AUTOMATIC SEGMENTATION OF MOVING OBJECT AND BACKGROUND *

ZHAN Jing-Feng QI Fei-Hu ZHAO Xue-Chun

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract A moving object detection method in video sequence was proposed, which is considered a fixed camera model. Firstly, a new method was presented for detection of changes of motion. After the change detection, the recursive higher-order mothod was used to extract the change from the Gaussian noise. Compared with the classical higher-order statistics method, the recursive higher-order statistics method uses the former n images information, so it can obtain better estimates of HOS to reduce the effect of additive noise and detect the small object and perform well on segmetation.

Key words video sequence segmentation, recursive higher-order statistics, morphology filtering.

Introduction

Detecting and segmenting moving objects in a static scene is an important computer vision task [1]. In object recognition applications, the segmentation of moving objects is the first step for system. Object-based coding of video sequences, which is currently under investigation in MPEG-4, is a new video compression technique used in wireless communication system such as mobile computing and very low bit-rate telecommunication application such as public-switched telephone network (PSTN). It is well known that the present block-based video coding methods such as H. 261 and MPEG-2 have block artifacts and mosquito effects in very low bit-rate environments. The reason is that the traditional video standards such as H. 261 and MPEG-2 are low-level techniques in the sense that no segmentation or analysis of the scene is required [2]. Object-based coding method partitions video sequence into moving object and still background. It can achieve high compression ratios. So the segmentation of moving object from background is an important step in object-based coding algorithms. In MPEG-4, the segmentation of moving object and still background is also used for content-based functions such as video editing.

* The project supported by the National Natural Science Foundation of China

Received 1999-05-23, revised 1999-07-02

There are a lot of approaches for the segmentation of moving object from still background. In general, these methods use the absolute differences of subsequent frames to extract the moving information of an object. S. Haddadi and C. Fernandez [3] used a Markovian approach combined with statistic parameters of second order. And Hotter et al. [4] proposed an iterative approach, in which an image is first divided into unchanged and changed areas. Leung et al. [5] proposed a voting scheme to get the regions of interest.

In this paper, we develop a segmentation method [6] for detecting the moving object in still background. Different from the former methods which only use the displaced frame difference (DFD) as the change detection, we use the discrete temporal derivation $G_T(x, y, t)$ of the discrete gradient magnitude $G(x, y, t)$ of image to detect the changed areas of image. Because the moving object's edge is stressed, the change of moving object is more obvious than the still background. After change detection, we use the recursive high-order method to extract the non-Gaussian structured inter-frame variations from the Gaussian noise. Because of using the former n images information, it can be used to obtain better estimates of HOS to reduce the effect of additive noise and detect the small object and perform well on segmentation.

1 Principle of the Segmentation Algorithm

Figure 1 gives an overview of the principle of the segmentation algorithms. The proposed segmentation method can be subdivided into the following steps:

First, the change areas of two different frames are detected. Without using the former method to get the difference of the different frame, we get the discrete temporal derivation $G_T(x, y, t)$ of the discrete gradient magnitude $G(x, y, t)$ of image (DTD-DGM) to detect the changed areas of image. Within textured image regions and along object contours, the discrete spatial derivatives $B_x(x, y)$ and $B_Y(x, y)$, as well as higher order spatial derivatives and gradient operators, approximately show an exponential or two-sided exponential (Laplacian) distribution [1].

Second, we use the recursive HOS method to extract the motion of the image sequence from the still background. The higher order method has a good attribution in extraction of the non-Gaussian signal against the Gaussian noise. During the HOS method, we decide the local threshold by the feature of the displaced frame difference (DFD) and use it to decide the automatic segmentation threshold.

In the third step, we use morphology filter to smooth the shape of mask and eliminate the small isolated areas.

