International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.4 (2016), pp. 187-196
http://dx.doi.org/10.14257/ijmue.2016.11.4.19

Background Motion Video Tracking of the Memory Watershed Disc Gradient Expansion Template

Yao Nan[1] and Shen Haiping[2]

[1]*Department of Jiangsu Electric Power Company Research Institute, NanJing Jiangsu, 211107, China*
[2]*Department of Wuxi Power Supply Company, Wuxi Jiangsu 214000, China*
yaon_js@aliyun.com

Abstract

In order to further improve motion video target tracking and detection accuracy, a motion video tracking algorithm based on memory watershed disc gradient expansion template is proposed in this article. Firstly, heterodromous diffusion is executed to preprocess the motion video image in order to reduce the noise interference in the motion target tracking and detection process, and meanwhile the differential operation and the morphological operation are executed to extract and operate the outer profile of the motion object. Secondly, in allusion to the excessive division in the watershed algorithm, target marking and division of the watershed algorithm are realized on the basis of characteristic memory, and the accurate detection of the motion target is realized on the basis of the disc gradient expansion template. Finally, according to the experimental comparison, the algorithm proposed in this article can be applied in the complex background detection of the motion target, and such algorithm can not only improve the motion target detection accuracy, but also significantly improve computation speed.

***Keywords**: Video tracking; Memory; Watershed; Disc gradient; Expansion*

1. Introduction

Along with the gradual maturity of the motion object research technology in the video monitoring field, the corresponding algorithm is widely applied in the fields related to national economy and people's livelihood, such as bank monitoring and community monitoring, but the motion object monitoring is the key and difficult point of the research in this field. Generally, in the complex environment with noise interference, whether the motion object can be integrally and really detected and represented is an important index for verifying the target tracking algorithm [1].

The motion target detection mainly includes motion target tracking, target classification, target behavior annotation, *etc.*, and functions as the basic and key part of video sequence processing [2]. At present, the motion detection research is mainly focused on the following three aspects:

Firstly, motion image light stream analysis method: as mentioned in literature [3], this detection mode mainly has the following problem that the algorithm realization is too complex to limit the practical application value, and literature [4] also points out that this algorithm has poor anti-noise ability; secondly, motion detection algorithm based on the consecutive frame gradient detection of an image: as mentioned in literature [5], this detection method mainly has the problem that the algorithm has poor target integrity detection ability and the information details can be easily lost [6]; thirdly, motion image background target division detection algorithm: as mentioned in literature [7], the basic thought of this algorithm is to distinguish the background as preposition target and postposition background, and due to the simple and convenient calculation, this algorithm is widely used in

ISSN: 1975-0080 IJMUE
Copyright ⓒ 2016 SERSC

present research for motion target detection and is most applicable to real-time monitoring application [8]. In practical use, due to the uncertainty interference of illumination variation and irrelevant interference events, the standard background division has poor effect and stability [9].

In allusion to the motion target detection problems in the above algorithms, scholars mainly focus on researching the watershed target detection method and the application thereof in recent years and also obtain certain achievements. Such watershed method has the features of simple design, convenient realization and relatively stable target detection [10]. Due to the existence of gradation disturbance noise in practical use, the watershed target detection algorithm can easily cause the "excessive division" problem in video image detection, and the detection accuracy shall be further improved [11].

In order to improve the deficiency of the watershed method and realize more ideal motion video sequence detection effect as well as solve the "excessive division" problem, a motion video tracking algorithm based on memory watershed disc gradient expansion template (DGETMW) is designed and realized in this article, wherein this algorithm can effectively combine morphology and watershed algorithm characteristics in order to effectively track, identify and detect the motion video sequence. Meanwhile, the simulation test is carried out to verify the algorithm effect, and the test result shows that DGETMW algorithm has fast motion target detection speed and can meet the practical application requirement, namely real-time detection.

