Object detection algorithm based on the video stream for complex environment

Zhaoxia Fu

Party School of Shanxi Provincial Committee of the C.P.C, Taiyuan, 030006, China

fzx2005@163.com

Keywords: Object Detection; Gaussian Mixture Models; Background Modeling; Frame-difference

Abstract. Due to the complexity of the environment and the diversity of the goal itself in the visual system, the technology of object detection has brought great difficulties. The practical experience shows that the technology of object detection is far from mature in the general sense and there are still certain gaps away from practical application. In this paper, for solving these problems of complex background and lighting mutation of the fixed monitoring scenes, a moving object detection algorithm based on the video stream is proposed and can be applied to the complex of indoor and outdoor environments. The convergence rate is improved by improved Gaussian mixture model. Taken into account the interference of light mutation and combined with the advantage of frame-difference method, the system detection capability for fast moving target is improved. Morphological method is used for noise processing of motion region, while remedying the generated hole during object detection. Experimental results show that the algorithm can complete detecting under the light mutation quickly and accurately and has a strong robustness.

Introduction

Moving object detection and tracking based on the video stream are both challenging and very broad application prospects of research topics. Moving object detection is the basis of various follow-up advanced processing such as target classification, target tracking, behavior recognition and understanding, which is at the bottom of the entire computer video intelligent surveillance system, thus it has become an important research topic in a video surveillance system [1,2]. The purpose of the motion detection in the image sequence is dividing a change area from the background, simplifying image processing, locating target region, thereby obtaining the motion information in order to track and identify objects. In the actual scenario, the factors of chaos (wave, shaking branches and other interference), weather and light conditions change, moving shadow and camera shaking make correct moving target detection and segmentation become a rather difficult task [4].

Currently, moving object detection algorithms have variety of ways endless, but most object detection theories are based on some prior knowledge, mainly based on motion information and feature information. The motion information entails some changes in the time series, and the main feature information refers to the obvious features of the detected target, such as color, shape and texture, etc. The applications of object detection based on motion information are relatively common. The conventional object detection algorithms based on motion information are frame-difference method, background subtraction method [3], optical flow method and so on. Generally existing algorithms do not have the versatility, it often need use different algorithms for different applications. In this paper, in order to accurately detect moving targets in complex scene and response to illumination mutation problems, we propose a detection algorithm based on the video stream. The algorithm improves the Gaussian mixture background modeling and statistically analyzes video frames characteristics of each pixel. According to different characteristics of each pixel, the final detection method is selected for the adaptive light mutations.

Gaussian mixture models

The process of Gaussian mixture background modeling [5] is as follows.

(1) Determining whether a pixel matches with a Gaussian mixture model

In a new frame, each pixel is compared with the ith Gaussian model of the pixel. If $|X_t - \mu_{i,t-1}| \leq \lambda\sigma_{i,t-1}$ is established, the definition of X_t is matched with the ith Gaussian distribution in the Gaussian mixture models, otherwise it is not matched, where λ is a constant set by experience (generally $\lambda = 2.5$).

(2) Background Learning and updating

If the pixels of new frame match with the distribution of a Gaussian model, the following steps may be updated Gaussian mixture model parameters.

① Re-correcting the weight ω of Gaussian mixture distribution

Since the weights reflect the appearing probability of the number of pixel values recently, the Gaussian distribution that matches with the new pixel need to appropriately increase its weight, its weight of the matched Gaussian distribution is updated in the t moment:

$$\omega_{i,t} = (1-\alpha)\omega_{i,t-1} + \alpha \tag{1}$$

where α is the update rate to reflect the correction amount of the weight. So it is the smaller, the background component is the more stable.

② Correcting the mean μ and the variance σ^2 of Gaussian distribution matching with new pixels in the t moment

$$\mu_{i,t} = (1-\beta)\mu_{i,t-1} + \beta X_t \tag{2}$$

$$\sigma_{i,t}^2 = (1-\beta)\sigma_{i,t-1}^2 + \beta(X_t - \mu_{i,t-1})^T(X_t - \mu_{i,t-1}) \tag{3}$$

where β determines the update rate of the background. So β is the more larger, the convergence rate of the background component is the more faster. The experiment finds that α is taken a smaller value, and β is taken a relatively larger value, and then the convergence rate can be well improved while keeping the stability of the background model.

