

레터논문-08-13-5-12

Object Detection by Gaussian Mixture Model and Shape Adaptive Bidirectional Block Matching Algorithm

Gooman Park^{a)‡}, Byung-Wan Han^{b)}, Tae-Ki An^{c)} and Kwang-Jeek Lee^{d)}

Abstract

We proposed a method to improve moving object detection capability of Gaussian Mixture Model by suggesting shape adaptive bidirectional block matching algorithm. This method achieves more accurate detection and tracking performance at various motion types such as slow, fast, and bimodal motions than that of Gaussian Mixture Model. Experimental results showed that the proposed method outperformed the conventional methods.

Keywords object tracking, object detection, motion estimation

1. Introduction

1. Backgrounds

There are previous works regarding the Block Matching Algorithm(BMA)in moving object tracking. Gyaourova and et.al^[1] introduced the zero motion bias threshold to separate the background from objects. To avoid the false motion caused by small camera movements, they combined the frame history information. Compared to this method, we applied the shape adaptive bidirectional BMA only to the detected objects.

Laptev and et. al^[2] presented and evaluated a set of motion descriptors that exploit the spatial and the temporal co-

herence of motion measurements between corresponding events in image sequences. The performance of local motion events strongly depends on correct detection. Stauffer and Grimson^[3] proposed an adaptive on-line parametric color model in which the background color of each pixel is modeled as a Gaussian mixture model(GMM). This is one of the most commonly used approaches. They modeled the history of each pixel in the image by K Gaussian density distributions. The GMM method has a hole problem at the interior part of detected object when an object is slow or static, and in a severe case it suffers unwanted partition. In order to solve the problem and improve the detection capability, the shape adaptive bidirectional BMA with frame history is suggested. Firstly, we separated the foreground and the background by GMM method. For the detected objects, additionally, the shape adaptive bidirectional BMA is applied, which gives high accuracy in detection. During the proposed BMA operation, some blocks in the object that were previously classified as background are corrected as object.

The weights of those blocks are adjusted as the object category. The proposed system consists of: (i) object and

a) 서울산업대학교 매체공학과

Dept.of Media Engineering, Seoul National University of Technology

b) 동원대학 컴퓨터애니메이션과

Dept.of Computer Animation, Tongwon College

c) 한국철도기술연구원

Korea Railroad Research Institute

‡ 교신저자 : 박구만(Email: gmpark@snut.ac.kr)

※ This research was supported by a grant(07 Urban Transit Standardization A01) from "the 2nd phase of R&D on the urban transit standardization" funded by Ministry of Land, Transport and Maritime Affairs of Korean government.

background separation, ii) noise reduction with accumulated edge information, (iii) shape adaptive bi-directional BMA with history information, and (iv) updating weights.

2. The property of BMA

BMA is a standard technique for finding similar parts between temporally adjacent frames in video compression algorithms. In compression algorithms the motion vector indicates the block that has smallest difference between current block and reference block. Some examples of the matching criteria based on pixel difference are mean absolute difference(MAD), or mean squared difference(MSD), etc. In our experiments we used the mean absolute difference(MAD) as the matching criterion

$$MAD = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |f_t(i, j) - f_{t+\tau}(i+V_x, j+V_y)| \quad (1)$$

where, M and N are the block size, and f is the pixel value. i and j mean the horizontal and vertical coordinate. V_x and V_y are the displacements from the position i and j . τ indicates the previous or future frame. The matched block having the smallest difference does not always indicate the true part of an object. In order to find the true object motion, the detection should precede the block matching. We used the GMM method for moving object detection.

3. Object Detection by GMM^[3]

A GMM is used to model the pixel color. A pixel in the current frame is checked against the background model by comparing it with every Gaussian in the model until a matching Gaussian is found. If a match is found, the mean and variance of the matched Gaussian is updated, otherwise a new Gaussian with the mean equal to the current pixel color and some initial variance is introduced into the mixture. Each pixel is classified based on whether the matched distribution represents the background process.

GMM well detects the new entries and exits of objects but gradually loses it when they are slow or static.