2. Motion Video Sequence Detection

2.1. Algorithm Procedure

The basic procedure of the motion video tracking algorithm based on memory watershed disc gradient expansion template is as follows: firstly, the motion video sequence is preprocessed on the basis of the disc gradient expansion template in order to reduce the interference of uncertain illumination and event on the motion target; secondly, the profile of the motion target sequence is extracted on the basis of the consecutive frame gradient detection, and meanwhile the extracted profile is morphologically processed, and after the above preprocessing algorithms are executed, the memory watershed algorithm is adopted for target sequence memory marking and target detection. The algorithm procedure is as shown in Figure 1.

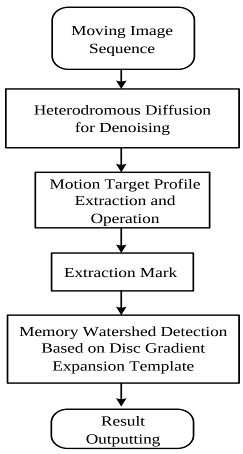


Figure 1. Procedure of Motion Video Tracking Algorithm Based on Disc Gradient Expansion Template

 Copyright ⓒ 2016 SERSC

2.2. Adoption of Heterodromous Diffusion for Motion Sequence Denoising

During the monition image sequence data acquisition and execution process, the influence of such uncertainty interference as uncertain illumination can cause various noises, and the effect of extracting noisy image sequence is certainly not good [12]. Accordingly, the heterodromous diffusion is adopted for the image to replace the noisy pixels, diffuse and ablate noise points and avoid the adverse influence of noise points in order to improve motion sequence detection accuracy. Therein, the mathematical principle of the heterodromous diffusion mentioned above can be described as follows:

$$\frac{\partial u(x,y,t)}{\partial t} = div[\, g(||\nabla u(x,y,t)||) \nabla u(x,y,t)] \tag{1}$$

In the formulae (1) and (2), $\nabla u(x,y,t)$ refers to the gradient amplitude of the consecutive frame of the pixel; $g(||\nabla u(x,y,t)||)$ is a nonincremental diffusion function and can meet the following conditions:

$$u(x,y,0) = u_0(x,y) \tag{2}$$

$$g_k(\nabla u) = exp(-(||\nabla u|| / K)^2) \tag{3}$$

In formula (3), K is the diffusion blockage factor of the gradient threshold value.

2.3. Profile Extraction and Morphological Operation

During the motion target video sequence extraction and the morphological operation, the basic process of using the pixel gradient of the consecutive frame is as follows: monitor the video, wherein once there is any motion target, the difference among the consecutive frames will become more obvious (the gradient value is more significantly fluctuated); calculate and judge whether the gradient of the consecutive frame is more than the preset gradient threshold value, thus to judge whether the video sequence includes any motion object [13]. In this article, the gradient differentiation operation adopted for the consecutive frame includes the following steps:

Step 1 (calculation of absolute gradation difference amplitude): the following three continuous source images are assumed to exist in the original motion video sequence, and the forms thereof are respectively as $I(X,n-1)$, $I(X,n)$ and $I(X,n+1)$, and the absolute gradation difference between two consecutive frames (absolute difference of gradation) is as follows:

$$D_{(n-1,n)}(x,y) = |W \times I(X,n) - W \times I(X,n-1)| \tag{4}$$

$$D_{(n,n+1)}(x,y) = |W \times I(X,n+1) - W \times I(X,n)| \tag{5}$$

In formulae (3) and (4), W is the noise point suppression parameter.

Step 2 (calculation of differential binary image): based on the selected gradient threshold value T_1 of the consecutive frame image, the image binaryzation is respectively executed for the absolute gradation differences $D_{(n-1,n)}(x,y)$ and $D_{(n,n+1)}(x,y)$ of the two consecutive frames mentioned above, wherein the corresponding binary image matrixes are respectively as $B_{(n-1,n)}(x,y)$ and $B_{(n,n+1)}(x,y)$, then relevant mathematical operation is executed for $B_{(n-1,n)}(x,y)$ and $B_{(n,n+1)}(x,y)$, similarly executed for the original image, and after the logical operation, the following differential binary image calculation formula can be obtained:

$$\begin{aligned}&if\ B(n-1,n)(x,y)\cap B(n,n+1)(x,y)=1,\\&then\,D_{S,n}(x,y)=1\end{aligned}\tag{6}$$

$$\begin{aligned}&if\ B(n-1,n)(x,y)\cap B(n,n+1)(x,y)\neq 1,\\&then\,D_{S,n}(x,y)=0\end{aligned}\tag{7}$$

