

Study on dynamic background generation of sequential images and detection of small moving object

Qi Wu, Yue Zhang, Zengping Chen

College of Electronic Science, National University of Defense Technology, Changsha 410073
785694438@qq.com

Abstract—In this paper, the motion estimation and compensation of interframe background, the restoration of complete background image, the detection of moving objects and the optimization of detection results of dynamic sequence images are studied. It is innovatively proposed to use the combination of variance, edge information and histogram that characterizes statistical characteristics as the appropriate characteristic to match blocks, and then use the block matching method to realize motion estimation and compensation of background. This algorithm has proven that it can effectively and quickly realize motion estimation and compensation of background between dynamic sequence images by experiments. Then the traditional background subtraction and three-frame difference method can be used to recover the whole background and detect moving target. However, the problem of target cavity and background residues appears in the target detection results. In this paper, the second threshold method of difference results, the background real-time update method and the subarea multi-threshold method are proposed. The experimental results show that the three schemes can effectively solve the problem of target void and background residue in the detection results. And they are also applicable in the situation where the gray level of the target is similar to the background, which can clearly extract the target submerged in the background, and extend the application scope of traditional background subtraction.

Keywords—*Sequential Image; Motion Estimation and Compensation; Block Matching Algorithm; Target Cavity and Background Residue*

I. INTRODUCTION

Moving target detection of sequential images is an important research topic of video monitoring and image processing. Its purpose is to determine whether there are moving targets in sequence images and extract them quickly and effectively from the background of dynamic changes, which is the premise and key to further analysis and processing of images. Moving target detection under static background has been widely used in military and civilian fields. At present, there are a large number of mature algorithms at home and abroad, mainly including: background model method [1], motion field method [2] and frame difference method [3]. However, the dynamic background greatly increases the difficulty of target detection, which makes most of the above algorithms impossible. For this problem, there is no mature and widely recognized algorithm for direct target detection of dynamic background. Since most of the pixels in the images

will change when the background moves, the detection of the target by the above algorithm will cause confusion between the target and the background. Most solutions are to use the global estimation and compensation method of background [4]. The motion vector of representative pixels is obtained by matching the feature quantity, and the global motion parameters of the background are obtained to estimate and compensate the motion information of the background. The problem of target detection under dynamic background is transformed into the problem of target detection under static background. In this paper, block matching algorithm is used to realize motion estimation and compensation of sequence images, which is transformed into static background, and then background reconstruction and moving target detection are carried out. Finally, the algorithm is optimized for the problem of target cavity and background residue, which makes the detection results of moving target more clear and complete.

II. MOTION ESTIMATION BASED ON BLOCK MATCHING

The basic idea of block matching method is to choose a certain size of reference block with rich gray level information and obvious matching characteristics. Then determine the position and size of an appropriate window in the current frame (next frame adjacent to the reference frame), and use the appropriate matching criteria to search the best matching block that is most similar to the reference block in the window. Then the motion information of the background in the current frame relative to the reference frame can be estimated by the position coordinate difference between the reference block and the best matching block (that is the offset direction and distance). This section divides the block matching algorithm into five steps:

- Determine the appropriate reference block (template) to be matched in the reference frame;
- Determine the search window in the next frame adjacent to the reference frame (current frame);
- Select the appropriate search strategy;
- Select the best matching criteria;
- Determine the best matching block and calculate the background offset by coordinate difference.

Block matching algorithm schematic diagram is shown in fig.1 below:

