

# Research on Video Object Segmentation Algorithm Based on Background Reconstruction

Feng-min Cheng

College of Intelligence and Information Engineering, Tangshan College, Tangshan 063000, China  
69670768@qq.com

**KEYWORDS:** Background Reconstruction, Shadow elimination, background frame

**Abstract:** Aiming at the problem of occlusion and shadow in the existing video object segmentation algorithm, a new video object segmentation based on background reconstruction is proposed. In the time domain, the motion area of a single frame image is obtained, and the method of changing the template with the frame difference or the two frame difference map is discarded. By using the two frame difference to obtain the motion information of the single frame, the threshold is obtained by the frame difference mask image, thus improving the accuracy of background reconstruction. The threshold value of the process is obtained by LHS method. The method can generate adaptive threshold in time. Finally, a method of eliminating the shadow is proposed.

## Introduction

In MPEG-4 standard, the video object is defined as a unit in the scene to allow users to access (search, browse) and operation (cut and paste), namely video object is the accumulation of the area and at least one common characteristic of consistently appeared in the video object. Video object segmentation is according to certain standards will be video sequence segmentation for the combination of semantic entity has a certain significance. Due to the video reflects the objective world is complex, the object itself is also a variety of, the lack of the definition of a very apt description. Any semantic meaning in the real world entities, such as driving cars, people, etc., can be regarded as semantic video object. Moreover, on the same video scene, the application of different interest of video object is different.

Video object segmentation algorithm is based on the direct use of space and time characteristics to segment and track the moving object, which directly use the sequence of two frames or multi frame detection and object segmentation method for the object segmentation method is simple and has been widely used. The background part of the object movement is detected as a result of the change. The formation of occlusion is the inevitable result of video object motion, and it is a difficult problem in video processing. Some scholars have proposed the sprite technology according to the MPEG4, realized the concrete improvement algorithm, and obtained the good segmentation effect. But the algorithm is still relatively complex, and although it has solved the problem of occlusion, it will appear the problem of the video object.

Based on the full study of all kinds of background algorithms, a new algorithm based on background reconstruction is proposed, which is simple in principle and has solved the problem of shadow.

## Algorithm introduction

Basic idea is to use two frames directly after change detection of movement object cannot eliminate the shade area, using more reliable enough background information to get a complete background

of heavy composition, using the current frame secondary frame difference "and" method can eliminate the influence of background shade. First, the basic steps of method applies the secondary frame difference "and" said video object motion area of the frame of the change detection mask, then use the binary mask map to obtain the background images, through a certain algorithm to initialize the background frame in the end, the current frame subtracting reconstruction background image to eliminate the influence of background shade, get more accurate video object. In which in the process of closing value, this paper puts forward a eliminate the shadow of worshipping the value increment, so as to solve the problem of shadows.

Algorithm is divided into two parts: background reconstruction and video object extraction. Content is as follow: (1) Background reconstruction: firstly, the two frame difference mask is extracted from the video frame stream by using the two frame difference. Then the background of the reconstruction is obtained by using the frame difference mask. (2) Video object extraction: first from the stream of video frames readout to be segmented video frame, the reconstruction and background subtraction, and combining the shadow elimination method, obtained initial without shadow mask, at last, combined with the spatial segmentation results, get the final segmentation mask.

### **Background frame initialization**

When the object is moving, the occlusion of each background pixel is exposed for a period of time.  $B(x, y)$  as the background image, the  $s(x, y)$  and a counter  $c(x, y)$  are set at each pixel of the  $B(x, y)$ .  $s(x, y) = 0$  indicates that the point is not initialized,  $s(x, y) = 1$  indicates that the point has been initialized. The counter records the number of changes in the detection results of the change detection results in the change of the frame. The counter is designed to take full account of the complex motion of the object.

Firstly, the value of the first frame in the video sequence is assigned to the background frame, that is,  $B(x, y) = f(x, y, 1)$ . For any  $(x, y)$ ,  $s(x, y) = 0$ ,  $c(x, y) = 0$ . Using the inter frame change detection technology, the detection of the change of  $CDM_{n,n-1}(x,y)$  between  $n$  and  $n-1$ . The pixel value of  $CDM_{n,n-1}(x, y)$  is 0, the corresponding counter value plus 1; otherwise the counter is cleared. When the value of the counter is greater than or equal to  $L$ , it is explained that the value of the point is not changed more than  $L$  times. If the value of  $s(x, y)$  is 0, which indicates that the background pixel point is not initialized, the current frame is assigned to the background corresponding pixel, and the corresponding  $s(x, y)$  is set to 1. That is:

$$\text{If } s(x,y)=0 \quad \text{and} \quad c(i,j) \geq L \quad \text{then} \quad B(i,j)=f(i,j,n), \quad s(x,y)=1 \quad (1)$$

When the value of the counter is greater than  $L$ , and the corresponding  $s(x, y)$  value is 1, then the updated background pixel value is updated. When the  $s(x, y) = 1$ , or the number of non-initialized pixels is less than a certain threshold value, the background frame initialization is finished.