The recent history of each pixel, $\{X_t, \dots, X_1\}$, is modeled by a mixture of K Gaussian distributions. The probability of observing the current pixel value is

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2)$$

where K is the number of distributions, $\omega_{i,j}$ is an estimate of the weight of the i th Gaussian in the mixture at time t , $\mu_{i,t}$ 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 η is a Gaussian probability density function.

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \quad (3)$$

If none of the K distributions match the current pixel value, the least probable distribution is replaced with a distribution with the current value as its mean value, an initially high variance, and low prior weight. The prior weights of the K distributions at time t , $\omega_{k,t}$ are adjusted as follows

$$\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \quad (4)$$

where α is the learning rate and $M_{k,t}$ is 1 for the model which matched and 0 for the remaining models. The mean and variance parameters of the matched distribution are updated, but unmatched distributions remain the same.

II. Proposed System

1. Shape adaptive bidirectional BMA

The regular or deforming motion types are detected by GMM, and the additional accuracy for object's motion is found by shape adaptive bidirectional BMA(SAB-BMA).

If a block is at the object's boundary, it takes the contour shape rather than square. When objects are moving slowly, the foreground detected by GMM includes holes in its interior part, which causes unwanted partition and it is regarded as part of background. Our block matching helps to decide the inner area as part of object. After object detection at time t , each object at time $t-1$ is partitioned into blocks. Fig.1 shows the block patterns on detected object.



그림 1. 블록모양(정방향 블록내에서 흰 부분만 블록정합에 사용)
Fig.1 Block shape(white color in a square block)

For each block, the SAB-BMA searches the best matching blocks in the previous and future frame, i.e., at time $t-2$ and at time t . The SAB-BMA finds additional blocks that are not detected by GMM. When the motion vector is not zero, the matched block is regarded as part of an object in each direction. If the motion vector is zero, the decisions in the history buffer of the several previous frames are examined. If at least one of them was the object,

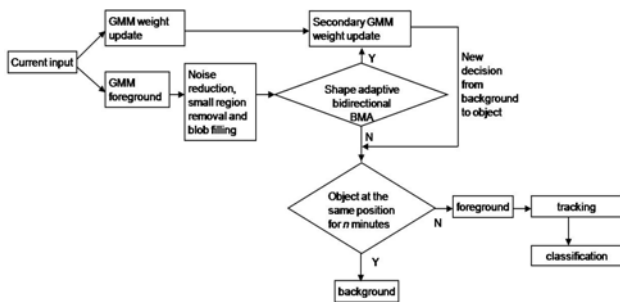


그림 2. 제안하는 객체검출시스템의 블록도
Fig. 2 Block diagram of proposed object detection system
then the block is also decided as such. If a new decision

is made, the newly updated weight is assigned to the collocated pixels at time t or at time $t-2$. If a new decision is made, the newly updated weight is assigned on the pixels in that block at time t or at time $t-2$, as in (4). As a result, the detection capability of GMM would increase at future frames after time t by help of added blocks. The proposed SAB-BMA is shown in Fig.2.

2. Noise Reduction

Before applying the proposed BMA, noises and small blobs are removed in foreground image. These are located mostly near the edges in the frame. Examples of meaningless blobs are stripes, rods or unlikely shapes compared to meaningful ones. Using the accumulated edge information, they are removed. At every frame, the edge component is obtained by edge operator and then the accumulated edge information is obtained by non-recursive averaging filter. During the removing process, however, the contour lines also tend to be lost. The proposed method maintains the contours.

III. Simulation and Analysis

We evaluated our method using luminance image and the size is 352x288 with the block size of 8x8 pixels. People in the 'hall monitor' sequence have various moving speeds such as normal walking, slow motion and shortly static appearance as shown in Fig. 3(a). Because of irregular light-sources, noises are largely distributed in the whole image as shown in Fig. 3(b). As we can see, the results after GMM include noises in large areas. As shown in Fig. 3(c), most noises and small blobs are removed by the noise removing process, but some parts of the contour were also deleted as the undesirable effect. As in Fig. 3(d), the proposed method shows the good object shape by help of SAB-BMA. The inner part of an object is well estimated.