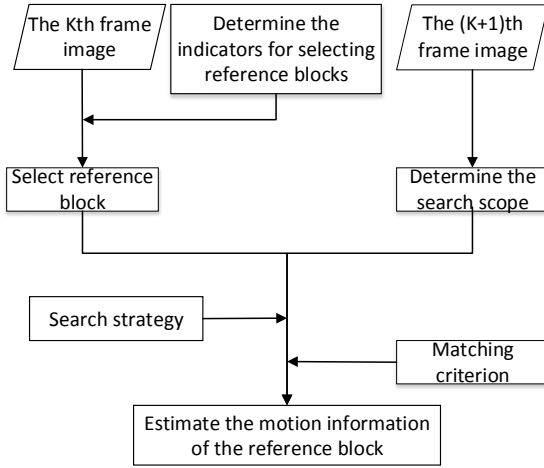


Fig. 1. Block matching algorithm schematic diagram

A. Selection of Reference Blocks

Because of the diversity of images, the complexity of image content and multifarious background, to match the images' translation moving of two frames accurately and quickly, reference blocks with abundant gray information and obvious matching characteristics should be selected as far as possible. It is innovatively proposed to use the combination of variance, edge information and histogram that characterizes statistical features as characteristic to choose the appropriate matching blocks, and select the square for reference block. Assume the size of the block is $m \times n$, $f(i, j)$ representing the gray value at the pixel point (i, j) , and the mean and variance of the block are defined respectively as:

$$\mu = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n f(i, j) \quad (1)$$

$$\sigma^2 = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n [f(i, j) - \mu]^2 \quad (2)$$

The histogram of the image shows the coverage of gray value and the proportion of each gray value pixel in the whole image. The mean and variance of the gray histogram are defined respectively as:

$$Mean = \sum_{k=0}^{255} k * p_k \quad (3)$$

$$\sigma^2 = \sum_{k=0}^{255} (k - Mean)^2 p_k \quad (4)$$

In the formula, k represents the gray value of the pixel, p_k represents the probability when the gray value of the pixel is k .

According to the threshold segmentation theory of histogram, for an image, if the histogram has obvious double peaks and the variance of each peak is large enough, the possibility of edge information included in the image is great. If the histogram of the image has only single peak and its variance is very small, the image does not have edge or edge does not protrude. The above analysis shows that, for an image, if the histogram of a region in the image contains double peaks and the greater the variance, the greater the dispersion degree

of the gray value of each pixel in the region will be, and it is likely to contain edge information, which is suitable for the reference block to be matched. Therefore, the reference block to be selected should be a region containing double peaks and large variance of gray histogram in an image.

B. Search Strategy and Matching Criteria

Because the background and target in the image have the continuity between frames, the search scope in the current frame only needs to be near the original position of the reference block $s(x, y)$. With $s(x, y)$ as the center, it expands d_x and d_y in horizontal and vertical directions respectively, to get a search window, whose size is $(2d_x + 1) \times (2d_y + 1)$. The full search method is adopted in the search window. Five matching criteria commonly used are shown as follows:

Definition 1. The normalized correlation function NCFF is expressed as:

$$NCFF(i, j) = \frac{\sum_{m=1}^M \sum_{n=1}^N f_k(m, n) f_{k-1}(m+i, n+j)}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N f_k^2(m, n)} \sqrt{\sum_{m=1}^M \sum_{n=1}^N f_{k-1}^2(m+i, n+j)}} \quad (5)$$

Definition 2. Sum function of absolute differences SAD is expressed as:

$$SAD(i, j) = \sum_{m=1}^M \sum_{n=1}^N |f_k(m, n) - f_{k-1}(m+i, n+j)| \quad (6)$$

Definition 3. Mutual Information MI, Normalization Mutual Information NMI, Entropy Correlation Coefficient, ECC

In information theory, Mutual Information (MI) describes the correlation between two systems, or the amount of information contained in each other [5]. The mutual information MI is used to represent the degree of mutual inclusion of R and F of two images, namely the degree of similarity. The mathematical expression is as follows:

$$MI(R, F) = H(R) + H(F) - H(R, F) \quad (7)$$

The higher the similarity between the two images is, the higher the correlation is and the greater the mutual information is. Therefore, the candidate block corresponding to the maximum of MI is the best matching block. However, when the reference block has similar gray scale distribution with multiple candidate matching blocks, the phenomenon of false matching is easy to occur when matching according to mutual information. To solve this problem, Studholme and Maes [5] improved MI and proposed the normalized Mutual Information (NMI) and the Entropy Correlation Coefficient (ECC) respectively.

