2011 International Conference on Signal, Image Processing and Applications
With workshop of ICEEA 2011
IPCSIT vol.21 (2011) © (2011) IACSIT Press, Singapore

Moving Objects Detection Algorithm Based on Visual Attention Mechanism and Gaussian Mixture Model

Xuhui Chen and Shaoke Zhu

School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050

Abstract. To deal with the disadvantage of noises and the problem that moving objects were replaced with background when the objects stay a long time in the scene, this paper introduced visual attention mechanism to improve the algorithm performance. Firstly, space-time dynamic hybrid algorithm was adopted to deal with the moving images detected by Gaussian mixture model and the saliency map interested by visual attention mechanism. Secondly, logical "and" operation was done between the significant moving objects of space-time dynamic hybrid mechanism and the saliency map region of visual attention mechanism to remove the isolate noises produced by Gaussian mixture model. Experimental simulations showed that the improved algorithm overcome the shortcoming effectively that moving objects were updated as background when the objects stay for a long time, and eliminated many isolate noises significantly.

Keywords: moving objects detection, gaussian mixture model, visual attention mechanism

1. Introduction

The common approach for discriminating moving objects, which is based on images captured by camera, is to create a mapping relationship between image and image descript, so that computers can understand the contents from video screen by digital image processing and analyzing. It plays an important role in the research fields of objects tracking, incident detection and analyzing of objects behaviour, and it aroused the interest of relative researchers from the world in recent years [1].

The optical flow method, frame difference method and background difference method are the three frequently-used methods for discriminating moving objects. The real time and the practicability of the optical flow method [2] is range. In addition, the optical flow method requires a high hardware and its computational complexity is bigger than other common approaches. The frame difference method is robust on real-time of discriminating moving objects and its computational complexity is smaller. However, the method can hardly discriminate the pixels with unconspicuous changes [3]. The background method [4] is most frequently used in discriminating moving objects. And the most representative method is Gaussian mixture model [5].

To deal with the disadvantages discussed in discriminating moving objects, this paper proposed visual attention mechanism [6] to improve the algorithm performance. Firstly, space-time dynamic hybrid algorithm, which is based on the ideology of movement is preferred, is adopted to deal with the moving objects images detected by Gaussian mixture model and the saliency map interested by visual attention mechanism to solve the problem of leading to objects lost for the reason of staying a long time. Secondly, Do logical "and" operation between the significant moving objects of space-time dynamic hybrid mechanism and the saliency map region of visual attention mechanism to remove the isolate noises produced by Gaussian mixture model.

2. The Model of Gaussian Mixture

The Gaussian mixture model is proposed by Stauffer [7]. It is one of the most concerned methods solving the problem of discriminating objects of dynamic scene. The ideology of Gaussian mixture model is that the algorithm will model multi-Gaussian mixture model for each pixel of the image. At time t , the value set of pixel (x_0, y_0) is $\{x_1, x_2, \cdots, x_t\} = \{(x_0, y_0, i) \mid t_0 \le i \le t\}$. We chose to model the recent history of each pixel as a mixture of K Gaussian distributions. The probability of observing the current pixel value is

$$p(X_t) = \sum_{i=1}^{K} w_{i,t} \times \eta_i(X_t, \mu_{i,t}, \Sigma_{i,t}) \tag{1}$$

Where, $\eta_i(X_t, \mu_{i,t}, \sum_{i,t})$ is the i th Gaussian distribution, and it was defined as follows:

$$\eta_i(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T \Sigma \sum_{i,t}^{-1} (X_t - \mu_{i,t})} \tag{2}$$

And K is the number of distributions of the Gaussian mixture model, X_t is the pixel value at time t. $w_{i,t}$ is an estimate of the weight of the i th Gaussian in the mixture at time . $\mu_{i,t}$ is the mean value of the i th Gaussian in the mixture at time t. $\Sigma_{i,t}$ is the covariance matrix of the i th Gaussian in the mixture at time t, and it was defined as $\Sigma_{i,t} = \sigma_K^2 I$. The prior weights are adjusted as follows:

$$w_{i,t} = (1-a)w_{i,t-1} + a \tag{3}$$

The mean value and the covariance matrix are updated as equation (4) and (5), respectively.

$$\mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho X_t \tag{4}$$

$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(X_t - \mu_{i,t-1})^T (X_t - \mu_{i,t-1}) \tag{5}$$

Where, $\rho = a \cdot \eta(X_t \mid \mu_{i,t-1}, \sigma_{i,t-1})$ is the learning factor for adapting current distributions. a is the constant iterating rate distributed on [0, 1], which stands for refresh rate of the background.