The outer profile of the object in the motion video sequence can be obtained during the above algorithmic process, but due to the edge discontinuity and the noise point interference, the outer profile extracted from the motion object may have the problem of information incompleteness. In order to solve the above problem, the morphological method is adopted in this article, and the specific process is as follows: give image data sets A and B, wherein A is expansion coefficient matrix and B is element structure, and record the process as $A \oplus B$; if B corrodes A, then record the process as $A\Theta B$. Additionally, opening and closing operations are the two key operations of the morphological method, and the specific process can be defined as follows:

$$\begin{cases}A\circ B=(A\Theta B)\oplus B\\A\bullet B=(A\oplus B)\Theta B\end{cases}\tag{8}$$

Through the above morphological operation, the motion sequence image can be denoised, and the elimination of the noise points can make the motion object have more smooth and continuous outer profile, thus to further improve the motion object identification and detection accuracy.

2.4. Accurate Division of Background and Target

The essence of dividing the motion video sequence by watershed algorithm is to regard the motion video sequence as the landform topology in topography for relevant processing, and the gradation corresponding to the pixel point in the image denotes the altitude at the position, wherein the surrounding area covered by the local gradation minimum point and the present point is called reception basin, and the boundary area of the reception basin is namely the so called watershed structure. Substantially, the watershed conversion has the features of simple algorithm, direct principle and algorithm execution parallelization potential, so multiple independent local reception basins will be formed at the same time when the watershed algorithm is adopted to divide the motion image sequence, and such phenomenon is called the “excessive division” phenomenon in the watershed algorithm. In order to solve the “excessive division” problem, a watershed algorithm with memory function is designed in this article [14].

The local minimum reception basin area extracted by the memory watershed algorithm through the above method is substantially the boundary with maximum gradient obtained from the motion image sequence, namely the outer profile of the motion image. Therefore, the substantial function of the watershed algorithm is in allusion to the image gradient sequence. In this article, the disc gradient expansion template is adopted to process the morphologically operated pixel points in order to realize the correspondence between the template center and the pixel point. Specifically, the morphological gradient form can be expressed as follows:

$$G(x,y)=I(x,y)\oplus B(x,y)-I(x,y)!\ \ B(x,y)\tag{9}$$

 Copyright ⓒ 2016 SERSC

In formula (9), $B(x,y)$ is the disc gradient expansion template with the size of three pixel point radiuses, the operators $\oplus$ and ! respectively denote the expansion and the corrosion in the morphological operation; $I(x,y)$ and $G(x,y)$ respectively denote the morphological processing result of the motion sequence image and the corresponding gradient distribution diagram.

The memory watershed division process mentioned above includes the following steps:

Step 1: Define the internal memory marks of the motion video image sequence: design the disc gradient expansion template with the size of two pixel radiuses, then execute the morphological corrosion process for the motion object, and then regard the corroded motion object as the internal memory mark of the watershed algorithm;

Step 2: Define the external memory marks of the motion video image sequence: design the disc gradient expansion template with the size of twenty pixel radiuses, then execute the morphological expansion process for the motion object, and then regard the expanded motion object as the external memory mark of the watershed algorithm;

Step 3: During the process of setting the gradient image of the motion video sequence, take the internal memory mark area and other areas out of the external memory mark area as the global reception basin minimum value of the motion video image sequence, and then remove most local extreme points in the above gradient image in order to make the outer profile of the motion video object located in the annular distribution area between the internal and external memory marks;

Step 4: Accurately divide the outer profile of the object in the motion video sequence according to the watershed operation in the annular distribution area between the internal and external memory marks.