Definition 4. Normalization Mutual Information, NMI is expressed as

$$NMI(R, F) = \frac{H(R) + H(F)}{H(R, F)} \quad (8)$$

Definition 5. Entropy Correlation Coefficient, ECC is expressed as:

$$ECC(R, F) = \frac{2MI(R, F)}{H(R) + H(F)} \quad (9)$$

Simulation experiments were conducted to analyze and compare the five matching criteria of SAD, NCFF, MI, ECC and NMI. The results are as follows:

TABLE I. MATCHING RESULTS OF SAD、NCFF、MI、NMI、ECC

Matching criterion	Lena images		The Measured Data	
	Matching Time(s)	Matching Results	Matching Time(s)	Matching Results
SAD	3.5982	failed	—	—
NCFF	31.4171	succeed	29.1113	failed
MI	313.7451	succeed	424.4288	failed
NMI	310.8318	succeed	410.1703	succeed
ECC	323.7715	succeed	420.5756	succeed

Simulation experiments on five matching criteria were carried out to compare their matching time and matching effects. It can be seen from TABLE.I that NMI matching criterion is superior to the other four in both the matching effects and time. It is determined that NMI is the most suitable matching criterion for block matching under the complex natural background studied in this paper.

C. Experimental Results and Analysis

In the reference frame, the area with dual peaks and large variance of gray histogram is selected as the reference block to be matched. In the search window of the current frame, the global search method and NMI criterion are used to match the block, and the best matching block is obtained, whose top left pixel coordinate is $s' = (x', y')$, The coordinate difference mv between the best matching block and the reference block is the offset of the two frames of image, which is defined as:

$$mv = s' - s = (x', y') - (x, y) = (\Delta x, \Delta y) \quad (10)$$

Taking 9 frames sequence images collected in the natural environment as the experimental objects, their size is 656×492 in pixels (656 is the width of image and 492 is the height of image), and taking the variance of the gray histogram with double peaks as the index of selecting the reference block, the size of the reference block is 64×64 . The search distance of horizontal and vertical extension d_x, d_y are 32 (that is, a half of the height and width of the reference block, experimental results show that the search distance is reasonable), then the search window is $(2d_x + 1) \times (2d_y + 1) = 65 \times 65$, matching criteria is NMI. Experiment results are shown in table as:

TABLE II. THE OFFSET DISTANCE AND DIRECTION OF EACH FRAME IMAGE RELATIVE TO THE PREVIOUS FRAME

Number	Reference block coordinates	Best match block coordinates	The offset	Background motion direction
1	(492,93)	—	—	—
2	(482,93)	(482,93)	(-10, +0)	left
3	(464,95)	(464,95)	(-18, +2)	Lower left
4	(462,95)	(462,95)	(-2, +0)	Left
5	(458,95)	(458,95)	(-4, +0)	Left
6	(458,94)	(458,94)	(+0, -1)	Up
7	(456,93)	(456,93)	(-2, -1)	Upper left
8	(453,94)	(453,94)	(-3, +1)	Lower left
9	—	(446,92)	(-7, -2)	Upper left

The migration direction and distance of each frame image relative to the previous frame are obtained from the TABLE II, that means, the motion information of the background in the dynamic sequence images, and the motion estimation of the dynamic background are realized. The common part of each frame image is taken out to transform the dynamic background into the static background, and then the restoration and reconstruction of the whole background image in the static stable background and the detection of moving target can be realized by using the three-frame difference method.

Experiments results are shown as follow:



Fig. 2. Experiment Results(a)(b)(c)(d)(e)(f)

In the figure, (a), (b) and (c) are dynamic sequence images with three frames; (d) are complete background images obtained according to the algorithm in this paper; (e) and (f) are target detection results obtained with a threshold of 50 and 20 respectively. It can be seen that the influence of the threshold on the target detection is large, which leads to the incomplete detection of the target and the inner cavity. A small threshold results in a large number of background pixels left. For this, the target detection algorithm needs to be optimized.

III. SOLUTIONS OF TARGET CAVITY AND BACKGROUND RESIDUE

According to the above analysis, it can be concluded that the reason of target cavity and background residue is related to threshold selection. Three solutions are proposed to solve the problem of target cavity and background residues in the detection results: the second threshold method of difference results; background real-time update method and the subarea multi-threshold method. Three solutions were simulated and

their advantages, disadvantages and applicable conditions were summarized.

