

Research on Background Segmentation Algorithm of Moving Image in Speed Skating Sport

Lei Xu

Harbin Sports University, Harbin 150008, China
xlfxd@163.com

Abstract

An important purpose of our research is to endow the computer with the ability being similar to the human vision. So, the computer can feel the motion object in the view and apprehend the behavior of the human more easily. It is studied on the representative theories and algorithms of the background subtraction with human motion monocular image. And this paper analyses the predominance and deficiency of these theory and algorithm. These algorithms include differential images, Running Gaussian average, the Mixture of Gaussians. The basic principle and steps of realization are expounded. Also the data of the evaluation are given. Experiment shows that the proposed algorithm of background subtraction is highly effective and it can cast the reflected light, shadow and inverted image well.

Keywords: Background subtraction, Gaussian mixture model, The difference method

1. Introduction

Background segmentation of human motion images is one of the leading research topics concerned too much in computer vision field in recent years, an important technique with the combination of modern biomechanics and computer vision [1-2]. The technology gains very wide and important applications, especially extensively applied in the filed of intelligent monitoring, man-machine interaction, motion analysis and virtual reality. Although in the last decades, lots of useful work were done about it, till now no general algorithm has been developed for motion image segmentation in various cases [3-4]. Existing researches on background segmentation of human motion images were presented to specific application problem. A universal algorithm needs exploration [5].

In segmenting interesting target in the scenario, the adopted method is closely associated with relative movement of both camera and target [6-7]. In the practical application system, the motion state of camera and interesting target in scene includes four combination methods:

1. Camera static-target static, which is static scene; the processing method is the one for static images, extensively applied in human recognition, iris recognition;

2. Camera static-target motion: this is a very important dynamic scenario; its processing refers to motion target detection, classification, tracking and behavior understanding, applied in pre-warning, security monitor field [8-9];

3. Camera moving-target static: mainly used in robot vision navigation, automatic generation of electronic map and 3D scene understanding; 4. camera motion-target motion: this is the most complicated situation in the moving target detection, such as surveillance system in satellite or airplane. In the paper, we discuss how to implement moving human segmentation in specific environment from the perspective of camera static-target motion condition [10].

We introduce two algorithms:

(1) one based on difference image, including inter-frame differential and background difference method;

(2) one based on background model. Here we discuss mainly online single Gaussian background model.

The method based on inter-frame difference is the one getting moving target foreground by making the difference between the current image and neighboring frame image. The method based on background difference is the one obtaining moving foreground by making difference between current image and pre-stored background image. Usually it chooses empty scene image which does not include moving target as background image and uses the first frame as background image for background differential. The key to background model method is descriptive model of background image, *i.e.*, background model. It is the foundation of background segmentation motion foreground.

2. The Difference Method

The difference method is a technique utilizing the value between current image frame and reference image to complete segmentation of moving target through subtraction in the way of pixels. The value of image can be expressed by gray value or gradient value of pixels. The result of subtraction reflects different pixels included by current image frame and reference image, such as motion target and noise. With different used reference images, difference method is divided into inter-frame difference method and background difference method. The reference image of the forward method refers to the previous one frame of the current one or combination of several foregoing frames; for background subtraction method, the adopted reference image is referential background image of the current scenario. Most existing systems employ background subtraction method because with it, the background reference image of a denoised, non-motion target can be easily acquired in static condition. To improve background subtraction method, it needs to update background reference image with the use of current image frame, making such image adaptive to changing scenes. Hereunder we discuss two kinds of segmentation methods.

2.1. Inter-Frame Difference Method

This method utilizes pixel-based temporal difference and thresholding to fetch motion area from images in two or three adjacent frames in a consecutive image sequence. The segmentation method for motion target based on time difference has strong adaptability to dynamic environment; however, it's unable to extract all relative feature pixels, and that cavitation can easily appear in moving object body. The variation between the k th frame image $f_k(x, y)$ and the $(k+1)$ th frame image $f_{k+1}(x, y)$ can be presented by a binary difference image $D(x, y)$ as:

$$D(x, y) = \begin{cases} 1 & \text{if } |f_{k+1}(x, y) - f_k(x, y)| > T \\ 0 & \text{others} \end{cases} \quad (1)$$

We choose one frame of motion video in ice arena place to get segmentation result as Figure 1, and Figure 2, after binarizing the frame with neighboring frame difference method. It can be seen from the Figure 1, and Figure 2. In spite of desirable segmentation effect, empty space incurs inside the moving human body.

