set at 0.5). The splicing results are shown in Table 2, the number of feature points is identical in the two methods, however, the correct rate of matching points obtained through the improved method is much higher.

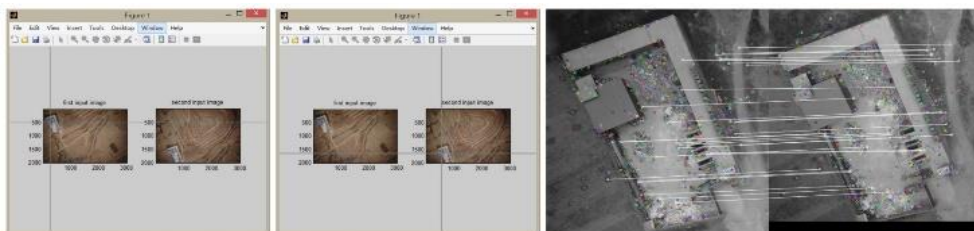


**Figure 8. The Results Affected of Stitching Frames**

What is shown in Figure 8 is the mosaic frame-by-frame test results of 25 video frames of the UAV which considers the first frame as the target, in Tang County, the city of Shijiazhuang, Hebei Province in August ( $405 \times 376$ ). With the increasing of the number of splicing frames, accuracy rate of matching point gradually decreases, which finally tends stabilized. This is because splicing error is gradually accumulated, tending to average splicing errors. Stitching quality of the improved SIFT algorithm is much higher than traditional SIFT algorithm and the stable time is earlier. The entirety performance increases about 30 percent.

### 5.3 The Improved Artificial Choice of Site Registration

As is shown, the Figure 9 is the registration test results with the target of picture on the left of two aerial photos ( $2000 \times 3000$ ) taken by UAV in Tang County, the city of Shijiazhuang, Hebei Province in August. When artificially choosing sites, the general area is only picked out (Figure (a)), the crosshairs cursor is the regional center, and the size of the area is 5% of the total picture. Then using the improved SIFT method to find corresponding points in the selected area (Figure (b)) to improve registration accuracy of artificial choosing the site, and spend less time. Each area includes a number of feature points of the improved SIFT method, Figure (c) and Figure (d) are comparison results between the improved and traditional artificial registration figures. The results show that the improved artificial choice of site registration method significantly enhances the speed and the accuracy of the registration.



(a) The Chosen Region

(b) Matching Points in Improved SIFT Algorithm



(c) Image Using the Improved Artificially Points-Choosing Method (d) Image Using the Traditional Artificially Points-Choosing Method

**Figure 9. Registration Results Improved of Manually Selecting Point**

## 6. Experimental Verification

Preparation stage: Investigating in Tang County, the city of Shijiazhuang, Hebei Province in August by launching the UAV, then taking eight aerial photos of the certain air band and Google digital map of this area. What's more, using of the improved artificial choice of site registration method, the UAV aerial video is registered thinking of digital map as the target to achieve the aerial mosaic. At the same time, the coordinate information of any points are gotten, and the time is 119.37s (Figure 11).



(a) Eight Aerial Photos of the Chosen Region Taken by the UAV (b) Google Digital Map in this Region

**Figure 10. UAV Aerial Imaging and Digital Map**

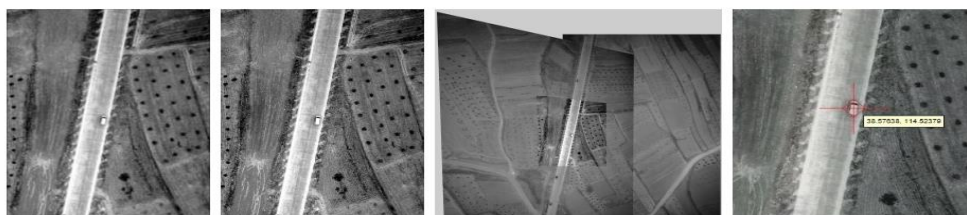


(a) Results of the Joints of the Aerial Photos (b) Results of the Matching of the Aerial Photos and the Digital Map (c) Details of the Matching (d) Obtain the Coordinate Information

**Figure 11. Stitching and Registration Results**

Application stage: As is shown in Figure 12, in order to locate the moving car, the UAV is secondly launched to shoot videos in the same area in Tang County, the city of Shijiazhuang, Hebei Province to obtain the original video frame that contains the target (Figure (a)), and the resolution is  $512 \times 384$ . Then use of aerial that has been ready to shoot in preparation stage, as is shown in Figure (10) a Section 8. Enhancing the resolution of the video by the improved super-resolution reconstruction method (Figure (b)), the resolution is  $1024 \times 768$ , and the whole time is 75.58s. Then considering mosaic picture as reference, the video frame containing the target is registered through the SIFT

algorithm to read latitude and longitude (38.57638, 114.52379) of the car (Figure c), spending 8.74s.