A. The Second Threshold Method of Difference Results

Firstly, low threshold value is selected to realize background subtraction, and threshold $T=20$ is selected to ensure the integrity of target detection. Then the difference result graph is divided into blocks to calculate the brightness sum of all pixel points of each sub-block. Based on the fact that distribution of the target pixels is relatively concentrated and the distribution of remaining background pixels is random, independent and evacuation, sub-blocks are segmented secondly according to the pixel brightness and threshold. The size of block is very important, sub-blocks are so large that they contain multiple targets and background pixel, making it difficult to distinguish between background and target; sub-blocks are too small, leading to increase the computational amount while it is not easy to distinguish between target and background.

B. Background Real-time Update Method

Background real-time update method is to update the background frame in real time before the input frame is subtracted to detect the moving target, and to establish a new background frame corresponding to the input frame. In the complex natural environment, there are always various interferences, such as the sway of leaves in the background, ripples on the water surface, and the slight jitter of the camera, which cause the slight rotation of scenes in the collected image, all of which will affect the detection effect of the target. Especially for dynamic background images with periodic motion, it is necessary to update the background frame in real time before implementing background subtraction.

C. The Subarea Multi-threshold Method

The subarea multi-threshold method is to partition for the background frame and input frame, and to set different local thresholds for different regions. Low threshold value is selected for the target part to ensure the integrity of the target detection. High threshold value is selected for the background part to filter out most of the background, so as to resolve the contradiction of threshold selection.

D. Experimental Results and Analysis

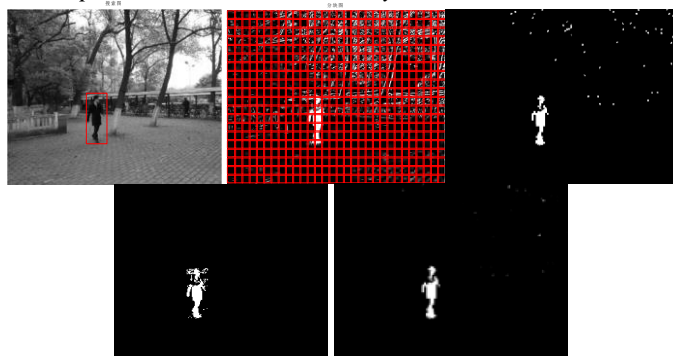


Fig. 3. Experiment Results(a)(b)(c)(d)(e)

Where, (a) is the experimental original graph, (b) is the block diagram of the initially obtained difference binarization image, and (c)(d)(e) are respectively the detection results obtained after the background subtraction optimized by the second threshold method of difference results, background real-time update method and subarea multi-threshold method. It can be seen from the experimental results that the advantage of the second-threshold method of difference results is that the complete and detailed target contour can be obtained, and the background filtering near the target is relatively clean, but there are still a few discrete background pixel residues for the periodically changing background. The scheme is suitable for occasions where the target is large and internally connected. The advantage of the background real-time update method is that it can remove most of the background cleanly, but there will still be a few residues near the target, and the background update needs some time. The scheme is suitable for occasions with periodic variation of background or external interference. The subarea multi-threshold method is applicable in most cases. Here are three options to compare:

TABLE III. A COMPARISON TABLE OF THREE SOLUTIONS

Solutions	The second threshold method of difference results	Background real-time update method	Subarea multi-threshold method
Basic process	Two threshold segmentation	Background update and threshold segmentation	Subarea and multi-threshold segmentation
Object	Difference binarization result map	The input frame and the updated background frame	Input frame and the original background frame
Threshold	Global threshold	Global threshold	Local threshold
Suitable conditions	Occasions where the target is large and internally connected	Occasions with periodic variation of background or external interference	Most occasions

All three solutions can optimize the detection effect of traditional background subtraction to a certain extent, solve the problem of target cavity and background residue, and make the detection target clearer and more obvious.