3. Experiment and Analysis

3.1. Setting of Experimental Environment

The hardware setting of the experimental environment is as follows: processor: CPU $\mathrm{i3-2440\ 2.20GHz}$; operating system: win7 flagship; internal memory: $\mathrm{RAM\ 8G\ ddr3\ 1333}$; simulation software: Matlab R2012a. In order to verify the algorithm performance, the motion video sequence object detection algorithm designed in literature [15] is adopted as the comparison verification algorithm. The following four video images are selected as the experimental objects, as shown in Figure 2.

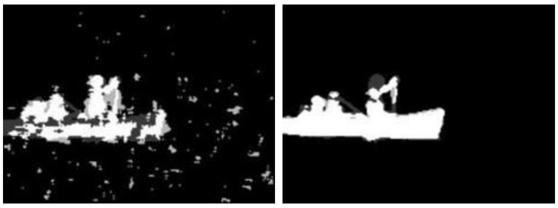
(a) Motion Object 1 (b) Motion Object 2

(c) Motion Object 3 (d) Motion Object 4

Figure 2. Experimental Objects

3.2. Experiment Result and Comparison Analysis (Qualitative)

The object division and extraction experiment results respectively obtained from DGETMW algorithm and the algorithm proposed in literature [15] for the above four video images are as shown in Figure3~ Figure6. According to the comparison results, it can be visually seen that compared with the algorithm proposed in literature [15], DGETMW algorithm can more accurately divide the motion target and can avoid the adverse influence brought by the "excessive division" phenomenon on the division accuracy to obtain relatively ideal detection effect, and such algorithm is obviously superior to the algorithm proposed in literature [15].

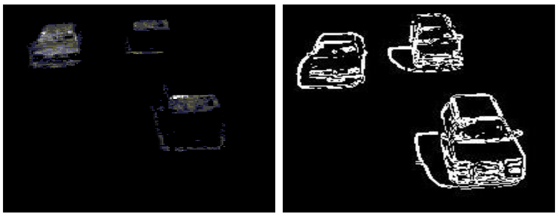


(a) Detection Result of Algorithm in Literature [15] (b) Detection Result of DGETMW Algorithm

Figure 3. Detection Result of Motion Object 1

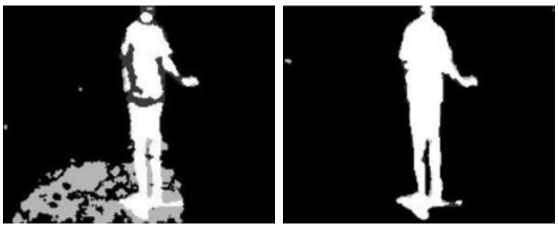
(a) Detection Result of Algorithm in Literature [15] (b) Detection Result of DGETMW Algorithm

Figure 4. Detection Result of Motion Object 2



(a) Detection Result of Algorithm in Literature [15] (b) Detection Result of DGETMW Algorithm

Figure 5. Detection Result of Motion Object 3



(a) Detection Result of Algorithm in Literature [15] (b) Detection Result of DGETMW Algorithm

Figure 6. Detection Result of Motion Object 4

Copyright ⓒ 2016 SERSC

3.3. Experiment Result and Comparison Analysis (Quantitative)

In order to more accurately compare and analyze the division effect on the motion video sequence objects, the three parameters, namely accuracy (AC), true positive rate (TPR) and false positive rate (FTR), are selected as the evaluation indexes to quantitatively analyze the division effect on the motion video sequence objects. The above three indexes are defined as follows:

True Positive Rate: $$TPR = \frac{TP}{TP + FN} \tag{10}$$

False Positive Rate: $$FPR = \frac{FP}{FP + TN} \tag{11}$$

Detection Accuracy: $$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

In formulae (10) ~ (12), TP refers to the division quantity of the object pixels in expert division; FN refers to the division quantity of the object pixels not in expert division; TN refers to the quantity of the background pixels in the expert division achieved by the algorithm.

