Multiple Objects Detection Based on Background Subtraction and Gaussian Pyramid*

Liang Jiang
Computer Science & Telecommunication Engineering
Jiangsu University
Zhenjiang Jiangsu 212013, China
jiangliang0@gmail.com

Yong-zhao Zhan
Computer Science & Telecommunication Engineering
Jiangsu University
Zhenjiang Jiangsu 212013, China
yzzhan@ujs.edu.cn

***Abstract* - For the multi-object detection problems with complex background, a method which combines background subtraction with Guassian pyramid on objects detection is presented. The method detects the whole object with taking samples from the objects with Guassian pyramid, building background models, extracting foreground areas from background subtraction, and eliminating the shadow on the foreground. The detection that integrates Gauss model which concerns the background renewal of calculation, overcomes the error resulted from the sudden change of background. A dynamic threshold concept is proposed to enhance detection effect, thus it increases the possibility of implementation. The experiment results show that moving objects detected by the proposed method is more reliable than the state-of-the-art.**

Keywords- Multiple objects detection; Background subtraction; Guassian pyramid; Dynamic threshold

I. INTRODUCTION

Moving objects detection is to segment movement area from background area. It is really a challenge for correctly detecting the moving objects, because it has lots of conditions affecting the detection, such as change of the illumination, disordered motion on the background, shadow of the moving objects, shakes of the vidicon, and occlusion problem of multiple moving objects. The accuracy of moving objects detection and segmentation affects the accuracy of moving objects tracking and classification, so moving objects detection is an important problem in computer vision. There are three types of methods on moving objects detection currently: optical flow, frame difference and background subtraction.

Optical flow[1] can detect the independence moving objects in the conditions of none information in the scene, but most of optical flow methods are time-consuming and complicated in computation so that they are difficult to meet the requirements of real-time detection.

Frame difference[2] is highly adaptive, but the choice of different frames required is very high. And the method depends on the velocity of moving objects. If the velocity is high and the time interval is too large, it will result in no coverage area between the two frames, which makes this method can not segment the moving objects.

Background subtraction[3] is easy and fast, but it is sensitive to the change of illumination, but shadow is often detected as a part of moving objects, affecting the accuracy of detection results.

In order to set a better tolerance toward the change of illumination and change of the environment caused by the moving objects detection failure, and to ensure that there are shadows of moving objects and cross-movement presented can also be accurately detected. A method which combines background subtraction with Guassian pyramid on objects detection is presented in the paper, a shadow removal method is leaded in the algorithm to solve multiple moving objects efficient detection problems in video sequences, also the algorithm is robust to the problems of the cross movement of targets and occlusion.

II. RELATED WORK

To make a general survey about motion detection techniques home and abroad, we come to the following conclusions:

In [4], a moving objects detection method is present. The method uses the Otsu method to segment successive three frames of difference images, the segmented moving objects are then processed with mathematical morphology in order to eliminate discontinuity and holes which appear after the segmentation. Finally, the moving objects can be located by picking up objects centroid. The method effectively separates the background and the moving objects, extracts the moving objects more accurately, and it has significant advantages compared with the conventional detection method. but the method is affected by environmental factors seriously, and the two similar objects will make the moving objects located not correct by picking up objects centroid failure.

In [5], a objects detection method based on adaptive threshold mixed difference with the improved OpenCV clustering algorithm is presented. The method was proved to have perfect detection result on motion objects and high computing speed. But the method could easily lead to the separation of moving objects detection, and the detection range needs to be improved.

In [6], a novel object detection algorithm based on color histogram and adaptive bandwidth mean shift is proposed. The algorithm is capable of detecting objects rapidly and

* This work is partially supported by National Natural Science Foundation of China Grant #60673190 and Provincial Natural Science Foundation of JiangSu Grant #BK2009199.

precisely. Position, size, and orientation can be simultaneously and precisely detected and tracked. But all the thresholds used are selected experimentally. They do not have the self-adaptive. The algorithm only uses color histogram to distinguish the moving objects and it can't reflect the characteristics of movement accurately.