Inter-frame difference reflects fewer pixels in target moving state, and object which is cut through inter-frame difference between two consecutive frame images is much bigger than real object. To improve neighboring inter-frame difference effect of successive images, we can segment motion target with the help of three consecutive frame images. At the moment, to judge whether a pixel belongs to moving target, the principle to be followed by is:

$$D(x, y) = \begin{cases} 1 & \text{if } |f_k(x, y) - f_{k-1}(x, y)| > T \text{ and } |f_{k+1}(x, y) - f_k(x, y)| > T \\ 0 & \text{others} \end{cases} \quad (2)$$



Figure 1. Source Color Image

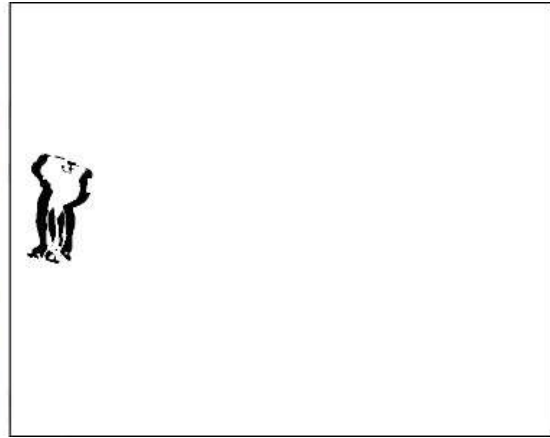


Figure 2. Result of Subtraction on Binary Image

The background segmentation method of moving image based on inter-frame difference features with less complicated computation and easy implementation. It's not sensitive to gradual change of lights in scenario. The detection is effective and stable, with low false drop rate and omission ratio, as well as strong robustness. However in general, it can't wholly segment and fetch all relevant feature pixels that cavitation is produced inside moving object. The segmentation result is not accurate especially when target moves fast and there's too big moving displacement between adjacent frames, resulting in big covered and uncovered background areas of moving changeable region in differential image. Moreover, virtual shadows easily appear; hence accurate extraction of the motion target area is greatly affected.

2.2. Background Subtraction Method

The background subtraction method is designed to get motion foreground by making difference between current image and pre-stored background image. It's simple and easily carried out. Generally the method uses images which do not include motion target as background image. However, in most cases in reality, that condition is not met. The selection of background image has huge influence on segmentation effect. Some algorithm regards the first frame image as background image, requiring camera to capture scene graphs without moving target. That technique has fatal shortcoming. When moving target is included in the first frame, the segmented outcome has virtual image. Some improved approaches overcame that weakness, segmentation result forwarded to background image and that the improved methods are adaptive. But, they're too weak to deal with situations where background changes heavily or background elements repeat moving [11-14].

We choose a set of motion video in ice arena condition, with one frame whose image doesn't include moving body and the first frame which contains moving target as background image differential result. After binarization, the segmentation results of two images are acquired.

Through observations, we find better segmentation effect is realized in the first group by selecting scenario which does not include moving target as background picture; while interference area of moving target in the first frame appears and erroneous segmentation shadow is formed in the other group where the first frame picture which contains moving target is chosen as background image. In this case, it's seen that the selection of

background image has decisive influences on segmentation result. It is shown in Figure 3, and Figure 4.



