Research Article   **Open Access**

Motion Detection in Satellite Video

Aigong Xu[1]*, Jiaqi Wu[1,2,3], Guo Zhang[2], Shenlin Pan[3], Taoyang Wang[4], Yonghua Jang[4] and Xin Shen[2]

[1]*School of Geomatics, Liaoning Technical University, Fuxin 123000, China*
[2]*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China*
[3]*Zhuhai Orbita Control Engineering Co. Ltd., Zhuhai, 519080, China*
[4]*School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China*

Abstract

In view of the problem of satellite video motion detection, a background subtraction method of combining global motion compensation and local dynamic updating is proposed. In the first instance, the improved ViBE model method is used to establish the background model in the middle frame. The background model has one more dynamic update factor. Secondly, the motion model of global scene between frames is estimated by using uniform blocked forward-back LK optical flow, and the global motion compensation is performed. Last but not least, comparison between compensated frame and model, and connected domain analysis are employed to detect and segment the motion objects. Even more, we can correct the update factor of model according to the "pseudo motion" judgment. And then, the model would be updated locally and adaptively. "Target-wise" evaluation recall rate method is proposed which statistic the object entirety but not pixels. We do four experiments using Skysat and JL1H video. The results show that the proposed method perform a favorable effect on "Target-wise" recall rate and the error detection rate is low. The "Target-wise" recall rate is better than 80%. The error detection rate is reduced by at least 10 times, and even more than 160 times, compared with the classical method. The method could be suitable for advanced application and motion analysis in satellite video

Introduction

At the end of 2013, Sky Imaging (USA) launched SKYSAT-1. It is the first moonlet in the world which could film HD video. The observation of earth went into a brand-new and dynamic mode. In October 2015, Chang Guang Satellite Technology Co., Ltd (China) launched two HR video satellites JILIN No.1 and 2. These are the first ones which are able to film colour videos with high levels techniques including resolution ratio, width and stability. It was also the first-time China manages to prove the video data processing and moving object detection and tracking techniques on-orbit [1]. Different from the single static image shot by the conventional remote sensing satellites, video satellites employ the sensor on board and the maneuver capability of the satellite platform to form videos of the images shot in the targeting areas. Besides the information included in the classical optical remote satellite, it can also acquire the dynamic information directly, realizing the sustainable earth observation extensively within a certain period. The analysis and interpretation of the remote sensing information could be transferred from the static single image to the dynamic sequences, marking a significant step.

In a video clip, people focus more on the state and features of the moving targets. The detection of the moving targets is the core issue in the field of machine vision [2]. Background subtraction method [3] is frequently adopted because of its relatively better accuracy compared with the other motion detection methods. The key of this method is to build the background model by adopting the proper methods for pixel sampling. Meanwhile, an appropriate model updating should be carried out according to the segmentation information of the detection of the previous frame to adapt to the change of the background [4]. The classical methods for motion detection target on the processing of the video data on the ground. The experiments and practices have proved their higher accuracy as well as better effectiveness. They are critical in the field of the application of the video on the ground such as the monitoring for the important areas [5] as well as the statistics of the urban traffic information [6]. However, the detection of the moving targets is faced with the following challenges in the field of satellite

videos: 1) The scene of the satellite video is constantly changing due to the satellite traveling, platform wobbles as well as the adjustment of the sensor attitude. These changes could be categorized into two segments: the global motion and local pseudo motion caused by parallax. Therefore, there are big differences between two adjacent frames with relatively complex relations. The background which is actually static and the moving objects would move or get deformed and detection errors might occur. 2) The frame width of the video is large. And both the resolution as well as the contrast ratio are lower. The size of the moving object is small and the information related to its features and textures are lacking. Some detections might be missed since it is difficult to segment and extract in the videos. As seen in Figure 1 (the image taken from the overlapping display of two adjacent frames), factors such as scene motion and shape changes might result in the big differences between frames. The double images of the moving objects and the background as well as the pseudo color of edges are obvious. Figure 2 shows the frames difference of the adjacent frames. Residual errors could be detected including the residual noises, differences between frames as well as the pre-processing errors, etc.

Therefore, two major aspects should be involved in the moving object detection of the satellite video. One, for the dynamic scenes with large width, the estimation and compensation of the global motion should be applied. In a relative datum space, the background of each frame should be fixed and kept static. So the moving object could be highlighted. Two, an accurate and robust background model should be
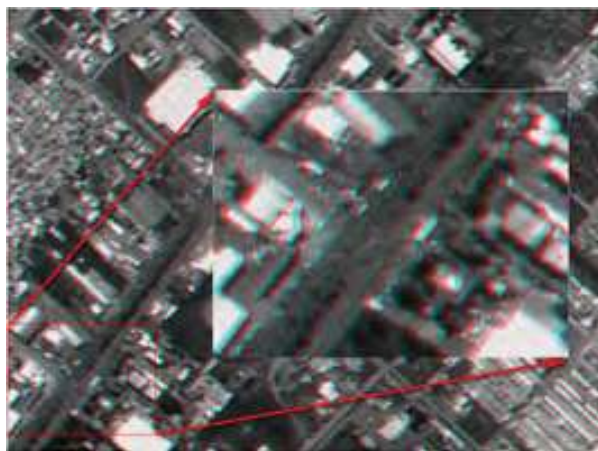
Figure 1: Adjacent frames overlaid display.

Figure 2: Frames difference.

established. And it is also a must to segment and extract the objects with small sizes. The processing of the object with local pseudo motion should also be considered. Meanwhile, proper strategy for background updating should be adopted according to the outcome of segmentation. As for motion processing, some scholars have carried out similar studies. Literature [7] carried out the motion compensation through feature matching and motion filtering. Literature [8] extracted the motion parameters through compressed video motion vector. However, the adoption of this two methods is targeted on the ground or the mobile platforms in the low altitude. In addition, the partial local pseudo motion caused by parallax has not been considered. Guo [9] team from Wuhan University employed the rational function model of the affine correction for processing the image video stabilization. They also considered the geometrical distortion. And the result of processing was good. However, the adoption of this method require us to solve the PRC geometric model. In the literature [10], a three-view geometry was adopted to restrict and block the pseudo motions. But a continuous tracking for the corresponding points in multiple frames is needed. So the requirements for matching precision is fairly high. And due to the adoption of the intersection principle of the three-view drawing, the accuracy is not high if the intersection angle is small. Besides, the model to be solved is too complicated. In the respect of motion detection based on background modeling, many high-quality and classical background

methods [11-13] have been adopted. These classical algorithms are targeting on the detection of the fixed camera on the ground. There is no need to consider the global motion or the local "pseudo motion" of the background. Therefore, appropriate improvements are required according to the features of the satellites video.