(a) The Original Video Frames (b) The Video Frames Using the Improved Super-Resolution Reconstruction Method (c) Matching of the Video Frames and the Aerial Photos (d) Obtain the Coordinate Information of the Target

### Figure 12. Targeting Results

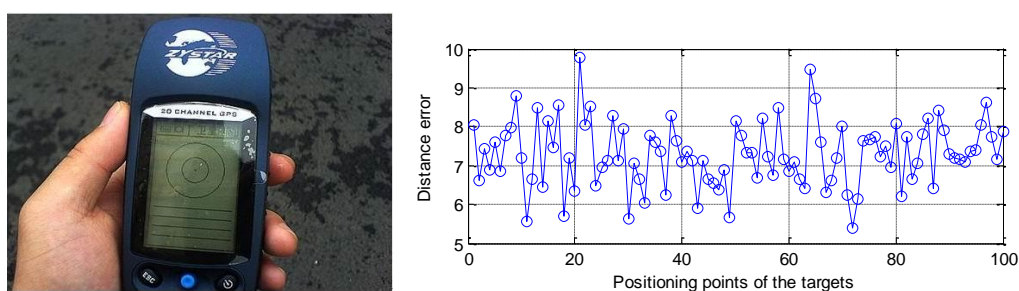
The Final positioning error of algorithm mainly depends on several factors<sup>16</sup>:

(1) The error of the digital map itself. The digital map error depends on digital orthophoto production error. Different regions and different companies make different positioning accuracies. For China, the positioning accuracy of Google satellite map is 0.6 m, the accuracy of Baidu and SOSO is much higher, which can reach 0.2 m.

(2) Manipulator target point measurement error. Generally, Manipulator target point measurement error is 0.5 mm, it is 7 m after transformed into the actual scale according to the photography scale.

(3) The improved SIFT matching algorithm error. General matching error does not exceed 1.33% of the entire image. It is 2 m after transformed into the actual scale according to the photography scale.

Based on the above three factors, the final positioning error of algorithm does not exceed 10 m. Using Tianxiang 500 GPS handhelds (error less than 1 m), the target is positioned on-site to obtain the coordinates information. The latitude and longitude of target is transformed into this coordinates to get deviation distance and solve the deviations for 100 different points, as is shown in Figure 13. The maximum deviation distance is 9.88m, the average is 7.35 m, and the average time of each point is 75.27s. The results show that this method significantly improves the accuracy of the target location and speed.



(a) Tianxiang 500 GPS Handheld Device (b) The Positioning Error

### Figure 13. The Results of Positioning Error

## 7. Conclusion

This paper proposes an UAV target position method based on image matching, utilizing modified super-resolution reconstruction method and SIFT algorithm to accomplish the position of target. The method don't need UAV's location and attitude information, at the same time avoid ground control points' introduction and orthographical correction, and enhance the target position's accuracy.

First, using the modified super-resolution reconstruction method to improve video's resolution with less time, the details of their built image are kept better, and there are no obvious saw tooth stripes at the edge.

Second, the modified SIFT algorithm enhance the matching points' accuracy, let the image montage quality improve stably about 30%.

Third, for the high differences between UAV aerial photo and digital map and the problem of difficult direct registration, the paper utilizes modified artificial source pitching method to make up for the shortage of high resolution of images and then no ability to pitch quickly and accurately in traditional artificial source pitching method, thus to solve the pitch problem of different source images.

Fourth, there are shortages in the algorithm in the paper: there is a limit in improving characteristic points of super-resolution reconstruction method for SIFT, when the video frames are improved to a certain degree, the characteristic point number will not improve with it, when applying the SIFT algorithm, there is error in the pitching for the high buildings in city. These issues will be improved in the following research.