Figure 3. Source Color Image

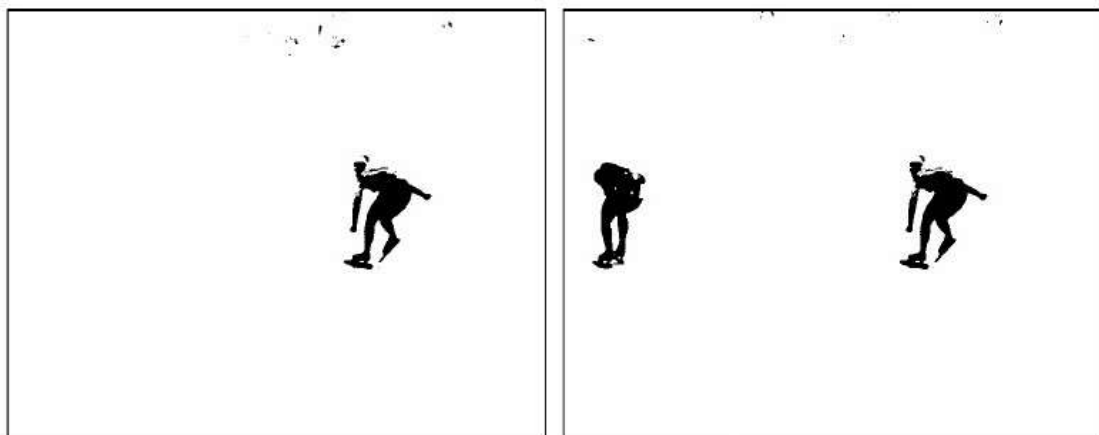


Figure 4. Result of Subtraction on Binary Image

2.3. Method based on Background Model

The method based on background model is the most popular background segmentation technology so far, which utilizes current image to match with background model and then splits out moving foreground through binarization. With fixed camera as premise, for background image sequence, this method assumes background follows Gaussian distribution and background noise is white noise. It can utilize consecutive N-frame images which initially don't include foreground to construct initial background model; then updates background model at certain renewal speed to make it compliant to slow illumination change. The background model conforms to natural condition. It can be adaptive to slow background change, with strong fitness. But it can't deal with sudden illumination change and background disturbance. For object which has shadowy image, it can't segment accurately out foreground object, causing inconvenience to subsequent processing. Most of the existing visual monitoring systems do foreground detection with background model. Up to now, a plenty of researchers have been focusing on the performance of various background models to minimize possibly the effect of scene

changes on precise division.

As mentioned above, the most difficult point in moving object analysis technology of background model method doesn't rest with t-point frame image and background; instead, the key is whether a good background model (description of real-time background and relative statistical features) can be maintained as always in the whole process. Background model is created on the whole through two ways: regression and non-regression. Generally speaking, during background extraction computation, non-regression algorithm requires certain buffer memory to maintain sample data of motion calculation during background extraction calculation. Regression algorithm models background of single-frame image based on all input frames, without saving buffer memory of sample data for supporting.

Non-regression background model method utilizes dynamically the observational data starting from one time point to the current point as sample to model background template. During background modeling, what's usually applied is the recently observational data saved in the buffer. Hence it's named non-regression background modeling method. Regression algorithm doesn't need maintain and reserve buffer zone of background estimation frame in the estimation. It updates background model at one time point based on each input frame image in the way of regression.

2.4. Online Gaussian Background Model

Background model has strong capacity of resisting disturbance against moving target because in the renewal process, the model is "trained" by each new acquired pixel value, which, in real application, whether belongs to the background or moving target. What we expect to see is background model is "trained" by using background pixel value. The training of moving target's pixel value can impair the accuracy of background model, particularly when color uniform moving object is big or moves slowly, such long-time "training" will result in false scene background.

For each pixel point in video image, the variation of its value in sequence image can be regarded as random process where pixel value is generated. The process is described as: $\{X_1 \dots X_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\}$.

Single Gaussian distribution model is suitable for single mode background. It uses single Gaussian distribution $\eta(X_t, \mu_t, \Sigma_t)$ to represent color value of each pixel point. In the distribution density function $\eta(X_t, \mu_t, \Sigma_t)$, lower subscript t means time; μ_t refers to mean value of Gaussian distribution at moment t; Σ_t is covariance of Gaussian distribution.

Gaussian mean model is an idealized Gaussian probability density function constructed with the last N pixel values. The renewal of single Gaussian distribution background model means that of Gaussian function parameters which describe scene background. We introduce learning rate $\alpha (0 \leq \alpha \leq 1)$ to indicate renewing speed of parameters. When a new frame image appears at time t, parameters of pixel Gaussian distribution are updated by following equation:

$$\mu_t = \alpha I_t + (1 - \alpha) \mu_{t-1} \quad (3)$$

$$\Sigma_t = \alpha d_{t-1} d_{t-1}^T + (1 - \alpha) \Sigma_{t-1} \quad (4)$$

Single Gaussian background model can deal with tiny and slow changing scene. When complex scene background changes greatly or suddenly, or background pixel value is of multimode distribution, e.g., tiny repetitive movement, background pixel value varies quickly, not transiting from one relatively steady single-modal distribution to another. For the moment, single Gaussian background model is too hopeless to depict background

accurately.