Considering the issues mentioned above, this thesis, targeting on the features of the satellite videos and the problems faced by the moving object detection, provided an effective method specialized for the satellite videos. The main approaches ideas are as follows. 1) When the compensation for the global motion occurs, the one unified model for the whole frame could not accurately describe the changes among frames because of the large width of the image and few features which are unevenly distributed. Therefore, a method of motion compensation of uniform blocked forward-backward optical flow is proposed. Through the constraint handling of the local scope, the problems of inaccurate estimation of the models triggered by the small number of features and uneven distribution could be solved. Meanwhile, the tracking of the forward-backward optical flow should be carried out to guarantee the robust matching of the correspondence points. 2) The background model should adapt to the dynamic changes of part of the local scenes. The updating frequency of the local areas bearing the relatively big changes should be increased. Therefore, the updating factor of each pixel in the background will be introduced while building the background model. The updating frequency of the each pixel for background should be dynamic and undertakes adaptive changes according to the detection results. 3) After the compensation for the global motion, the object extraction and detection should be carried out. The candidate target is the made up by two segments: the real moving objects and objects with local "pseudo motion". And the object with the local "pseudo motion" includes the object of parallax "pseudo motion" and "ghost" object. On one hand, the parallax object possesses the following characters: 1) Mainly appear in the edge areas; 2) the size of the object is relatively large and the length-width ratio of the minimum enclosing rectangle is high; 3) The motion direction is aligned with the direction of the translation vector of the inter frame motion model between frames. On the other hand, the "ghost" object is static which is located at the initial position of the object. Hence, we could take advantage of the features of the parallax and "ghost" object to extract the potential unreal moving objects. The scenes of the pseudo motion object areas are bearing big changes in the video and the update frequency of the background model of the corresponding areas should also be increased. Therefore, the adaptive adjustment on the background updating factors of the neighborhood peripatetic of the zone for unreal motions should be adopted to adapt to the local dynamic changes of the scenes. In this way, a more robust and accurate background model could be established to effectively improve the accuracy of the detection.

Description of the proposed algorithm

In terms of motion detection in the satellite video, the basic procedure of the proposed method in this paper is as follows: 1) Establishing the background modeling by using the middle frame; 2) Estimating the motion model between frames, conducting motion compensation based on this model, and getting the result of segmentation via the comparison between compensated frame after resampling and the background model. 3) Extracting potential object by employing the method of connected components analysis, while renewing background model in accordance with detected segmentation information and the candidate "pseudo-motion" region, and processing the next frame continuously in this way. The algorithm flow is shown

as the Figure 3. For detail description of the key technique, please see the subsections below.

The modified VIBE background model

A scene consists of foreground and background in video processing. The detection of moving pixels is actually a procedure to distinguish and categorized each pixel into background or foreground. Therefore an accurate background modeling is significantly critical. In lots of classic algorithm [4,6], using the first N frames for background modeling can build up a precise model via the mathematical statistics analysis of the correlation information and contextual sequences between frames. However, this method fails in extracting dynamic information form the first N frames, especially in satellite video. The duration of satellite video is short, usually less than 2 minutes (even just tens of seconds). So N frames will cause heavy information loss. Besides, the model sampling for the moving regions is a hybrid of both foreground and background, in which some targets with small size might probably be misunderstood as part of the background, causing the decrease of the detection rate. Hence, this thesis adopted the background model theory of VIBE [12] algorithm, completing the model only based on one frame. VIBE employs the pixel-wise random neighbor-sampling of single-frame, to build up the background model. For each pixel position, N pixels would be selected randomly from the pixel and its 8 neighborhood zones as the background samples at the position. The background model at each pixel position is marked as: $M(x) = \{v_1\ v_2...v_n\}$, v here stands for background sampled pixels. This random sampling method promises an equal probability to every pixel in selection, which results in a reliable background model, which is objective and interference-free. On the other hand, because the model just saved the pixel samples from 8 neighborhood zones, the detection of the target with small sizes could be guaranteed and the detection errors caused by the background wobble slightly could also be avoided. But considering the problem of the local "pseudo motion" issue in satellite video, some adjustments for VIBE model should be conducted. In this paper, the updating factor, **a**, has been introduced to this model, representing the renewal frequency of the $M(x)$ in background model. The setting and usage of factor **a** will be illustrated

in following Chapter. As a result, The background model of each pixel position here modified into: $M(x)_m = \{v_1\ v_2...v_n\ a\}$.

Global motion compensation

Motion between frames results in the displacement and deformation of the static background in a video, causing severe interference for the motion detection. The differences caused by the inter-frame motion are the greatest challenge to the motion detection in the satellite video. So the motion compensation for inter-frame is a critical step. Inter-frame motion of the satellite video consists of the global motion as well as the local motion, and this Chapter will focus on the compensation for the global motion. A global motion compensation method of blocked forward-backward LK optical flow here is proposed, and the basic ideas are as follows. Firstly, conduct the uniform blocking processing and extract Good Features in each block. Secondly, track the Good Features with forward-backward LK optical flow [14] to realize the matching of the corresponding point sets. And then, estimate the motion model of each block based on the corresponding point sets. Last, collect and systemize all the transformation results to complete the motion compensation for the whole frame. The details are illustrated as follows:

Image blocking and Good Feature extraction: Broke up the reference frame uniformly into M*N blocks. In this case, the reference frame can be regarded as a new image made up of image blocks in M rows and N columns. Mark each sub-block as B_{mn}, referring to the image block at row m(1<=m<=M) and column n (1<=n<=N). While extracting features in blocks, we use the method—Good Features [15], with a filter window of 3×3. The value of the minimum factor and distance could be set to be slightly smaller, ensuring more enough Good Features could be extracted in the regions with low frequency and weak characteristics, and be equally distributed. Figure 4 has shown a single frame from the video shot by Jilin No. 1. It was divided into 10×10 blocks, and Good Features had been extracted in $B_{10,1}$.