The comparison data of the motion video object detections respectively realized by DGETMW algorithm and the algorithm proposed in literature [15] are as shown in Table 1. According to the experimental data in Table 1, compared with the motion video division detection algorithm proposed in literature [15], DGETMW algorithm can effectively improve the detection accuracy and the true position rate of the motion video object and meanwhile reduce the false positive rate of the motion object, thus to have strong universality to the interference object detection in multiple complex scenes.

Table 1. Quantitative Experiment Result Comparison

Video Image No.	Algorithm in Literature [15]			DGETMW Algorithm		
	TPR	*FPR*	*AC*	*TPR*	*FPR*	*AC*
1	0.9889	0.0539	0.9465	0.9962	0.0131	0.9884
2	0.9293	0.0378	0.9534	0.9887	0.0036	0.9931
3	0.9745	0.0508	0.9598	0.9889	0.0086	0.9935
4	0.9595	0.0541	0.9793	0.9904	0.0059	0.9941

3.4. Algorithm Operating Speed Comparison

Algorithm operating speed is an important index in motion video detection process. Therefore, the simulation experiment is carried out to compare the operating speeds of the algorithm proposed in literature [15] and DGETMW algorithm proposed in this article, and the experiment results of the two algorithms are as shown in Figure 7. Compared with the algorithm proposed in literature [15], the operating time of DGETMW algorithm is less than that of the algorithm proposed in literature [15]. Obviously, compared with the algorithm proposed in literature [15], DGETMW algorithm proposed in this article has faster operating speed, thus to be more applicable to practical real-time detection.

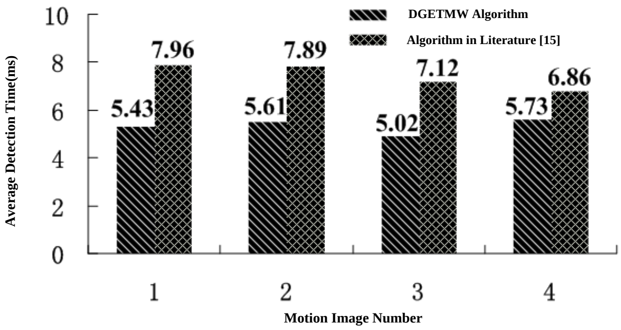


Figure 7. Detection Speed Comparison of the Two Algorithms

4. Conclusion

In order to further improve motion video object division detection accuracy and algorithm operating time and meet the real-time requirements, a motion video tracking algorithm based on memory watershed disc gradient expansion template is proposed in this article. According to the experiments carried out for the four standard video images, DGETMW algorithm can not only accurately detect the motion objects under the interference of noise and illumination, but also effectively solve the "excessive division" problem in the standard watershed algorithm. Due to the improvement of detection accuracy and calculation speed, DGETMW algorithm has strong practical application value.

References

[1] Z. Su, X. Zhang and X. Ou, "After we knew it: empirical study and modeling of cost-effectiveness of exploiting prevalent known vulnerabilities across IAAS cloud", Proceedings of the 9th ACM symposium on Information, computer and communications security. ACM, **(2014)**.

[2] W. Gu, Z. Lv and M. Hao, "Change detection method for remote sensing images based on an improved Markov random field", Multimedia Tools and Applications, **(2016)**.

[3] Z. Lu, C. Esteve, J. Chirivella and P. Gagliardo, "A Game Based Assistive Tool for Rehabilitation of Dysphonic Patients", 3rd International Workshop on Virtual and Augmented Assistive Technology (VAAT) at IEEE Virtual Reality 2015 (VR2015), Arles, France, IEEE, **(2015)**.

[4] Z. Chen, W. Huang and Z. Lv, "Uncorrelated Discriminant Sparse Preserving Projection Based Face Recognition Method", Multimedia Tools and Applications, **(2016)**.

[5] Z. Lv, A. Halawani, S. Feng, H. Li and S. U. Rehman, "Multimodal Hand and Foot Gesture Interaction for Handheld Devices", ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM). 11, 1s, Article 10, **(2014)**, pp. 19.