### **Shadow elimination**

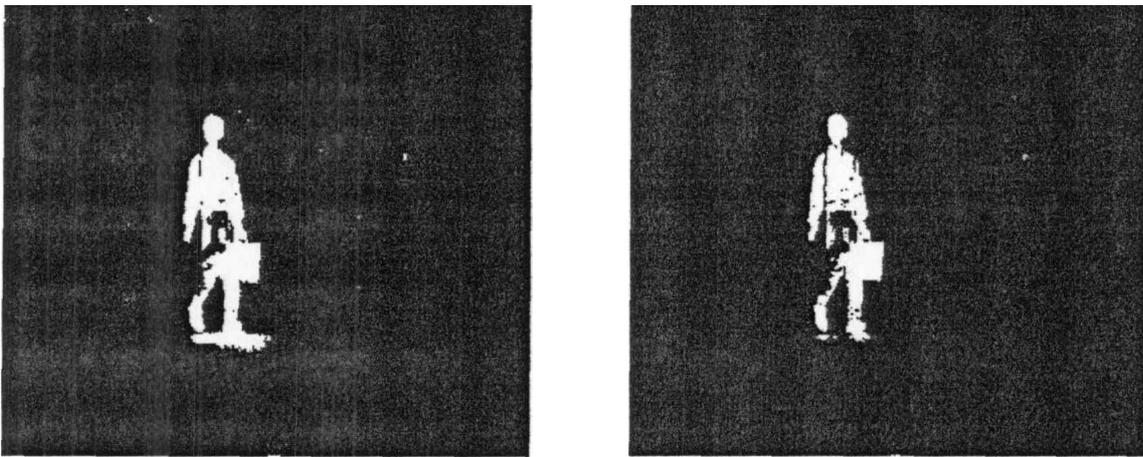
Using background frame as a reference frame, the shadow of the video object region will also be testing as part of the video object, additional mechanisms must be used to eliminate the shadow. For shadow problems, and puts forward a new method of shadow elimination.

According to the Phong model, the intensity of the reflected light is reflected by the environmental light intensity and the light intensity of the light source. For the point in the shadow, the response of the sensor is only included in the response of the light source, and the luminance component of all

points in the shadow region is less than that of the corresponding point in the light condition. To determine whether the brightness of each pixel is less than the corresponding point in the background frame for each pixel in the change region. If it is less, the pixels are for the candidate shadow regions; otherwise, the object pixels.

Is divided into two steps to eliminate the shadow region, first of all according to the luminance information to detect the regional variation may belong to the shadow pixels. Then, the difference of corresponding pixels of the background shading and video frames do thresholds for background subtraction.

The case of Before and after the shadow elimination as shown in figure 1, brightness information to detect the change area is used to calculate its shadow elimination of threshold of 50. For the increase of the threshold, it has caused the segmentation, and the segmentation results are used to reduce the effect of the segmentation.



(a) Before the shadow elimination

(b) After the shadow elimination

Figure 1 Comparison of the shadow elimination

## Conclusion

In the background frame, the value of  $L$  is 5, and the fifty-fifth frame is constructed in the thirtieth frame, and the initialization algorithm is well separated by 25, 35, 45 and 55 frames. This method can segment the video object more accurately and effectively remove the shadow.

## Reference

- [1] Y.Wang, K.Loe, T.Tan. Spatiotemporal Video Segmentation Based on Graphical Models, IEEE Trans. on Image Processing 2005, Vol. 14(7): PP.937-947
- [2] D.Xu, J.Liu, X.Li. Insignificant Shadow Detection for Video Segmentation, IEEE Trans. on Circuits System and Video Technology, 2005, Vol. 15: PP.1058-1064
- [3] Y.Liu, Y.F.Zheng, Video Object segmentation and Tracking Using  $\Phi$ -Learning Classification, IEEE Trans. on Circuits System and Video Technology, 2005, Vol. 15 (7): PP.885-899
- [4] Y Tsai, C. Lai, Y Hung, et al, A Bayesian Approach to Video Object Segmentation via Merging 3-D Watershed Volumes, IEEE Trans. on Circuits System and Video Technology, 2005, Vol. 15(1): PP.175-180
- [5] Minglum Gong, Yee-Hong Yang. Near real-time reliable stereo matching using programmable graphics hardware[J]. Computer Vision and Pattern Recognition. 2005. 1(1):924-931.
- [6] Faugeras O, Hotz B, Mathieu H, et al, Real-time correlation-based stereo: algorithm, implement and

applications, Technical Report[R].2013, 1993, INRIA.