The point tracking and matching of the forward-backward LK optical flow: LK optical flow, is a classical algorithm for tracking the inter-frame points in videos. This method is based on the hypotheses of micro-movement, brightness constancy, and motion similarity in local region. By calculating the offset of inter-frame to estimate the position of this point in the next frame, to realize the feature matching. In order to ensure the robustness and accuracy of tracking and matching of the feature points in domains with low frequency, we use the Good Features matching method with the forward-backward LK optical flow.

Figure 3: Flow chart of proposed method.

Figure 4: Image blocking and GoodFeatures extraction.
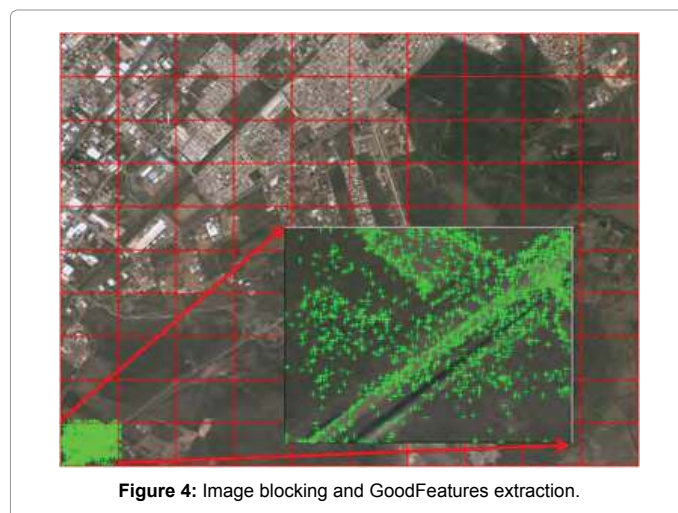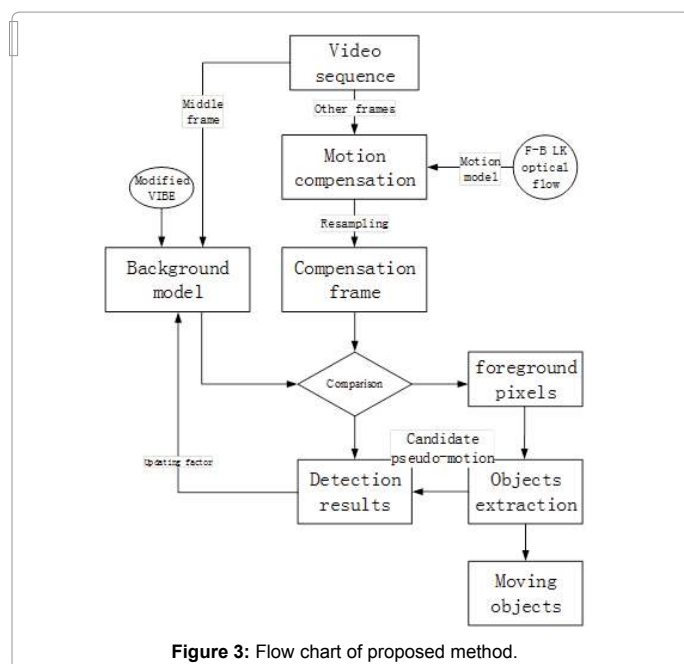
An accurate point tracking algorithm is supposed to hold the alignment between forward and backward direction. The tracking trajectory must be the same no matter it is tracking forward or backward. As shown in Figure 5, the forward tracking was conducted firstly: LK optical flow was conducted to track the X_t point (frame t) to X_{t+1} (frame t+1) and then to the X_{t+k} (frame t+k). Afterwards, the backward tracking should be adopted to track the X_{t+k} to t (frame t) also by LK optical flow. Two trajectory path ways were generated. Then the contrast of the distance between point X_t and $^\wedge X_t$ was conducted. If the distance is less than the fixed threshold, the outcome was reliable. In this paper, k equals 1.

According to the method mentioned above, the Good Features extracted in Block $B_{10,1}$ were tracked and matched. And then an overlapping display of the adjacent frames were conducted and the corresponding point sets were marked. As showed in Figure 6, a large amount corresponding points could be extracted accurately. Meanwhile, an even distribution of the points could also be guaranteed.

The estimation of model and motion compensation: After performing the matching, the corresponding points are used to estimate the inter-frame 2D affine model parameters of $B_{10,1}$ and to establish the geometric transformation relations between frames. In the meantime, the MSACM-estimator SAmple Consensus algorithm [16] can be used to remove the matching points with errors to further optimize the model's parameter, and improve the accuracy of the inter-frame transformation. The 2D inter-frame affine parameters are in the form of a 3 × 3 matrix:

$$\begin{bmatrix} a_1 & a_2 & 0 \\ a_3 & a_4 & 0 \\ a_5 & a_6 & 1 \end{bmatrix}$$

Then, the digital differential rectification method of "reverse solution" [17] is applied. A blank image is established as the compensated frame. The size of the image is the same as the reference frame. According to the transformation relationship, the corresponding sub-image $B_{10,1}$ of the original compensated frame is resampled into the blank frame. Then the new compensated frame is generated, and so as to complete the global motion compensation. Two Adjacent sub-images $B_{10,1}$ overlaid display after the motion compensation is shown in Figure 7, which also include the parameter-optimized high-accuracy Good Features matching point set.