Still we choose motion video used in previous part. We select image of moving human at the 17th frame, the segmentation result without after treatment by online single Gaussian background model. It is shown in Figure 5, and Figure 6.



Figure 6. Result of Subtraction on Binary Image

The observations can be found online single Gauss background model in complex ice environment, the segmentation effect of the athletes is better, but there is a certain interference area. In the segmentation results can clearly see the reflection of light and athletes on the ice caused by shadow.

3. Experimental Analysis and Results

3.1. Experimental Video Image Acquisition

Data samples here are videos of highway speed skating team athletes doing curve

training in AVI format got by monocular static camera, at 25fps; each frame image is level 256 colorful image of 576*720 pixels. Apart from static background information and dynamic target information (skating players) in data samples, there are numerous noises, typically such as athlete shadow, obscure inverted image of players who do not appear in other motion detection scene, as well as other dynamic non-target areas. The background features of those video data samples agree with the described properties of complicated ice arena scene in the above.

3.2. Experimental Result Analysis

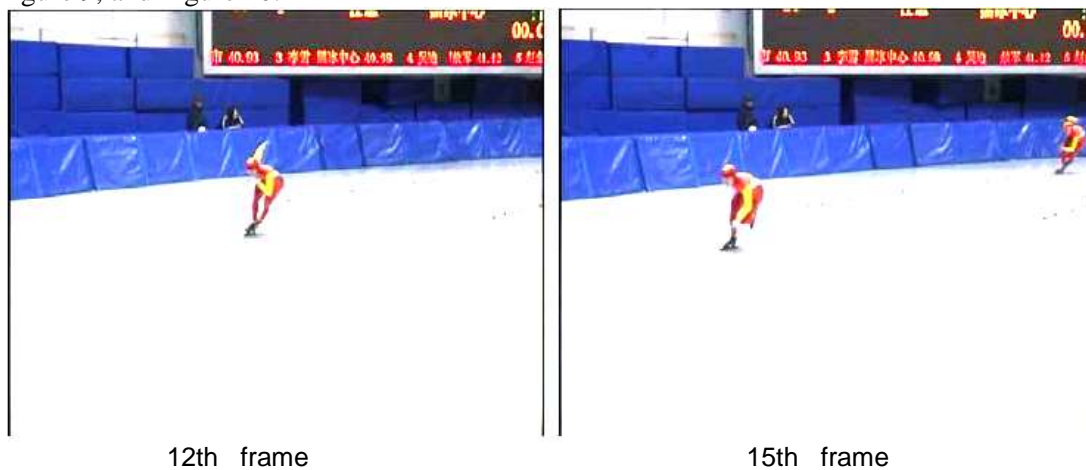
To validate effectiveness of the proposed algorithm, we perform simulations with a few kinds of videos and compare it with Gaussian mean method, hybrid Gaussian model. The experimental environment is IV 3.0GHz CPU Pentium, RAM is 512MB, Simulation platform is C++6.0 Visual.

Data samples used here are videos of highway speed skating team athletes doing curve training in AVI format got by monocular static camera, at 25fps; each frame image is level 256 colorful images of 576*720 pixels. The selection of the parameter values in the algorithm is as shown in Table 1.

Table 1. Choice of the Parameter

Parameter name	Parameter symbol	Experimental selection value
The number of Gauss distribution	K	4
learning rate	α	0.3
Luminance deviation	τ_{alo}	0.85
Chroma bias threshold	Th_CD	50
Luminance deviation threshold	Th_BD	0.9

The experiment data of twelfth frames, Fifteenth frames, seventeenth frames and nineteenth frames in the experiment video are selected, it is shown in Figure 7, Figure 8, Figure 9, and Figure 10.