3. The Model of Gaussian Mixture

3.1. The introduction of visual attention mechanism

The characteristic that brain responds to outside important information and control it is called visual attention mechanism. Visual psychology researches show that when the brain analyze complex input scene, the human visual system adopts a serial calculation strategy, namely selective attention mechanism. Selective attention mechanism, according to the local characteristics of image, choose a particular area of the scene through rapid eyes' movement scanning, move to the region of high resolution central retinal sag area, realize more elaborate observation and analysis to the area's attention. Currently one of the most representative visual attention mechanisms is ITTI model [8], shown in Fig.1.

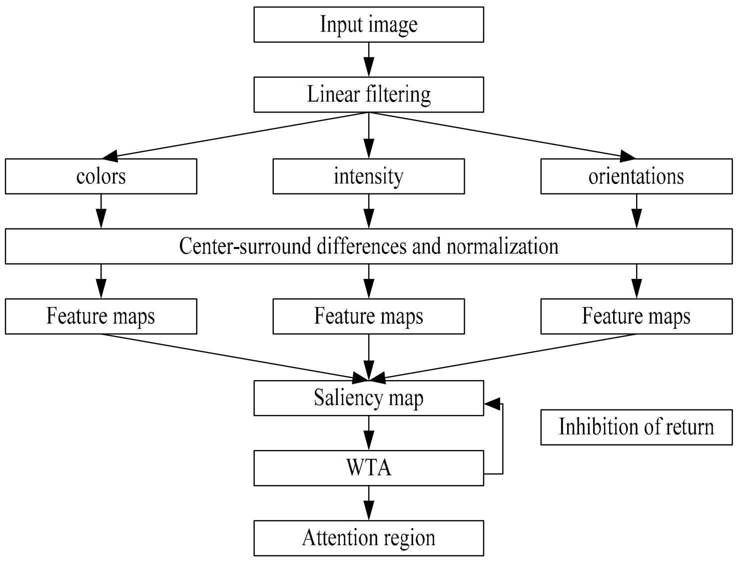

Fig. 1: Visual attention mechanism framework(ITTI model).

The model was proposed by Laurent Itti and Christof Koch (ITTI model).Visual features such as color, intensity and orientation are computed in a massively parallel manner, with a set of pre-attentive feature maps based on retinal input. Activity from all the feature maps is combined at each location, giving rise to activity in the topographic saliency map. The winner-take-all (WTA) network detects the most salient location and directs attention towards it, such that only features from this location reach a more central representation for further analysis. An inhibition-of-return mechanism transiently suppresses this location in the saliency map, such that attention is autonomously directed to the next most salient image location. Finally, model generates the visual attention region.

3.2. Space-time dynamic hybrid algorithm

By the principle of GMM, it is known that moving objects will be replaced by background finally when the moving objects stop and stay in the scene for some time. Therefore, only by modifying the value of α, the GMM algorithm can't essentially solve the problem.

To deal with this problem, this paper runs the space-time dynamic hybrid algorithm to form an overall attention map in a motion priority fashion. Psychology research indicates that the movement significance has a rapid increase first along with the increase of contrast of movement and then the process becomes gently. Therefore, the movement contrast is used to decide weight figure of the results from both Gaussian mixture model in time domain and visual attention mechanism in space domain.

Movement contrast is defined as follows:

$$O_T = \frac{1}{O}\sum_{x=1}^{O}\sum_{y=1}^{B}|\theta_x - \theta_y| \tag{6}$$

Where, O is the feature point number of moving objects, B is the point number of background pixel, θ_x is the direction vector of the x th moving feature point, θ_y is the direction vector of the y th background pixel point. Decided by movement contrast, the weight figure of the results from both Gaussian mixture model in time domain and visual attention mechanism in space domain are defined as follows:

$$W_T = \frac{O_T}{(O_T + \sigma)}, \quad W_S = \frac{\sigma}{(O_T + \sigma)} \tag{7}$$

Where, σ is a constant parameter, used to adjust the sensitivity of the space-time dynamic hybrid model. Here σ value is 0.4. W_T is the weight figure of the moving objects result from Gaussian mixture model in time domain, W_S is the weight figure of the saliency map result from visual attention mechanism in space domain. When movement contrast O_T strengthen, the weight figure W_T of moving objects result from Gaussian mixture model increases rapidly, and the weight figure W_S of saliency map result from visual attention mechanism decreases rapidly.