Motion detection and segmentation

After the global motion compensation, the moving targets pixels could be extracted by the contrast between the compensated frame and background model. The method is: for each pixel, compare the similarity between pixel and its corresponding sample pixels in the background
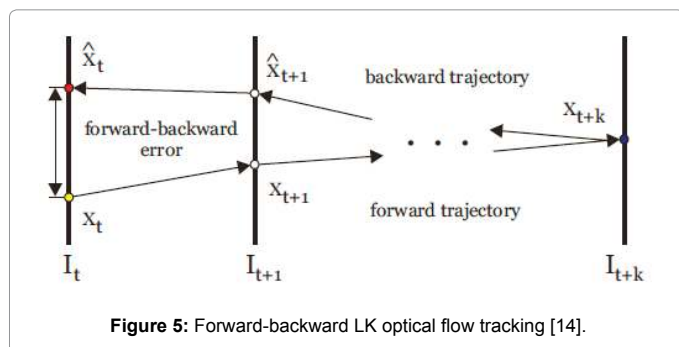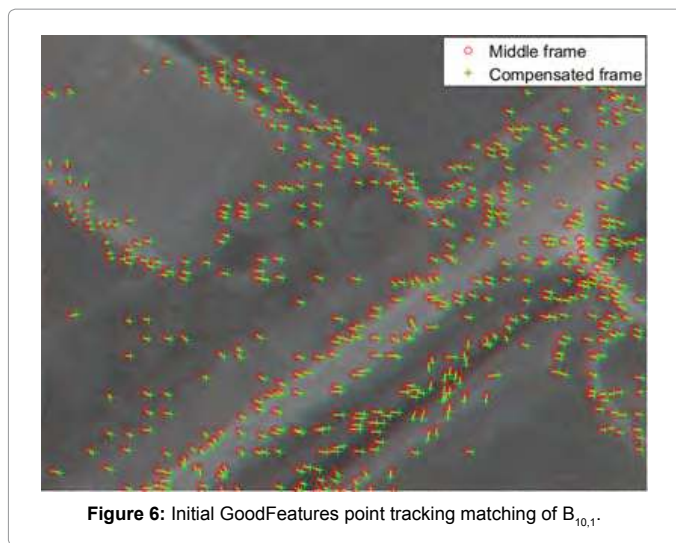

Figure 6: Initial GoodFeatures point tracking matching of $B_{10,1}$.
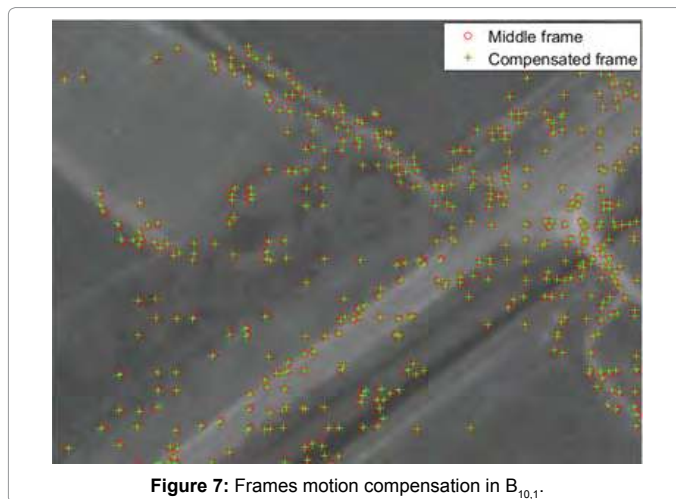

Figure 7: Frames motion compensation in $B_{10,1}$.

model, and if numbers of the designate similarity is larger than a certain threshold, the pixel can be categorized as a background pixel, or else it is a moving pixel. This judging process for the similarity can be explained by the 2D space's Euclidean distance. As shown in Figure 8, V(X) is the pixel to be confirmed, SR(V(X)) is the Euclidean distance space centered with V(X) and having R as the radius. $\{v_1\ v_2...v_n\}$ is the sample pixels of the background model. If the intersection number of SR(V(X)) and $\{v_1\ v_2...v_n\}$ is equal or greater than T, then V(X) could be regarded as the background pixel, the grayscale of which is recorded as 255; or else V(X) is regarded as foreground pixel, the grayscale of which is recorded as 0. It is represented in the following formula as:

$$\begin{cases} Num = S_R(V(\mathrm{x})) \bigcap \{v_1\ v_2...v_n\} \\ Num >= T \end{cases}$$

When judging, Euclidean distance is represented by the inter-pixel gray difference. When all moving pixels classification is completed, a fore- and back-ground binary image is obtained. On this basis, the connected component analysis and extraction could be conducted. Each connected component is regarded as a candidate moving object and the segmentation then was completed.

Figure 5: Forward-backward LK optical flow tracking [14].

The judgment for "pseudo-motion" and background model updating

The candidate objects segmented is made up of three major components: the authentic moving targets, parallax "pseudo-motion" objects, and the "ghost" objects. The judgment of "pseudo motion" is targeting on the judgment of the local parallax "pseudo motion" and the "ghost" objects. Firstly, based on the three characteristics of parallax "pseudo-motion", all the candidates are further processed to separate out potential local "pseudo-motion" objects.

The selection criteria are as follows:

1)	Taking advantage of the coordinate of the nodes of connected component, the minimum enclosing rectangle of the target can be extracted. Calculate the length-width ratio, which, if is more than 3.5, makes this target a potential "pseudo-motion".

2)	Extracting the edge from the local region of the target candidate. To confirm whether the overlapping area between the target and the edge surpasses 90% of the acreage of the target. If so, then the target can be regarded as a potential "pseudo-motion".

Extend the envelop rectangle of potential "pseudo-motion" targets 1 pixel outwards, and in the meantime, extend 10 pixels outwards towards the direction of the translation vector of the motion model to get the updating factor rectangle A. Figure 9 shows the detection results of a frame. The red box represents the minimum bounding rectangle. Based on the "pseudo motion" judgment, this target has a high possibility to be a target of "pseudo-motion. The green dotted box is the extended updating factor rectangle A. The reason for choosing 10 pixels is to take the change of target motion direction into consideration. It is also noted that the orbits of video satellites are usually the sun synchronous orbit, and as a result, the translation vector's direction is usually upwards or downward.