III. THE METHOD COMBING BACKGROUND SUBTRACTION WITH GUASSIAN PYRAMID ON OBJECTS DETECTION

A. Basic ideas

The basic idea of the method combing background subtraction with Guassian pyramid on objects detection is as follows: firstly, one frame image as the background frame is selected and each frame is sampled with Guassian pyramid. Secondly, we detect the images from low resolution, then calculate the frame difference and distinguish between the background points and moving regions after compared with the threshold value, so we complete the background modeling. Thirdly we extract foreground areas with background subtraction, and eliminate the shadow on the foreground considering changes in the image brightness. Finally the moving objects detection is complete. Background renewal is also carried out in the process, the background point is deemed to have rules to renew, and the area corresponding to the moving objects is not renewal.

B. Multiple objects detection

The algorithm combing background subtraction with Guassian pyramid on objects detection is as follows:

(1) The images are captured in a time period of video sequences $\{A_k \mid k = 1, \ldots, N\}$, Where k is the frame number, N is the total number of frames in video sequences.

(2) The images are sampled with Guassian pyramid, if the resolution reduced by half, the objects' velocity at least reduced to 1 / 2, so that the accuracy of detection will be increased. We sample the images with Guassian pyramid in two steps: first, do Gaussian smoothing on the images; second, continue down-sampling. We can get a thumbnail of an image with subsampling. However, if we need to reduce the size of an image, it will lose a lot of information with subsample only. According to sampling theorem, it requires that all the structures obtained that are smaller than the shortest wavelength of 1 / 4 to eliminate out through the smoothing filter, so as to obtain a correct sample images. We suppose that we set M×N image frame k A_k^0 inputing as a low-level Gaussian pyramid, and its l-level image is A_k^l, by convolution of Gaussian window function W and l-1-level image A_k^{l-1} and down sampling we have the following formula:

$$A_k^l(i,j) = \sum_{m=-p}^{p}\sum_{n=-p}^{p} W(m,n) A_k^{l-1}(2i+m, 2j+n), \tag{1}$$

Where $0 \le i < M/2^l$, $0 \le j < N/2^l$, $0 < j \le t$ (t is a positive integer representing the decomposition of the layers), [-p,p] is the range of m,n , (2p+1)×(2p+1) is the size of the window function W.

(3) We suppose we capture two l-level images in video sequences A_k and A_r. By two images differencing, we have the following difference formula at the pixel (i, j):

$$FD_{k,r}^l(i,j) = s_k^l(i,j) - s_r^l(i,j), \tag{2}$$

Where $s_k^l(i,j)$ stands for the intensity value at pixel (i, j) in frame k, $s_r^l(i,j)$ stands for the intensity value at pixel (i, j) in frame r.

Detecting the pixels belong to the l-level background could use the image intensity to subtract, we have the following formula about background function $\alpha^l(i,j)$:

$$a^l(i,j) = \begin{cases} 1; & \left|FD_{k,r}^l(i,j)\right| > T^l \\ 0; & else \end{cases}, \tag{3}$$

where T is an appropriate threshold. Here, $\alpha^l(i,j)$ is called a segmentation label field, which is equal to "1" for changed regions and "0" otherwise.

(4) Start detection from the lowest resolution image after resampling each image with the Gaussian pyramid. We calculate the l-level frame difference for each pixel:

$$FDN_{k,r}^l(i,j) = \frac{\sum_{x\in N}\left|s_k^l(i,j) - s_r^l(i,j)\right|\left|\nabla s_r^l(i,j)\right|}{\sum_{x\in N}\left|\nabla s_r^l(i,j)\right|^2 + c}, \tag{4}$$

Where N denotes a local neighborhood of the pixel (i, j), $\nabla s_r^l(i,j)$ denotes the gradient of image intensity at pixel (i, j) in l-level frame r, and c is a constant to avoid numerical instability, the value of c can determine by the experiments in the actual scene or adaptive learning. If the normalized difference is high (indicating that the pixel is moving), replace the normalized difference from the previous resolution level at that pixel with the new value. Otherwise, retain the value from the previous resolution level.

(5)Repeat these steps for all resolution levels. Finally, we subtract the detected frame and background at the highest resolution level, and then threshold the value of the image intensity difference at each pixel. If the value exceeds the threshold, the pixel with that value is the pixel of a moving object, or the pixel is the pixel of the background. At this point, we complete the work of the moving objects detection.