2 Change Detection

The previous work about the segmentation of moving object is mostly based on the displaced frame difference techniques. The displaced frame difference techniques are to compute the pixel-by-pixel absolute difference of the two different frames. The pixel locations differing from zero indicate "change" regions. However, because of the presence of observation noise, the frame difference hardly ever becomes exactly zero. In this paper, we use a new method to get the change areas of the two differnet frames. From Ref. [1], we know that within textured image regions and along object contours, the discrete spatial derivatives $B_x(x,y)$ and $B_Y(x,y)$, as well as higher order spatial derivatives and gradient operators, approximately show an exponential or two-sided exponential (Laplacian) distribution. Especially along the contours of moving objects, the discrete temporal derivative of the discrete gradient magnitude of the image is expected to approximately show a Laplacian distribution in moving textured regions.

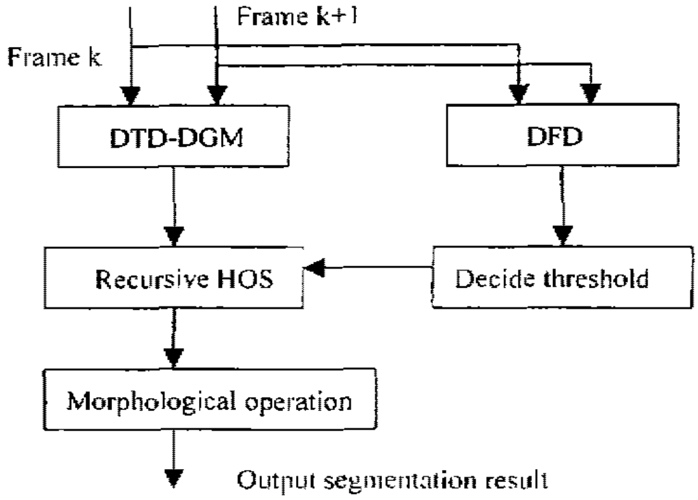

Fig. 1 The principle of the segmentation algorithms
图 1　分割算法基本原理

2. 1　Probability distribution of image derivatives

As we show in the front of this section, the discrete spatial derivatives of image approximately show an exponential or two-sided exponential (Laplacian) distribution

$$P\{O(B)(x,y)\} \approx \frac{1}{2\lambda}\exp(-\frac{|O(B)(x,y)|}{\lambda}) \qquad (1)$$

where $B(x,y)$ is the local image brightness at location (x,y), O is a spatial derivative operator and λ is an experimental constant. By testing a number of images, we can confirm this qualitative law. Figure 2 illustrates an example of real image from MPEG-4 test sequences. The Fig. 2 shows an exponential distribution.

2. 2　Motion detection

The object of motion detection is to distinguish temporal variations caused by object moving from noise and background. Here, we use the difference of the discrete gradient magnitude for motion detection. The advantage of the method is that the motion of the edge of moving object is stressed. However, the textured moving object points, especially along the moving contours show a Laplacian distribution. It can be extracted by HOS against Gaussian noise.