When the input pixel value is not matched with a Gaussian model, the current pixel doesn't make any contribution to the model, and therefore we do not need to change the parameters of the model, and only need to change their weight values as shown in Equation 3. After updating in every time, the weight value must be done the normalization processing.

$$\omega_{i,t} = (1-\alpha)\omega_{i,t-1} \tag{4}$$

Video object extraction algorithm

The specific process of object detection algorithm proposed in this paper is as follows.

(1) The initialization of Gaussian model

Change each frame image into the gray image, and establish the Gaussian mixture model in each pixel. Each pixel value of the first frame is as the mean μ, given the variance σ^2 and the weight ω, which is as the first of Gaussian mixture background model. The variance σ^2 should be larger and the weight ω should be smaller.

(2) The matched criteria and parameters update

Each point is compared with various Gaussian distribution corresponding to the Gaussian mixture model. If $|X_t - \mu_{i,t-1}| \leq \lambda\sigma_{i,t-1}$, the point belongs to the background, otherwise it belongs to the foreground. The foreground and the background are made the binarization processing, which obtains a binarized image. Generally β is updated by the following equation:

$$\beta = \alpha + \frac{1-\alpha}{c_k} \tag{5}$$

(3) The detection of light mutation

Detect whether the light takes place a larger change. In the use of Gaussian mixture model detection at the same time, the statistical number of the foreground pixel is more than 50% of the

entire image, and then the scene light has the larger change and the detection of Gaussian mixture model is not valid. Reinitializing the Gaussian model parameters, the use of frame-difference method is to detect moving objects before n frames. Since the frame-difference method is the one based on detecting adjacent frames, it is not sensitive to light change. Therefore, this method is effective to detect moving object in the light change.

(4) Denoising and filling voids of the motion area

After the detecting results of moving region may have some noises and voids, it will affect the accurate determination of regional connectivity, so the detected foreground region needs to carry through morphological processing [6] for distinguishing between the background and the object, and ultimately moving object is correctly segmented.

Experimental results and analysis

In order to verify the effectiveness of the algorithm, we use an actual 130-frames outdoor pedestrian video clips to carry through, camera acquisition speed is 24 frames / sec, with a resolution of 320×240, in the experiments of this algorithm by Matlab7.0 testing. In the beginning of the 52 frames, due to the statistical number of foreground pixels is more than 50% of the pixels in the entire image, it indicates a mutation at this scene illumination, and frame-difference method is for testing, which the threshold is taken 12. After the 78 frames, the statistical number of foreground pixels does not exceed 50% of the pixels in the entire image, the improved Gaussian mixture model is used to detect the foreground objects, K is taken 3 in the Gaussian mixture model. The results of our algorithm are shown in Fig.1: the original video images are in the left side respectively for the 70th frame, the 90th frame and the 110th frame, and the results of this proposed algorithm are in the middle, and the results of the algorithm of GMM are in the right side. Seen from the experimental results, the proposed method can suppress noise while effectively detecting moving objects.
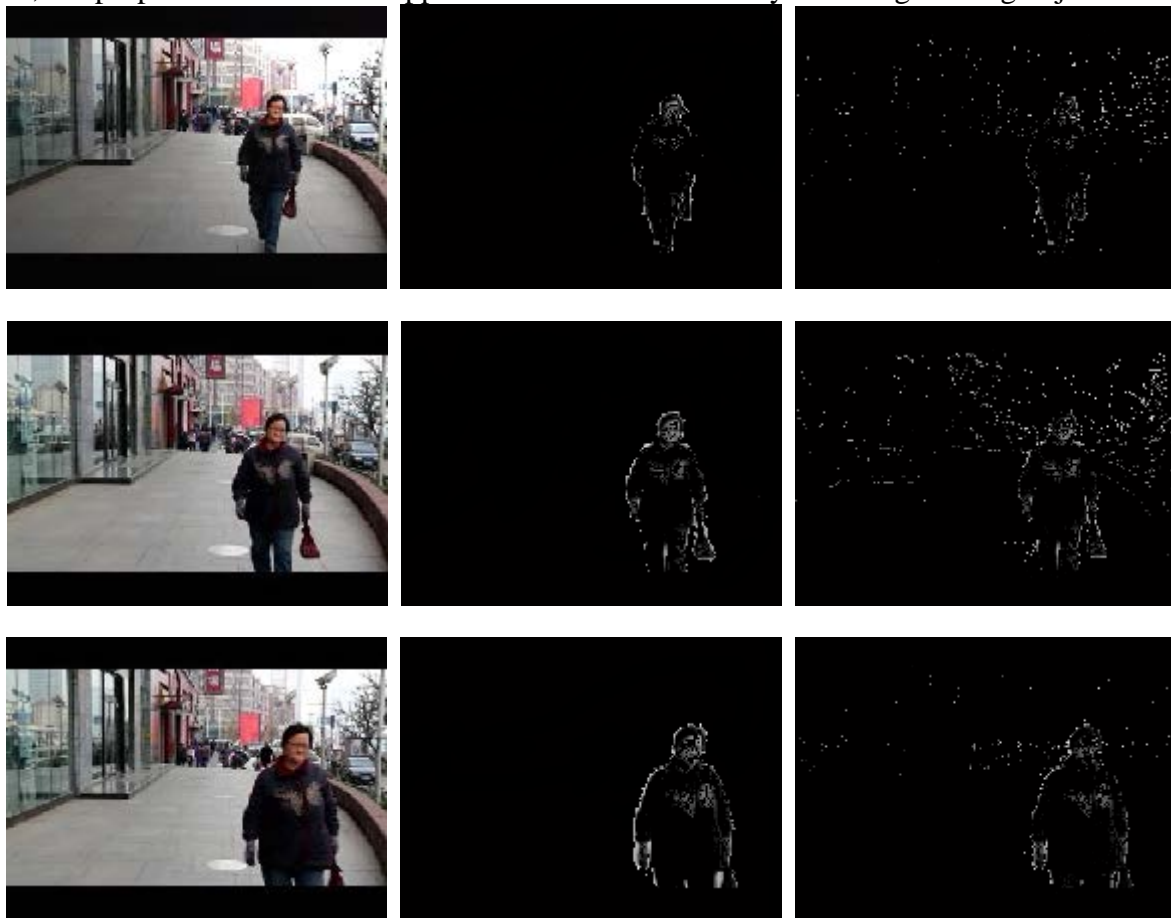