C. A dynamic threshold determination algorithm

Moving objects detection threshold value is very important. If it is too large, there will be some of the moving objects pixels mistaken for background pixels; if it is too small, the background pixels will be mistaken for moving objects pixels. We don't like to see that. In [7], author compares the frame difference with threshold value ε to obtain the background and foreground image. But the threshold in that paper is got from experience and experiment, it can not dynamically adapt to changing scenarios. Therefore, an algorithm use of the current image average intensity to determine the dynamic threshold is present:

(1) Find the minimum and maximum image intensity values in the l-level image, so that the initial threshold value T_0^l is equal to the average of these two values.

(2) According to the threshold we segment the l-level image into two parts: object and background, calculate the average intensity s_1^l and s_2^l of these two parts separately. We set the sum of the image intensity S_1^l at the pixel which the image intensity at the pixel is bigger than T_0^l in the l-level image, and we set the total number of pixels N_1^l which the image intensity at the pixel is bigger than T_0^l. We set the sum of the image intensity S_2^l at the pixel which the image intensity at the pixel is smaller than T_0^l in the l-level image, and we set the total number of pixels N_2^l which the image intensity at the pixel is smaller than T_0^l, we have the following formula about s_1^l and s_2^l :

$$s_1^l = \frac{S_1^l}{N_1^l}, s_1^l = \frac{S_1^l}{N_1^l}, \tag{5}$$

(3)Make use of the average value of s_1^l and s_2^l as separate threshold T^l. T^l is a global threshold value, we use T^l instead of each point with its threshold value can not reflect changes at all the pixels in the scene, but lots of experiments have shown that we still can detect moving objects as long as the proper threshold selection and adapt to the constantly changing in the new scene. Therefore, the above algorithm to obtain the threshold T^l is feasible.

D. Background renewing

The background model extracted may be subject to the interference of some unforeseen circumstances, such as illumination changes, the scenery originally belongs to the background suddenly moves, etc. In order to overcome these obstacles and change the background model adaptive to external changes, we must renew the background model in real time. Suppose the frame k is the current frame, B_k^l is the background of the l-level Gaussian pyramid image, then the method to renew B_{k+1}^l the background of the frame k+1 is as following: the pixels judged belonging to the objects still maintain the original image intensity, the pixels judged belonging to the background renew the background model in accordance with the following rules:

$$B_{k+1}^l(i,j) = \alpha^l(i,j)B_k^l(i,j) + (1-\alpha^l(i,j))s_k^l(i,j), \tag{6}$$

Where $\alpha^l(i,j) \in \{0,1\}$ is the value of the background model at pixel (i, j). $s_k^l(i,j)$ is image intensity at pixel (i, j) in l-level frame k.

Thus, due to renew the background model is not entirely dependent on the current pixel value, but relate to the previous frame, the background model can be relatively stable over time. This method not only effectively avoids some of burst phenomenon on the background, but also takes into account the impact of noise. By renewing the value of the background pixels intensity, the method can adapt to light, weather, etc, the impact from the changes in the external environment.

E. Shadow removal

Learning through observation, the shadow of movement moves with the moving objects, so after the above-mentioned method for moving object segmentation, it mostly cases the existence of shadow effects. This allows moving objects segmentation areas larger than the actual, or even may cause the moving objects segmentation areas sticking together. If as a basis for follow-up, it will lead to mistake, even error.

After detecting the boxes that surrounded the moving objects, we further consider whether the pixels of the scope are the shadow pixels. Then we build shadow modeling, and finally amend the scope of the boxes.

As the HSV color space has the good consistency of color perception, it is suit for computing the similarity between the images. By analyzing, the brightness of the region covered by the shadow changes greatly, and its hue changes little[8]. Therefore, we can eliminate the shadow by the following formula:

$$sp(i,j) = \begin{cases} 1, & \alpha_s \le \dfrac{I_v(i,j)}{B_v(i,j)} \le \beta_s \\ & \wedge \left|I_s(i,j) - B_s(i,j)\right| \le T_s \\ & \wedge \left|I_H(i,j) - B_H(i,j)\right| \le T_R \\ 0, & \text{else} \end{cases}, \tag{7}$$

Where $I_H(i,j), I_s(i,j), I_v(i,j)$ respectively represents the HSV component in the current frame. $B_H(i,j), B_s(i,j), B_v(i,j)$ respectively represents the HSV component in the background.