Then, the judgment of the "ghost" target is conducted. While establishing the background model, the moving pixels was adopted as the samples to build the background model at the initial position of the target. When the target left the initial position, the actual background pixel of this position was regarded as the target pixel. Figure 10 shows the detection results of the plane. The dual image of plane appeared at the initial position, namely, "ghost". According to the features of the "ghost" object, the judgment methods were defined as follows: record the position where the moving target appears for the first time. If the moving target keeps still for continuous ten frames, it would be regarded as a potential "ghost" target.
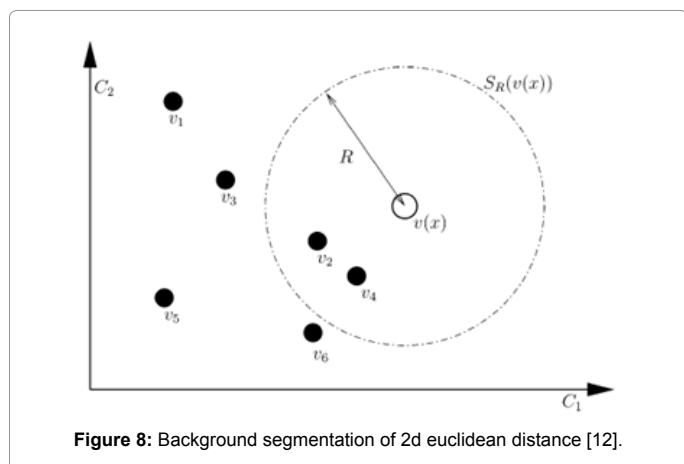

Figure 9: Parallax "pseudo-motion" object.

Figure 10: "Ghost" object.

Figure 8: Background segmentation of 2d euclidean distance [12].
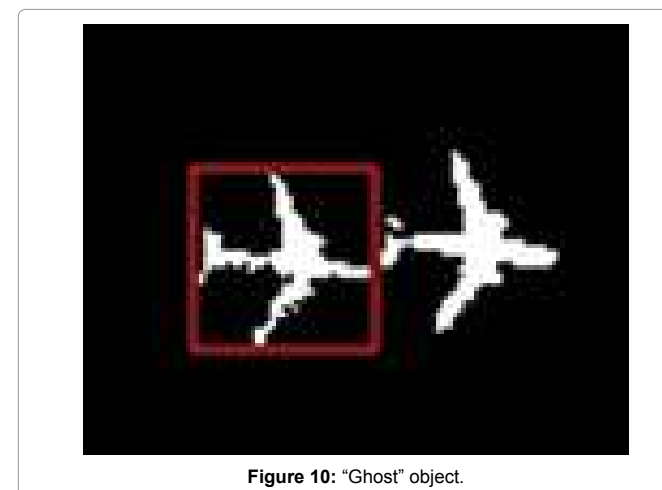
The areas where the moving target and "pseudo-motion" object are located might change as well as its adjacent areas. To adapt to the dynamic changes of the scene and reduce the detection errors, a self-adaption local updating should be conducted following the segmentation of the moving target and the "pseudo motion" judgement. The updating strategy is as follows: every pixel background model possesses an updating factor **a**. For each detected background pixel, we randomly select one number, **s**, from the natural number [1,a]. If s=1, a sample could be selected randomly from its background model $M(x)_m$ and replace the sample with the detected background pixel. Therefore, the updating frequency of each pixel background model is around 1/a. The probability that the sample of each background model is replaced is 1/N. It could be shown in the following formula:

$$\begin{cases} Random(v_1 v_2 ... v_N) = V(x) \ \ if \ \ s=1 \\ s = Random(1,a) \end{cases}$$

Having been processed with the above method, the local "pseudo-motion" regions is updated, while other areas are updated according to the probability. This enables the background model be able to descript the concurrent scenario in a more precise manner to adapt to the local change of "pseudo-motion". It can effectively reduce error detection and increase the preciseness of detection in the premise of

not decreasing the detection rate.

In this experiment, the updating factor is set as follows:

$$\begin{cases} a = 1 & if\ a \in A, Num(Obj) > 20 \\ a = 5 & if\ a \in A, 3 < Num(Obj) <= 20 \\ a = 1 & if\ Num(Obj) <= 3 \\ a = 1 & if\ Obj = Ghost \\ a = 10 & else \end{cases}$$

Here A is the extended updating factor rectangle which mentioned above. Num(*) represents the number of pixels in target. If the number of pixels is less than or equal to 3, this candidate target can be categorized as the background directly, and also the background area of this target is updated; Ghost means "ghost" target, and when the "ghost" target areas are confirmed, they will be refreshed by a=1 regionally. After 20 frames, the update is nearly done, making a=10.

Experiments

Illustration

In order to prove the validity and correctness of the proposed method, four experiments were carried out under the condition of satellite video. Firstly, conduct the global motion compensation, and then, the detection results of two methods, 1) the classical methods without considering the 'pseudo motion' and local updating, and 2) the method proposed in this paper, were compared. Data for experiments were borrowed from the videos shot by Jilin No.1 and Skysat Satellite with a resolution of 1.1m. The size of the targets were small and the contrast against the background was low. In experiment 1, there are a large amount of moving vehicles and certain noises were included in the video. The purpose is to test the detection performance of the algorithm, as well as the performance of blocking the "ghost" and noises. In the second experiment, a car in the video was in a static condition at first. After a while, the car got initiated and swerved with an acceleration. This test mainly was conducted to examine the capability of detecting the object from static condition to the moving condition. It was also employed to block the "ghost". As for the video of the third experiment, no moving target appeared, but there are a large amount of parallax "pseudo motion". The main aim of this experiment is to check the processing effectiveness of the parallax "pseudo motion", namely, the measurement of the precision of the algorism. In the 4th experiment, both the real moving target and the local "pseudo motion" target are included which could check the detection capability and precision of the algorithm comprehensively.

Method of numerical evaluation

The method of numerical evaluation contain two way. The first, "Target-wise" Recall Rate. In the evaluation of detection capability (numerical quantization) towards moving object, the RR (Recall Rate) [18] is a common method. RR represents the ratio of number of real moving pixels in the classified foreground pixels to total moving pixels, a pixel-wise method. Classic detection rate measures the algorithm's performance in detecting moving target pixels. Nevertheless, under the satellite condition, including lower contrast between the object and background, blurred edge of motion targets, synchronized moving objects in small size and pixel quantities, low rate among the total, even if algorithm showing high detection performance, the classic pixel-wise RR method might not read high statistic value and well describe the

performance of the algorithm in inspection. On account of this, "target-wise" Recall Rate evaluation method is put forward in this paper.