Fig.1. Video object detection results: the original video images are in the left side respectively, and the results of this proposed algorithm are in the middle, and the results of the algorithm of GMM are in the right side.

Conclusion

The vast majority of visual features are affected illumination changes, such as changes in the weather, indoor lights turned on or off. Adjust the brightness of the camera will result in large changes of light, and result in a significant change in the image, and then the extracted movement area and moving targets does not correspond, which can often show large areas of false alarms. Robust moving object detection algorithm should be in response to light changes.

This paper proposes an object detection algorithm based on the video stream for complex environment. The pixel value of the image is as a combination of foreground Gaussian distribution and background Gaussian distribution, and the use of improved learning method of mixture model makes that the model can promptly follow pixel value's change and improves the convergence capability. The statistical number of foreground pixels of the current frame determines the number of pixels of light mutations and detects moving objects combined with frame-difference method. Simulation results show that this method can accurately detect foreground moving targets with more flexibility.

Acknowledgement

References

[1] Coifman B, Beymer D, Mclauchlan P, Malik J: A real-time computer vision system for vehicle tracking and traffic surveillance, Transportation Research Part C, Vol. 6, issue 4, pp. 271-288. (1998)

[2] Magee D: Tracking multiple vehicles using foreground, background and motion models, Image and Vision Computing, Vol. 22, issue 2, pp. 143-155. (2004)

[3] McKenna S et al: Tracking groups of people, Computer Vision and Image Understanding, Vol. 80, issue 1, pp. 42-56. (2000)

[4] Koller D, Daniilidis K, Nagel H: Model-based object tracking in monocular image sequences of road traffic scene, International Journal of Computer Vision, Vol. 10, issue 3, pp. 257-281. (1993)

[5] Stauffer C, Grimson W: Learning patterns of activity using real-time tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, issue 8, pp.747-757. (2000)

[6] Johan Van Horebeek, Ernesto Tapia-Rodriguez: The approximation of a morphological opening and closing in the presence of noise, Signal Processing, Vol. 81, issue 9, pp.1991-1995. (2001)