The formula of space-time dynamic hybrid algorithm to form an overall attention map in a motion priority fashion is defined as follows:

$$D_{(t)}(x,y) = \sum_{x,y\in F}\{W_T * f_T(x,y) + W_S * f_S(x,y)\} \tag{8}$$

Where $f_T(x,y)$ is the pixel point (x,y) luminance value of moving objects result from Gaussian mixture model in time domain, $f_S(x,y)$ is the pixel point (x,y) luminance value of saliency map result from visual attention mechanism in space domain, F is all the pixels of entire area for the frame image.

3.3. Eliminate isolate noises

Gaussian mixture model models for isolate points, therefore, it will cause isolate noises. Thus, Visual attention mechanism can be used to eliminate isolate noises of saliency moving images by calculating the saliency map region. The result image from space-time dynamic hybrid algorithm is $D_{(t)}(x,y)$ .The saliency map region of visual attention mechanism is $D_{(s)}(x,y)$. $I_k(x,y)$ is significant moving object image, which is obtained from the logical "and" intersection operation between $D_{(t)}(x,y)$ and $D_{(s)}(x,y)$.

$$I_k(x,y) = \begin{cases} 1; & if D_{(t)}(x,y) \cap D_{(s)}(x,y) = 1 \\ 0; & otherwise \end{cases} \tag{9}$$

3.4. The model of improved algorithm

Initialize the background learning model of the Gaussian mixture model first. Second, a new frame image is to be calculated and analyzed by both GMM and ITTI in a parallel way. Moving object images are produced by GMM through matching the template, and overall attention maps are produced by ITTI using space-time dynamic hybrid algorithm. Finally, noises of significant moving objects are eliminated using saliency map region from saliency map. The flow chart of the improved algorithm is shown in Fig.2.

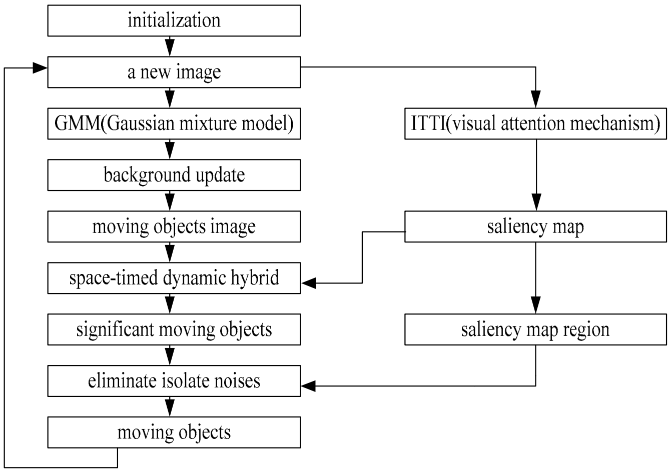

Fig. 2: The improved algorithm flow chart.

4. Simulations and Analysis

In order to verify the algorithm, the experiment use an image and two video to test. And the improved result and the result from Gaussian mixture model are compared to further clarify the advantage of this method.

4.1. Significant moving objects

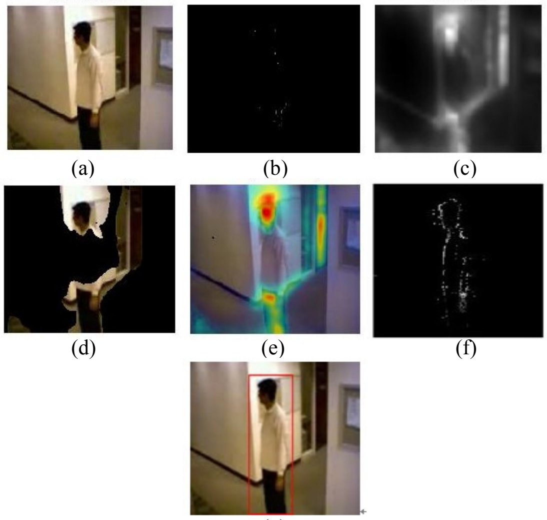

Fig. 3: The process and results of space-time dynamic hybrid algorithm.