With respect to the features of the satellite videos, the whole object should be regarded as the statistical object in the accuracy assessment. Meanwhile, it is also necessary to measure the detect ability of the calculation of the whole object. The calculation formula is:

$$RR_{obj} = TP_{obj} / (TP_{obj} + FN_{obj}) \qquad (1)$$

In this formula, TP_{obj} and FN_{obj} represents the number of the objects which satisfied the following conditions:

$$\begin{cases} TP_{obj} + 1 & if\ Area(GT\ i \cap I(j)) > T * Area(GT\ i) \\ FN_{obj} + 1 & if\ Area(GT\ i \cap I(j)) <= T * Area(GT\ i) \end{cases} \qquad (2)$$

In formula (2), Area (*) refers to the acreage, $\cap$ is the intersection operator, GT(i) represents the i-th object in image GT, Ground-truth, I(j) refers to object j in the image of the detection results. GT(i) and I(j) are matched according to their locations. Ø represents no object; T is the threshold with a value from 0 to 1. After the ergodic of the whole image, we could get the object number of TPobj and FNobj. Then we could employ the formula (1) to calculate RRobj to make accuracy assessment. The value of T in this experiment is 0.4.

Second, The error detection rate was adopted as the measurement of precision of the moving objects detection. The statistical method of the error detection rate is as follows: evaluate the background pixel as the percentage covered by the error classified foreground pixel in total pixels.

Results and Analysis

Experiment 1

The experimental data is a clip of video filmed by JILIN No.1 Video Satellite in an urban area. There are several small-sized vehicles on the road. The contrast between vehicles and the background is low. Figure 11 shows one frame of the video and the results of the detection without motion compensation. It can be seen that the difference among different frames of the satellite video is obvious. The effectiveness of the detection without motion compensation is fairly poor. So the global motion compensation is the significance step in motion detection. For observation, three ROIs was extracted where the moving objects more concentrated. The detection result is shown by Figure 11. The grey scale of the motion pixel is 225 while that of the background pixel is 0. The result of the "target-wise" recall rate are illustrated in Tables 1 and 2.

The result of the observation of the test reveals a good effectiveness of the moving object detection. The "Target-wise" detection accuracy has reached 80% or beyond. The detection rate of the processing results by the classical methods is slightly higher. High similarity between the target and background, low contrast of the grey scale as well as the small size caused the invalidation of the model comparison. This also trigger the missed detections. However, the effectiveness of the classical methods was not good when it comes to the measurement of the precision ratio, especially a large number of detection errors of the "ghosting" pixel. As what we can see from the red square area of the second rank in Figure 12 a wide range of "ghosting" pixel has not been removed for the objects with bigger sizes and higher grey scale contrast. The inaccurate background model is the main reason. A timely background updating was failed to be carried out after the objects were moved. However, the effectiveness of the detection method proposed in this paper is fairly effective. Firstly, the target "ghost" region was figured out. And then update part of the background model in this area. As

Figure 11: A frame and detection result.
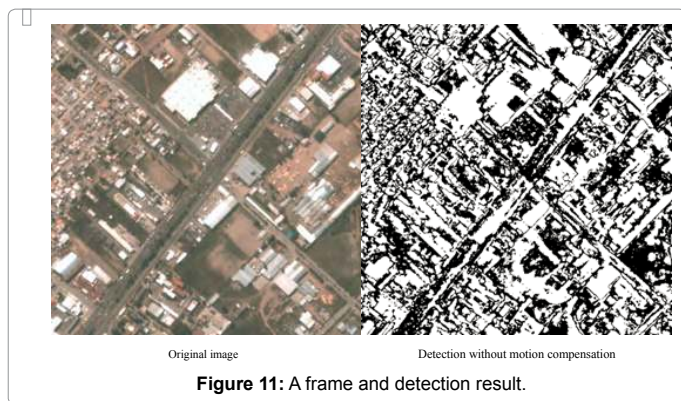
	Classic method			Proposed mothod		
	RRobj	**TPobj**	**FNobj**	**RRobj**	**TPobj**	**FNobj**
ROI1		4	1		4	1
ROI2		8	2		8	2
ROI3		7	0		6	1
total	**0.86**	19	3	**0.81**	18	4

Table 1: "Target-wise" RR of experiment 1.

	Classic method		Proposed method	
	Error detection rate	Errors number	Error detection rate	Errors number
ROI1		56		5
ROI2		41		2
ROI3		30		1
total	**0.00149**	127	**0.00009**	8

Table 2: Error detection rate of experiment 1.

a result, the detection errors of the "ghost" pixel could be effectively decreased. Meanwhile, the analysis of the connected component boost the blocking effectiveness of the noise pixels.

Experiment 2

The experimental 2 data is the video filmed by JINLIN No.1. Figure 13 revealed the processing results. In the 39th frame, the vehicle in the red square area was in a static condition, waiting for turning. This could be regarded as the background. And all the other moving vehicles could be detected. The test outcome shows that the detection error of the "ghosting" existed in the processing results of the classical methods. However, few "ghosting" targets could be observed by using the methods provided. In about 100th frame, the target in the red square was still at the static condition. By upgrading the global background model, the "ghosting" was removed by using classical methods. However, this cost a longer time. In about 207th frame, the objects begin to move. Since it had been regarded as the background when it was in a static condition, the "ghost" occurred in the static situation. In around 234th frame, the object came into the condition of turning and began to speed up. The detection errors of the "ghosting" get increased. And because the methods proposed had removed the "ghost" by performing the pseudo motion processing, the detection errors are obviously lower than the classical methods without processing for pseudo motions. Meanwhile, the effectiveness of the detection of the object silhouettes is good. In about 254th frame, the "ghosting" was still obvious in the classical methods. In the 300th frame, the turning was completed and the detection effectiveness was good. This experiment reveals that the provided methods takes good use of the processing for the pseudo motion to upgrade local background in time. As a result, the detection errors of the "ghosting" could be eliminated rapidly and guarantee the

Figure 12: Detection results of experiment 1.