## References

- [1] M. Zhaojun and W. Dehu, "A model of target position for UAV based on attitude measuring laser range-finder", *Fire Control & Command Control*, vol. 28, no. 5, (2003), pp. 334-3379.
- [2] S. Saripalli and J. F. Montgomery, "Vision based autonomous landing of an unmanned aerial vehicle", *Proceeding Of IEEE International Conference On Robotics And Automation*, May (2005).
- [3] D. G. Lowe, "Distinctive image features from scale invariant key-points", *International Journal of Computer Vision*, vol. 60, no. 2, (2004), pp. 91-110.
- [4] Z. Xiangyang and L. Min, "An Automatic and Robust Image Mosaic Algorithm", *Journal of Image and Graphics*, vol. 9, no. 4, (2004), pp. 417-422.
- [5] L. Lingling, L. Cuihua, Z. Xiaoming and L. Bao, "An automatic image registration method based on SIFT and Harris-Affine features matching", *Huazhong Univ. of Sci. & Tech. (Natural Science Edition)*, vol. 36, no. 8, (2008), pp. 13-16.
- [6] W. Suyu, Z. Li and S. Lansun, "A simple & effective video sequence super-resolution algorithm", *Journal of Beijing University of Technology*, vol. 35, no. 6, (2009), pp. 742-747
- [7] Z. Li, W. Suyu and L. Xiaoguang, "Image/Video super resolution. Beijing: Posts & Telecom press, (2011), pp. 100-102.
- [8] A. J. Patti, M. I. Sezan and A. M. Tekalp, "Super resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time", *IEEE Trans IP*, vol. 6, no. 8, (1997), pp. 1064-1076.
- [9] L. W. Tao and L. Gaohuan, "Application of scale invariant feature transform to SAR image matching", *Information And Electronic Engineering*, vol. 8, no. 4, (2010), pp. 388-392.
- [10] L. M. Junjie, Ng M. K. and C. Yuming, "Agglomerative Fuzzy K-Means Clustering Algorithm with Selection of Number of Clusters", *IEEE Trans. Knowledge and Data Engineering*, vol. 20, (2008), pp. 1519-1534.
- [11] L. Heng, L. Yongshu and H. Jing, "An automatic mosaic method in unmanned aerial vehicle images based on feature points", *Geography and Geo-information Science*, vol. 26, no. 5, (2010), pp. 16-19.
- [12] H. Jing, L. Yongshu and L. Heng, "Research of UAV Aerial Image Mosaic Based on SIFT", *Opto-Electronic Engineering*, vol. 38, no. 2, (2011), pp. 122-126.
- [13] Waske B. and Vander L. S., "Classifying multilevel imagery from SAR and optical sensors by decision fusion", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, (2008), pp. 1457-1466.
- [14] S. Xiaoxia and L. Xiaoxia, "The implementation of the FUZZY C-MEANS clustering algorithm", *Computer Applications and Software*, vol. 25, no. 3, (2008), pp. 48-50.
- [15] C. Bradski and A. Kaebler, "Learning Open CV", *Free Online Edition*, (2009).
- [16] W. Zhizhuo, "Photogrammetric Principles", Wuhan: Wuhan University Press, 2007. Legat K. Approximate direct geo-referencing in national coordinates. *ISPRS Journal of Photogrammetry & Remote Sensing (S0924-2716)*, vol. 60, no. 44, (2006), pp. 239-255.

## Authors



**Yan Zhang**, (b. May 16, 1991) He received his Bachelor degree in Surveying and Mapping Engineering (2013) from Jilin University. Now he is studying in UAV information processing in Ordnance Engineering College for Master Degree. He has authored 3 papers.

Yan Zhang, Ordnance Engineering College, Num. 97, Heping West, Road, Shijiazhuang, Hebei Province, China, hillwind@126.com.

**Jianzeng Li**, Ordnance Engineering College, Num. 97, Heping West Road, Shijiazhuang, Hebei Province, China, ljz681@sohu.com.

**Deliang Li**, Ordnance Engineering College, Num. 97, Heping West Road, Shijiazhuang, Hebei Province, China, 271395592@qq.com.

**Zidong Zhou**, Ordnance Engineering College, Num. 97, Heping West Road, Shijiazhuang, Hebei Province, China, zzd michael@163.com.

**Yulong Du**, Ordnance Engineering College, Num. 97, Heping West Road, Shijiazhuang, Hebei Province, China, 1159316475@qq.com.

**Weidong Ji**, College of Computer Science and Information Engineering, Harbin Normal University, Harbin, China, 1209301925@qq.com.