In Fig.3, (a) is one frame original image of video 1;(b) is the lost moving objects of result from both Gaussian mixture model in time domain, which residence 3 minutes in the scene and be updated as the background lead to loss; (c~e) is the process of visual attention mechanism algorithm model to this original image ,including Graph-based Visual Saliency(GBVS), image thresholding ,and saliency map overlayed;(f) is the significant moving objects of result from space-time dynamic hybrid algorithm to form an overall attention map in a motion priority fashion;(g) is the detected moving objects at last. (f) and (b) are compared to turn out that space-time dynamic hybrid algorithm can effectively solve the moving objects lost problem due to moving objects stay for a long time.

4.2. Eliminate isolate noises

In Fig.4, (a) is one frame original image of video 2 and (b) is the significant moving objects of result from space-time dynamic hybrid algorithm to form an overall attention map in a motion priority fashion, including a lot of isolate noises. (c~e) is the process of visual attention mechanism algorithm model for this original image, including Graph-based Visual Saliency(GBVS), image thresholding, and saliency map overlay. (f) is the denoise moving objects of result from eliminating isolate noises by saliency map region;(g) is the detected moving objects at last. (f) and (b) are compared to turn out that saliency map region can effectively eliminate isolate noises due to no regional idea of Gaussian mixture model.

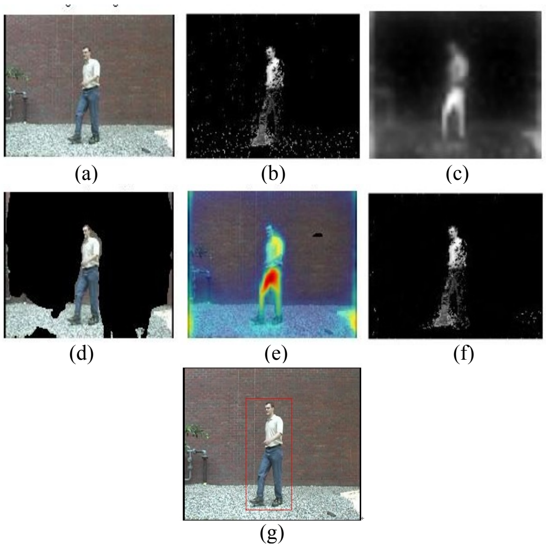

Fig. 4: The process and result of eliminating isolate noises.

5. Conclusions

To solve the problems of Gaussian mixture model, this paper introduced the regional space ideology of visual attention mechanism to improve the performance of the Gaussian mixture model. Space-time dynamic hybrid algorithm, which is based on the ideology of movement is preferred, was adopted to deal with the moving objects images detected by Gaussian mixture model and the saliency map that is interested by visual attention mechanism. Logical "and" operation was done between the significant moving objects of space-time dynamic hybrid mechanism and the saliency map region of visual attention mechanism. Typical simulations showed that the improved algorithm performs well in discriminating moving objects with solving the problems of leading to objects lost for the reason of staying a long time and the isolated noises produced by Gaussian mixed model.

6. Acknowledgements

The work was partially supported by the National Science Foundation of China (61065007), Foundation of Gansu Education, and the National Science Foundation of Gansu of China (0914ZTB098).

7. References

[1] K-T-P. Pakorn and B. Richard. A real time adaptive visual surveillance system for tracking low-resolution colour objectss in dynamically changing scenes. *Image and Vision Computing*. 2003, 21:913-929.

[2] Fejes S, Davis L S. Detection of Independent Motion Using Directional Motion Estimation[R]. *University of Maryland,Technical Report.*CAR-TR-866, CS 3815, 1997.

[3] Paragios N, Deriche R. Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects[J]. IEEE *Transactions on Pattern Analysis and Machine Intelligence*. 2000,22(3): 266-280.

[4] Kim J B, Kim H J. Efficient Region-based Motion Segmentation for a Video Monitoring System[J]. *Pattern Recognition Letters.* 2003,24(1-3): 113-128.

[5] Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking. *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition.* 1999,2: 246-252.

[6] Sun Y,Fisher R.Object-based visual attention for computer vision[J]. *Atrificial Intelligence.* 2003,146(1):77-123.

[7] Stauffer C, Grimson W E L.Learning patterns of activity using real-time tracking[J].*IEEE Transactions on Pattern Analysis &Machine Intelligence.* 2000, 22(8): 747-757.

[8] Itti L, Koch C, Niebur E. A model of saliency-based visual-attention for rapid scene analysis [J] . *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1998, 20 (11): 1254-1259.