The second sequence contains the repetitive motion in the tree caused by swaying branches and it includes a moving

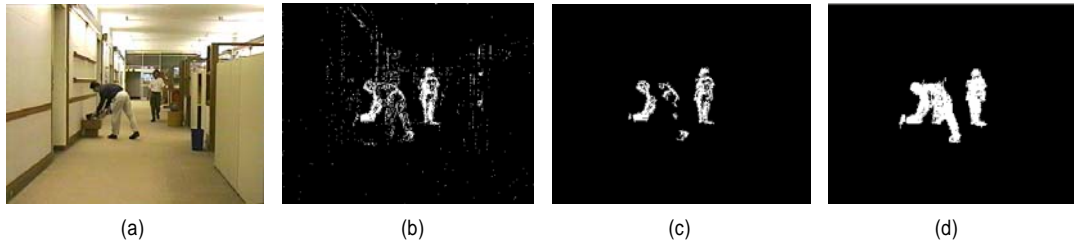


그림 3. 느린 움직임을 보이는 ‘홀 모니터’ 시퀀스

Fig. 3 Hall monitor sequence showing slow motion (a) Hall monitor sequence, (b) GMM result, (c) After removing noise and meaningless small regions on GMM result, and (d) Shape adaptive bidirectional BMA based on GMM

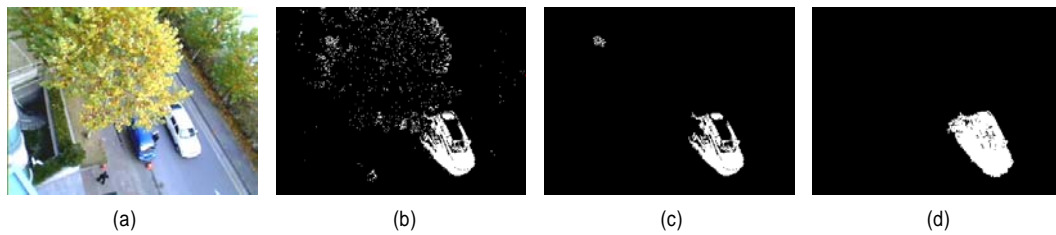


그림 4. 배경에 바이모달 움직임이 있는 ‘나무와 자동차’ 시퀀스

Fig.4 Tree and vehicle sequence showing the bimodal motion in background. (a) Street with a big tree and a moving car, (b) GMM result, (c) After removing noise and meaningless small regions, and (d) Shape adaptive bidirectional BMA based on GMM

vehicle under the tree as in Fig4.(a). Even though GMM method is good at bi-modal motion, the result still contains the swaying motion as in Fig4.(b). But the proposed method shows only the vehicle motion as shown in Fig. 4.(d).

For 300 frames in the ‘hall monitor’ sequence, we counted the case of contour line lost in detecting whole body of objects. Table1 shows that the proposed method outperforms the GMM. In order to count the case, we used the well known ‘hall monitor’ sequence to have objective assessment as in Table 1.

표 1. 300 화면동안 테두리선이 끊어진 횟수

Table1. Number of Contour line lost in 300 frames

Method	Number of contour line lost
Proposed	26
GMM+Noise reduction	208
GMM	198

Compare to the detection by GMM, the proposed method takes more operation time due to noise reduction, contour filling and SAB-BMA. We expect that future DSP chips overcome the computing delay and achieve real time processing.

IV. Conclusion

The shape adaptive bidirectional BMA based on GMM showed good detection performance in fast and slow motions as well as repetitive meaningless motions. The average value of motion vectors belonging to one object provides the information such as speed and direction. By preventing the unwanted segmentation in an object, the ratio of dividing one object into two or more has reduced and correct tracking capability has increased in successive frames.

References

- [1] A.Gyaourova, C.kamath, and S.-C. Cheung, "Block Matching for Object Tracking," Technical Report, Lawrence Livermore National Laboratory, UCRL-TR-200271, Oct.14, 2003.
- [2] Ivan Laptev, et. al., "Local velocity-adapted motion events for spatio-temporal recognition," Computer Vision and Image Understanding, vol.108, No.3,108, pp.207-229, Dec.,2007.
- [3] C. Stauffer, W.E.L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking," Proc.IEEE Conference on Computer Vision & Pattern Recognition, vol.2, pp.246-252, June, 1999.