E. Situation With Similar Gray Level of Target and Background

When the gray level of the target is close to that of the background, the target is often submerged in the background. At this point, the target cannot be detected only by the frame difference method and the traditional background subtraction method. The three solutions proposed in this paper are equally applicable to this situation. The specific results and analysis of experiments are shown as follows:

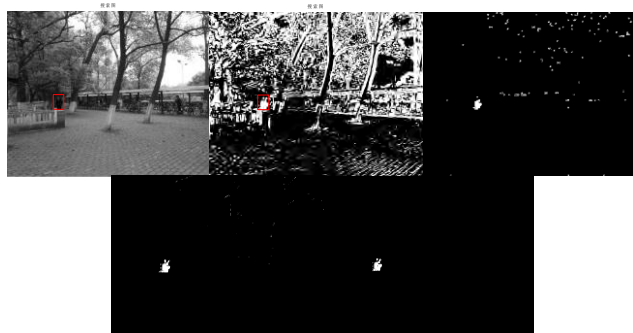


Fig. 4. Experiment Results(a)(b)(c)(d)(e)

Where, (a) is the original image where the gray level of the target is close to that of the background, and the target is mostly submerged in the background. The detection results are obtained by the traditional background subtraction method, as shown in figure (b). It can be seen from the results that for the image whose gray level are similar to the background, the traditional frame difference method or background subtraction method can't detect the target directly. (c)(d)(e) are the detection results obtained from three optimization algorithms proposed in this paper, it can be seen that the optimization algorithm can extract the targets submerged in the background effectively and clearly, and the detection results are much better.

The experiment verifies the effectiveness and robustness of the optimization algorithms. They can overcome the defect that the traditional background subtraction is difficult to adapt to various disturbances and periodic changes of the background in the complex natural background, and they apply to the situation that the gray scale of the target is close to the background, and extend the application scope of the traditional background subtraction.

IV. CONCLUSIONS

Motion target detection in video image is the key technology of follow-up target tracking and behavior analysis, which has important research value and application prospect. In this paper, motion estimation of dynamic background, restoration and reconstruction of complete background and detection of small moving objects are studied. Based on the traditional block matching, this paper proposes a new block matching algorithm based on the key selection of reference blocks, innovatively proposes the idea of combining the variance with easily recognized edge information and histogram that characterizes statistical characteristics, and increases the strong features of intra-block identification, so as to identify the block position change more accurately and quickly. Experimental results show that the algorithm can overcome all kinds of noise interference under natural conditions and improve the matching accuracy, so as to achieve the estimation of background motion information in dynamic background sequence images.

After transforming the dynamic background into the static background, three frame difference and background subtraction are used to realize the reconstruction of the whole background frame and the detection of moving target. For the problem of target cavity and background residue in the detection results, it

is concluded that the main reason is the contradiction of threshold selection, and three effective solutions are proposed to optimize traditional background subtraction and improve the detection results of moving targets. The experimental results show that the three solutions can effectively solve the problems of target cavity and background residue in the detection results and make the detected target more clear and complete. In addition, these three solutions are also applicable to the situation where the gray level of the target is similar to the background, which can extract the targets submerged in the background clearly, overcome the defects of traditional background subtraction that can't adapt to various disturbances and periodic changes of the background, and extend the application scope of traditional background subtraction.

REFERENCES

- [1] Han Hongzhe, Wang Zhiliang. Real-time human detection based on adaptive background model [J]. Journal of Beijing university of science and technology, 2003,25 (4): 384-386.
- [2] Sun Hui, Zhao Hongying. Image motion estimation method based on optical flow model [J]. Optical precision engineering, 2002,10 (5): 443-447.
- [3] Wang Hongjian, Li Zhimin. Vehicle flow real-time monitoring system based on video image [J]. Optical precision engineering, 2005,13 (add): 222-226.
- [4] Huang Shenzi. Research on detection and tracking technology of moving objects under moving background. [D] ZheJiang University: image processing, 2011,1.
- [5] Image Matching. Image registration algorithm based on mutual information: MI, NMI, ECC algorithm -HuJingshuang-blogchannel-CSDN.NET[N].<http://blog.csdn.net/hujingshuang/article/details/47910949>.