[6] K. Leng, W. Shi, J. Chen and Z. Lv, "Designing of a I-shaped less-than-truckload cross-dock: A simulation experiments study", International Journal of Bifurcation and Chaos, **(2015)**.

[7] Y. Lin, J. Yang, Z. Lv, W. Wei and H. Song, "A Self-Assessment Stereo Capture Model Applicable to the Internet of Things", Sensors, **(2015)**.

[8] J. Bing, X. Ke and L. Jiankun, "Analysis and management of the operation and management of power generation units during the low load period", Jiangsu electrical engineering, **(2015)**.

[9] H. Ying, "Research on transmission network planning based on improved hybrid differential evolution algorithm", Jiangsu electrical engineering, **(2015)**.

[10] S. Mingchi and Y. Dong, "Research and optimization of Nanjing electric power communication transmission network", Jiangsu electric engineering, **(2015)**.

[11] W. Ke, "Next generation job management systems for extreme-scale ensemble computing", Proceedings of the 23rd international symposium on High-performance parallel and distributed computing. ACM, **(2014)**.

[12] L. Tonglin, "Distributed Key-Value Store on HPC and Cloud Systems", 2nd Greater Chicago Area System Research Workshop (GCASR), **(2013)**.

[13] Y. Geng, J. Chen, R. Fu, G. Bao and K. Pahlavan, "Enlighten wearable physiological monitoring systems: On-body RF characteristics based human motion classification using a support vector machine", IEEE transactions on mobile computing, vol. 1, no. 1, **(2015)**, pp. 1-15.

[14] J. He, Y. Geng, F. Liu and C. Xu, "CC-KF: Enhanced TOA Performance in Multipath and NLOS Indoor Extreme Environment", IEEE Sensor Journal, vol. 14, no. 11, **(2014)**, pp. 3766-3774.

 Copyright ⓒ 2016 SERSC

[15] S. Zhou, L. Mi, H. Chen and Y. Geng, "Building detection in Digital surface model", 2013 IEEE International Conference on Imaging Systems and Techniques (IST), **(2012)**.
[16] W. Ou, Z. Lv and Z. Xie, "Spatially Regularized Latent topic Model for Simultaneous object discovery and segmentation", The 2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC), **(2015)**.
[17] W. Ke, "Using Simulation to Explore Distributed Key-Value Stores for Exascale System Services", 2nd Greater Chicago Area System Research Workshop (GCASR), **(2013)**.
[18] Y. Wang, Y. Su and G. Agrawal, "A Novel Approach for Approximate Aggregations Over Arrays", In Proceedings of the 27th international conference on scientific and statistical database management, ACM, **(2015)**.
[19] Z. Lv, A. Halawani, S. Feng, S. U. Rehman and H. Li, "Touch-less Interactive Augmented Reality Game on Vision Based Wearable Device", Personal and Ubiquitous Computing, **(2015)**.
[20] J. Yang, S. He, Y. Lin and Z. Lv, "Multimedia cloud transmission and storage system based on internet of things", Multimedia Tools and Applications, **(2016)**.
[21] Y. Su, "In-situ bitmaps generation and efficient data analysis based on bitmaps", In Proceedings of the 24th International Symposium on High-Performance Parallel and Distributed Computing, ACM, **(2015)**, pp. 61-72.
[22] D. Jiang, X. Ying, Y. Han and Z. Lv, "Collaborative Multi-hop Routing in Cognitive Wireless Networks", Wireless Personal Communications, **(2015)**.

Author

Yao Nan, he was born in Jiangsu, China, in 1976. He Received M.S. Degrees in Computer software engineering from NanJing University, JiangSu, China, in 2005, respectively. He is currently work in Jiangsu Electric Power Company Research Institute. His research direction includes the power image intelligent technology and information technology. He has received a number of scientific and Technological Progress Award.

Copyright ⓒ 2016 SERSC

Copyright ⓒ 2016 SERSC