Figure 13: Detection results of experiment 2.
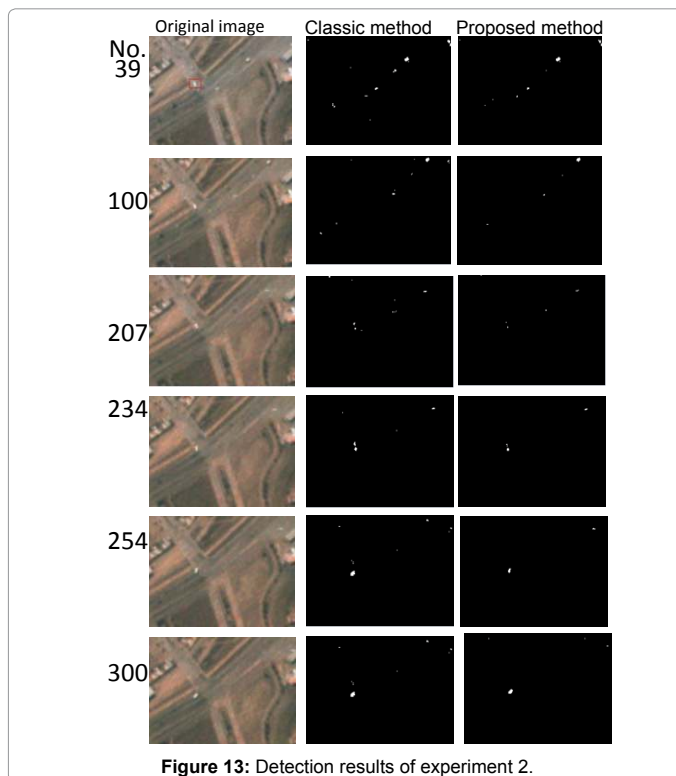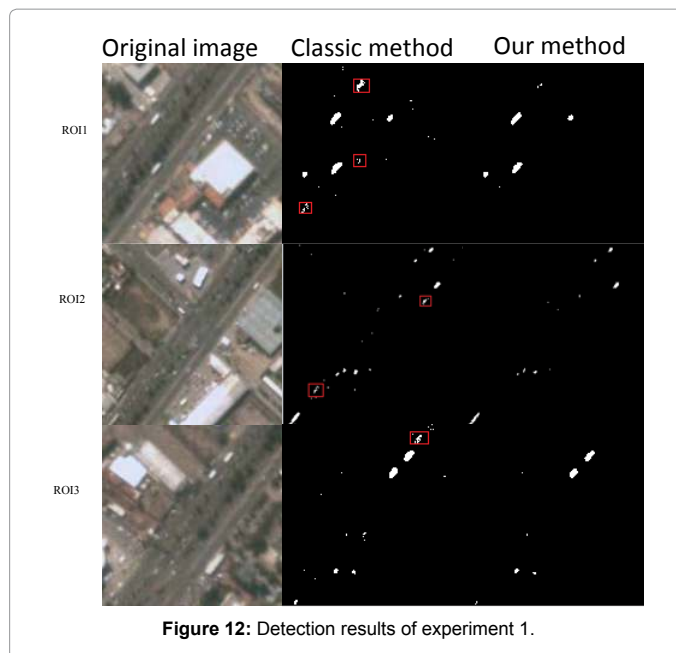
detection rate of the objects. The quality is superior to the classical method.

Experiment 3 and 4

The data of Experiment 3 and 4 is from the surveillance video of an open pit mine area filmed by SKYSAT. A large amount of side slopes consisted by the ladder mineral rock stratum existed in the open pit mine. The elevations of between slopes are greatly differentiated. The parallax were still serious even after the compensation for the global

motion. The local "pseudo motion" could be distinctly observed at the edges of the slope. The outcome of the Experiment 3 was shown in Figure 14. In around 20th frame, the changes of the background was relatively small and the pseudo motions were not obvious. The errors detection were less. From about the 50th frame, the changes of the background became obvious. The global update of the fixed frequency of the model created by the classical methods could not adapt to the "pseudo motion" of the background. And a large amount of detection errors occurred. However, the method provided could block most of the detection errors by processing the "pseudo motion". The accuracy is superior to the classical method. Table 3 illustrates the accuracy.

The outcome of the Experiment 4 is shown in Figure 15. Both the proposed method could block the errors caused by the "pseudo motions" and could also guarantee the detection and segmentation of the actual target.

Conclusion

Satellite video is a kind of new remote sensing data with dynamic and continuous characteristic, and it opened a new era of earth observation pattern and dynamic change information extraction. In this paper, aiming at the specialty and difficulty of motion detection in satellite video, we proposed a detection method of the global scene motion compensation and local dynamic updating. The method proposed could effectively solve the problem of high error detection caused by global scene motion and local pseudo-motion of parallax and ghost in the satellite video motion detection. Experiments demonstrated that our method have good effect on object detection and segmentation, meanwhile, can effectively remove the error detection caused by local

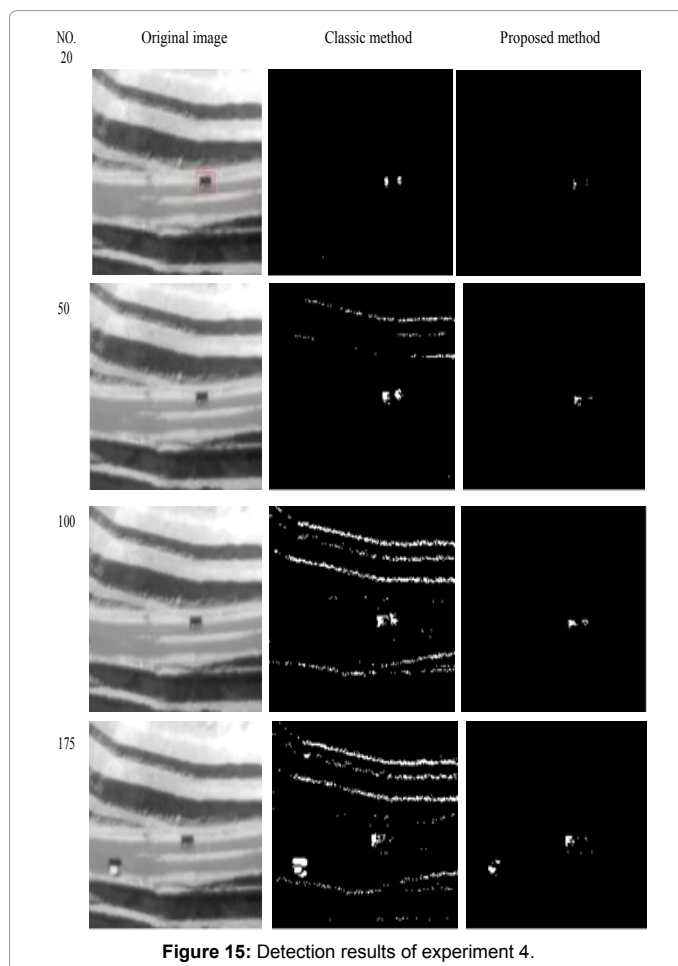| No.	Classic method		Proposed method	
	Error detection rate	Errors number	Error detection rate	Errors number
50	0.03686	4423	0.00055	66
100	0.05591	6709	0.00033	40
175	0.04671	5605	0.0031	372