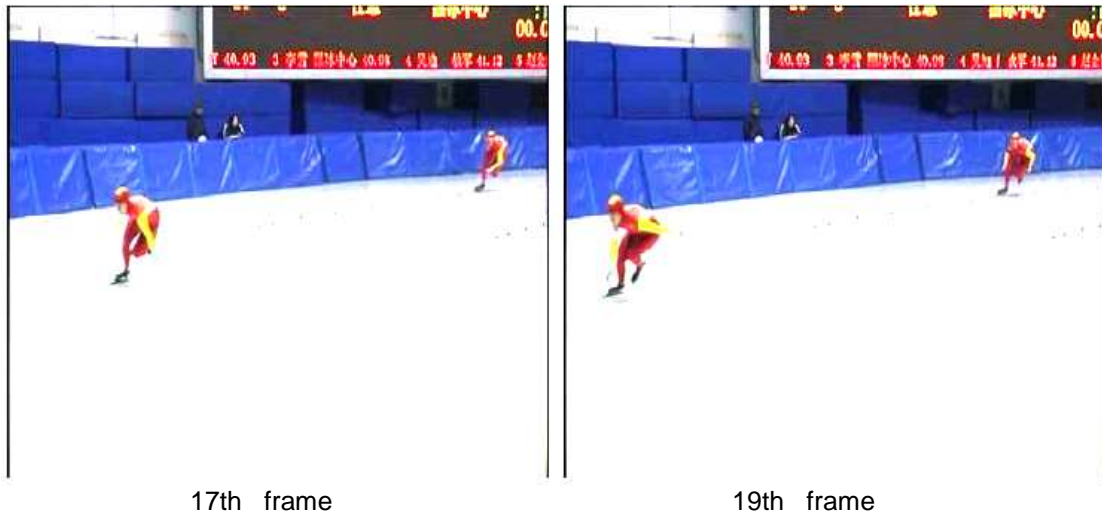


Figure 7. Source Color Image

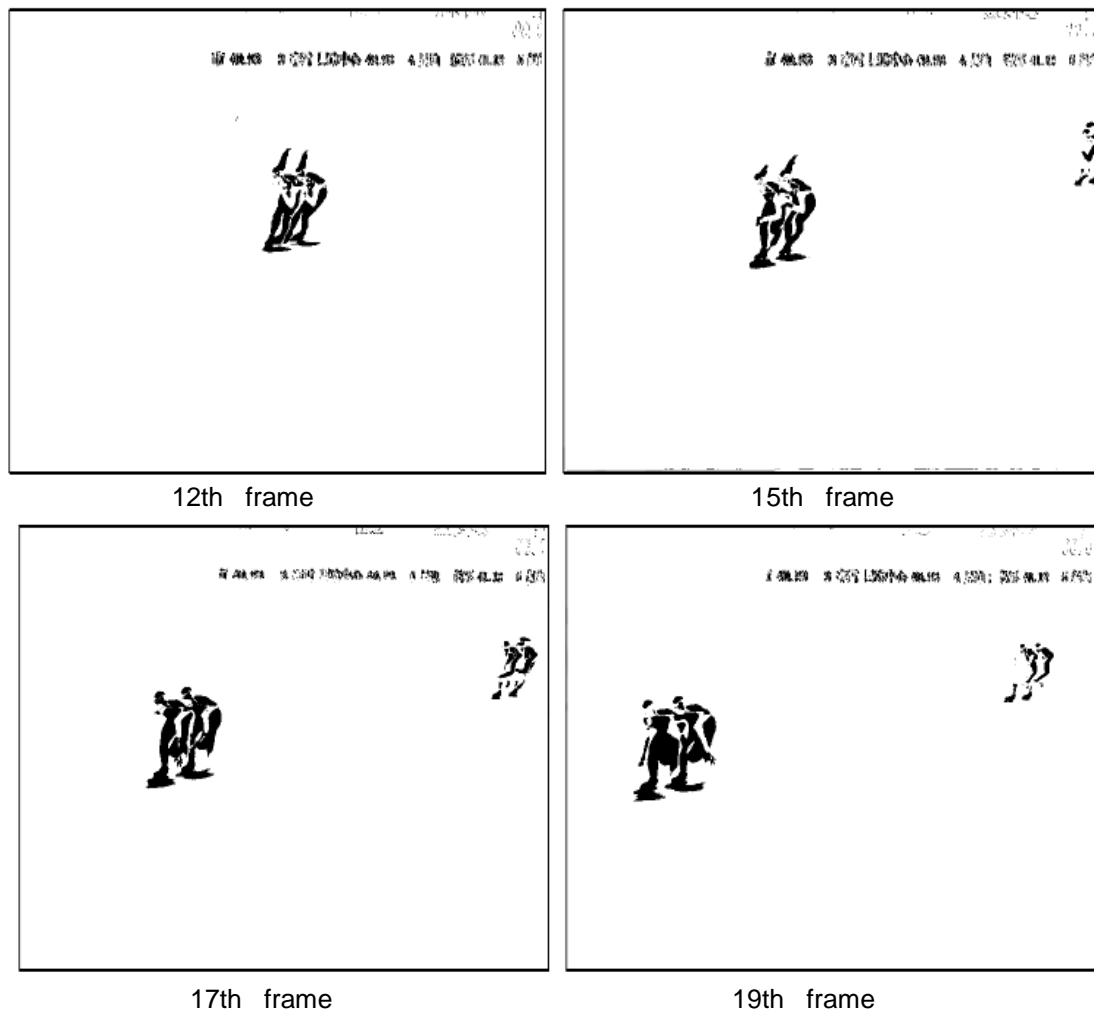