If sp (i, j) = 1, the pixel (i, j) is the shadow pixel. Thus by shadow modeling, we further amendments the scope of the boxes that surrounded moving objects.

Ⅳ. EXPERIMENTAL RESULTS AND ANALYSIS

To prove the effectiveness of the algorithm, we used MATLAB7.0 for programming. We use the computer ((Lenovo), Intel(R) Pentium(R) D CPU 3.00GHz 3.00GHz, 1.00GB RAM) for the experiment. Algorithm successfully detected moving objects marked with boxes in different colors. In Fig.1, the effect drawing of the detection based on background subtraction with Guassian pyramid on objects detection is presented. There are seven moving objects in the frame, in which two moving objects in the left will disappear, and the moving object underneath just appear. In Fig.2, The effect drawing of the multiple moving objects detection during cross-motion is presented. Although two moving objects occurred cross-movement and had become blocked, we detected the moving objects accurately. In Fig.3, The effect drawing of the multiple moving objects detection after cross-motion is presented. The effect drawing has shown the algorithm proposed in the paper is robust to the problem of moving objects occlusion.

Fig.1: the effect drawing of the multi-object detection

Fig 2: The effect drawing of the multiple moving objects detection during cross-motion

Fig 3: The effect drawing of the multiple moving objects detection after cross-motion

In addition, we do a lot of experiments on the ice hockey video in the simple background, statistics show that the accuracy of detection is up to 97%. Moreover we do experiments on the soccer video in the relatively complex background, statistics show that the accuracy of detection can be maintained above 92%. It can be seen that the algorithm in multi-object detection has better accuracy, and can eliminate the shadow of some of the moving objects, sequentially correctly detect the scope of the moving objects.

V. Conclusion

A method combing background subtraction with Guassian pyramid on objects detection is proposed in the paper. Compared with the related work, the method which combines background subtraction with Guassian pyramid on objects detection increases the moving objects detection accuracy and anti-jamming capability. We use Gaussian model for real-time updates on the background image to make background modeling more reasonable. At the same time we lead the dynamic threshold to the method to make moving objects detection in the changing environment more accurate.

Acknowledgment

The authors would like to thank the reviewers for their insightful comments and helpful suggestions.

References

[1] BARRON J, FLEET D, BEACHEMIN S. “Performance of optical flow techniques”. International Journal of computer Vision, vol. 12, no. 1, pp: 42-47, 1994

[2] MEIER T, NGUN K N. “Video Segmentation for Content-Based Coding”. IEEE Trans on Circuits and Systems for Video Technology, vol. 9, no. 8, pp:1190-1203, 1999.

[3] GUPT S, MASOUND O, MARTIN R F K, et al. “Detection and classification for vechicles”. IEEE Transcations on Intelligent Transportation Systems, vol. 3, no. 1, pp:37-47, 2002.

[4] Wang Jing, Bao Wen-xing. “A video supported moving object detection technique based on difference algorithm”. Computer Applications and Software, vol.26, no.12, pp: 68-70, 2009.

[5] Sun Ling, Wang Shuo, Li Qing-xia. “A target detection method based on adaptive threshold mixed difference”. Computer Applications and Software, vol. 26, no. 10, pp: 94-97, 2009.

[6] Xiaopeng Chen, Qiang Huang, et al. “Rapid and Precise Object Detection based on Color Histograms and Adaptive Bandwidth Mean Shift”. International Conference on Intelligent Robots and Systems, IROS 2009. IEEE/RSJ, pp: 4281-4286, 2009.

[7] Chaisorn L, Manders C, Rahardja S. “Video Retrieval-Evolution of Video Segmentation, Indexing and Search”. 2nd IEEE Internation Conference on Computer Science and Information Technology, ICCSIT 2009, pp: 16-20, 2009.

[8] Lv Chang-kui, Jiang Cheng-yu, Wang Ning-sheng. “A Novel Statistical Model for Motion Detection and Contour Tracing”. Geomatics and Information Science of Wuhan University, vol. 30, no. 8, pp: 724-727, 2005.