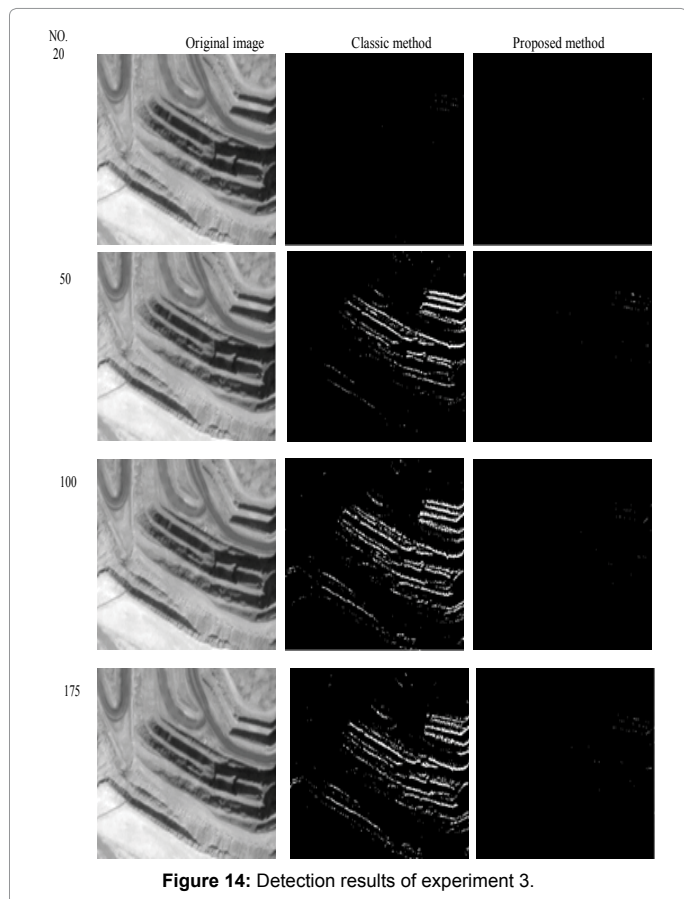**Table 3:** Error detection rate of experiment 3.

Figure 15: Detection results of experiment 4.

pseudo-motion. Comparing with the classic method, our method makes the precision improve significantly. The suggested method is computationally robust and effective, and could meet the needs of motion analysis under the condition of satellite videos. In addition, the proposed "target-wise", an evaluation norm, which is more suitable for the evaluation of small moving object recall rate and simplifies the statistical processing.

Acknowledgements

References

1. Guo Z (2016) Satellite Video Processing and Applications. Journal of Applied Sciences 34: 361-370.

2. Chen XM, Liao J, Li B, Chen QM (2014) Foreground detection based on modified ViBe in dynamic background. Optics and Precision Engineering 22: 2545-2552.

3. Sobral A, Vacavant A (2014) A comprehensive review of background subtraction

Figure 14: Detection results of experiment 3.

algorithms evaluated with synthetic and real videos. Computer Vision and Image Understanding 122: 4-21.

4. Evangelio RH, Patzold M, Keller I, Sikora T (2014) Adaptively splitted GMM with feedback improvement for the task of background subtraction. IEEE Transactions on Information Forensics and Security 9: 863-874.

5. Jeeva S, Sivabalakrishnan M (2015) Survey on Background Modeling and Foreground Detection for Real Time Video Surveillance. Procedia Computer Science 50: 566-571.

6. Kopsiaftis G, Karantzalos K (2015) Vehicle detection and traffic density monitoring from very high resolution satellite video data. Geoscience and Remote Sensing Symposium (IGARSS), IEEE International, pp: 1881-1884.

7. Ji SJ, Zhu M, Hu HP (2013) Aero-borne electronic image stabilization based on feature point matching. Chinese Journal of Optics 6: 841-849.

8. Mu C, Xie J, Yan W (2015) Detecting high-speed moving targets in moving camera environments. Journal of Image and Graphics 20: 349-356.

9. Wang X, Zhang G, Shen X (2016) Satellite Video Stabilization with Geometric Distortion. Acta Geodaetica et Cartographica Sinica 45: 194-198.

10. Xu F, Gu G, Ren K, Qian W (2015) Motion Segmentation by New Three-View Constraint from a Moving Camera. Mathematical Problems in Engineering 2015: 1-14.

11. Ye Y, Mingwei C, Feng Y (2014) EVibe: An improved Vibe algorithm for detecting moving objects. Chinese Journal of Scientific Instrument 35: 924-931.

12. Barnich O, Van Droogenbroeck M (2011) ViBe: A universal background subtraction algorithm for video sequences. IEEE Transactions on Image Processing 20: 1709-1724.

13. Su Y, Li A, Jiang K (2014) Improved Visual Background Extractor Model for Moving Objects Detecting Algorithm. Journal of Computer-Aided Design & Computer Graphics 26: 232-240.

14. Kalal Z, Mikolajczyk K, Matas J (2010) Forward-Backward Error: Automatic Detection of Tracking Failures. International Conference on Pattern Recognition. IEEE Computer Society, pp: 2756-2759.

15. Jianbo S, Carlo T (1994) Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition, pp: 593-600.

16. Lee KY, Chuang YY, Chen BY, Ouhyoung M (2009) Video Stabilization using Robust Feature Trajectories[C]. IEEE International Conference 30: 1397-1404.

17. Jiaqi W, Sun H (2015) Efficient Method of Oblique Image Correction [J]. Remote Sensing Technology and Application 30: 1006-1011.

18. Wang Y, Jodoin PM, Porikli F, Konrad J, Benezeth Y, et al. (2014) CDnet 2014: an expanded change detection benchmark dataset. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp: 387-394.