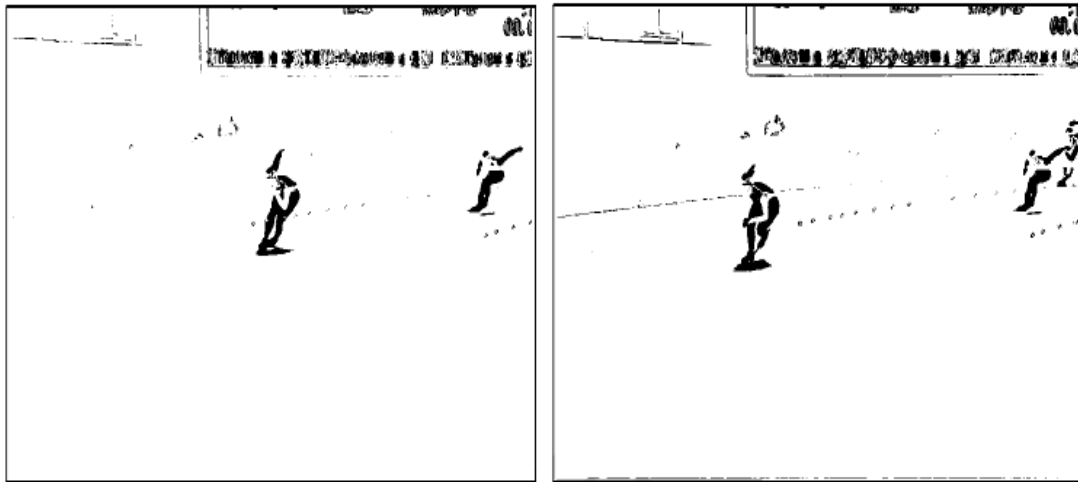
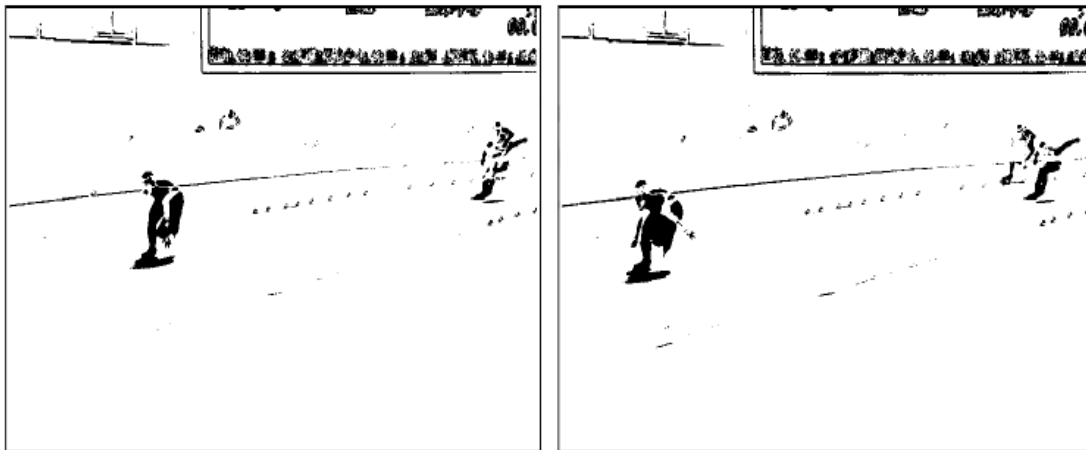


Figure 8. Result of Subtraction on Binary Image between Adjacent Frames



12th frame

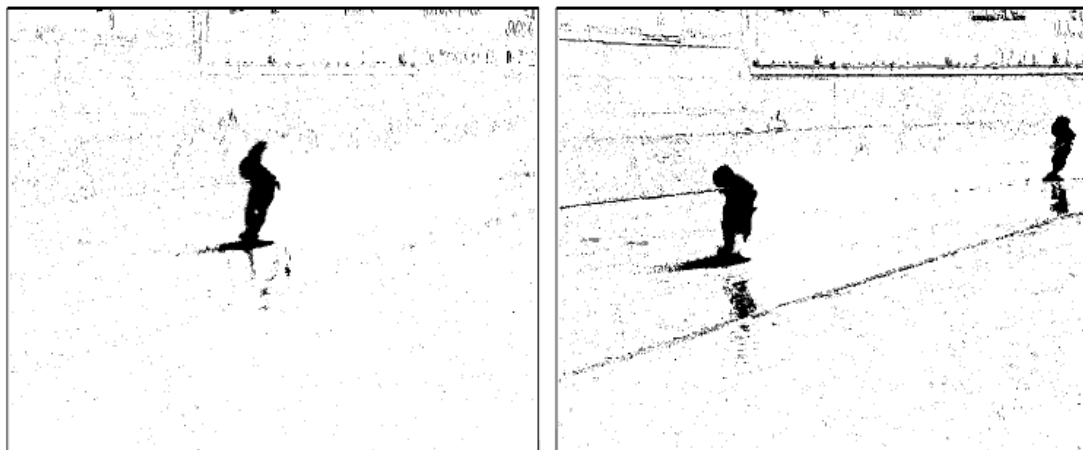
15th frame



17th frame

19th frame

Figure 9. Result of Subtraction on Binary Image with First Image



12th frame

15th frame

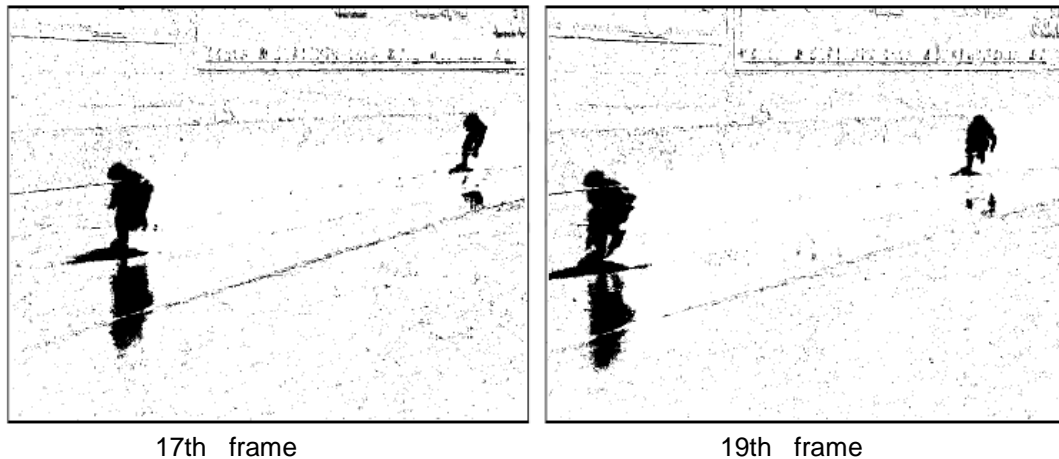


Figure 10. Result of Subtraction on Binary Image with the Method of Running Gaussian Average

With subjective observation method, *i.e.*, visual observation of segmentation results, we get:

1) In four pictures of segmentation results by inter-frame difference method in Figure 8, large cavitation occurs in motion human body; and because moving target skates too fast, double shadow forms, implying that in ice arena place, inter-frame difference method doesn't fit the scene owing to too fast moving speed of athletes;

2) Figure 9 is effect graph acquired by selecting the first frame as background image to make difference. Although target human segmentation is well achieved, moving targets existing in the first frame lead to the 1st frame moving target in segmented results, forming interference area. It reveals that with the method, the selection of suitable scene background picture affects a lot segmenting results. This method has special requirements for the acquisition of video information.

3) Four pictures in Figure 10, portray background segmentation without after treatment by online single Gaussian model method. We observe that too many interference areas appear in the picture and there are plentiful impulse noises. However it has the advantage: renewing dynamically background model and adaptive to background changes.

In short, as experimental results suggest, the method presented here outperforms other background segmentation algorithms, which not only segments out dynamically foreground information of players and also inhibits effectively shadow, intense illumination and can recognize inverted image on ice surface. Therefore, in complicated ice arena scene, our algorithm stands out by achieving better processing effect.

4. Conclusion

This paper mainly introduces the common background segmentation algorithm of human body moving image, including the difference method and background difference, based on the probability density background model method in the online Gauss background model method. Experiment shows that the proposed algorithm of background subtraction is highly effective and it can cast the reflected light, shadow and inverted image well.

References

- [1] F. Zhili, "Research on segmentation and recognition algorithm of moving human body image in complex environment", Changsha University of Science and Technology, (2013).
- [2] L. Feng, "Pedestrian motion analysis in intelligent video surveillance system", University of Science & Technology China, (2011).

- [3] C. Meiling, "Methods of 3D human motion analysis and motion recognition", Central South University, (2013).
- [4] M. Fanmin and J. Jiang, "A segmentation algorithm for the digital high speed photographic images of sand particles movement", Chinese Journal of mechanics, no. 1, (2012), pp. 82-87.
- [5] W. Wancheng and W. Juanli, "Research and implementation of human motion analysis system based on video sequence", Computer applications and software, vol. 23, no. 2, (2012), pp. 140-141
- [6] L. Jun Guo, L. Xi, Z. Jieyu and Zhongzhi, "Combined with motion information and apparent characteristics of pedestrian detection", Journal of software, no. 2, (2012), pp. 299-309.
- [7] M. Ming and Y. Fangbo, "Human motion detection based on depth image of the Kinect", Instrumentation journal, no. 2, (2015), pp. 386-393.
- [8] S. Z. Shiyu, L. Huai and Longfei, "Image sequences of a dynamic scene in moving target detection method", Journal of image and graphics, no. 9, (2007), pp. 1590-1597.
- [9] T. Yong and J. Yuming, "A human motion segmentation method based on color image", Microelectronics & computer, no. 3, (2006), pp. 70-72
- [10] W. Dian, "Research on background modeling and shadow suppression algorithm based on hybrid Gauss", Northwestern Polytechnical University, (2013), pp. 20-40
- [11] T. Jianhui and P. Baochang, "Human thermal image segmentation in complex background. Computer engineering and application, no. 14, (2011), pp. 13-16.
- [12] C. Wren and A. Azarbayejani, "Pfinder : Real-Time Tracking of the Human body", In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, (1997), pp. 780-795.
- [13] J. M. Stephen, S. Jabri and Z. Duric, "W4: Real-Time Surveillance of People and Their Activities", IEEE Transactions on Pattern and Machine Intelligence, vol. 22, no. 8, (2000), pp. 35-46
- [14] D. Gutches, "A background model initialization algorithm for video surveillance", In: Proc. of the 8th IEEE Int'l Conf. on Computer Vision Vancouver, (2001), pp. 733-740.

Author



Lei Xu, He has obtained his B.S and M.S degree from College of Education, University of Canberra, Australia. He is a lecturer at Harbin Sports University. He is specialising in speed skating, speed skating teaching and training.