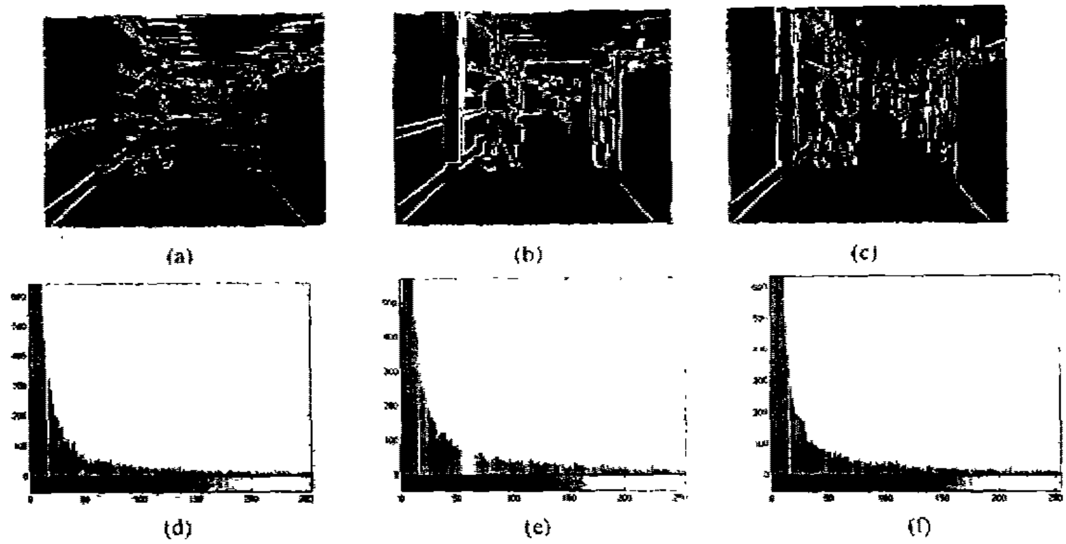

Fig. 2 (a) B_x using Sobel operation, (b) B_y using Sobel operation, (c) G using Sobel operatoin, (d) histogram of B_x, (e) histogram of B_y, (f) histogram of G

图 2 (a)使用 Sobel 算子得到的 B_x，(b)使用 Sobel 算子得到的 B_y，(c)使用 Sobel 算子得到的 G，(d) B_x的直方图，(e) B_y的直方图，(f) G 的直方图

We define the difference of the discrete gradient magnitude of the two frames as:

$$G_T(x,y) = G(x-\Delta x, y-\Delta y) - G(x,y) = -\Delta x \cdot \frac{d}{dx}G(x,y) - \Delta y \cdot \frac{d}{dy}G(x,y). \tag{2}$$

G(x,y) is the discrete gradient magnitude of the image. $G(x-\Delta x, y-\Delta y)$ is the discrete gradient magnitude of the next frame. In fact, $G_T(x,y)$ is also the discrete temporal derivative of the discrete gradient magnitude of the two frames. As we indicated in the front of the section, $G_T(x,y)$ is expected to approximately show a Laplacian distribution in moving textured regions and, especially, along contours c of moving objects:

$$E\{G_T(x,y)|c\} \approx \frac{1}{2\lambda}\exp(-\frac{G_T(x,y)}{\lambda}). \tag{3}$$

where λ is an experimental constant.

3 Recursive Higher-order Statistics

HOS-based methods already began to be used to estimate the motion in the presence of noise [7]. These include motion estimation and moving object segmentation. Higher-order cumulants of non-Gaussian processes are estimated in the presence of unknown deterministic and/or Gaussian signal. And noise can be realistically described as a colored Gaussian process. So, by HOS-based method we can estimate the motion of moving object in image sequence, which is severely corrupted by additive Gaussian noise. In practical situations, the fourth-order cumulant is used in the process of a symmetric probability density function. In particular, the fourth-order cumulant has a good Gaussian noise rejection capability.

The previous work used a traditional estimator of the fourth-order cumulants for moving object segmentation. But Elisa Sayrol showed in Ref. [8] that image information is repeated along the sequence. This redundancy may be used to obtain better estimates of HOS to reduce the effect of additive noise. Amblared indicated in [9] that the recursive higher-order statistics can overcome the disadvantage of the traditional estimator of the fourth-order cumulants.

3.1 Recursive higher-order statistics

The recursive higher-order statistics for moving object segmentation is derived as follows:

$$J_{43k} = \frac{N \cdot \sum_{m\in\Omega_m} G_T^4}{\sum_{m\in\Omega_m} G_T^2} - 3 \cdot \hat{E}_{k-1}\{G_T^2\} \tag{4}$$

where

$$\hat{E}_{k-1}\{G_T^2\} = \hat{E}_{k-1}\{G_T^2\} + \mu[\frac{1}{N}\sum_{m\in\Omega_m} G_T^2 - \hat{E}_{k-1}\{G_T^2\}] \tag{5}$$

where Ω_m denotes the spatial domain that contains the pixels from a region. N is the number of such pixels, e.g. for Ω_m 3 * 3 window, N = 9. The coefficient μ is decided by the change of scene. If the scene in the image sequences changes rapidly, the coefficient μ is chosen close to one, and close to zero otherwise.

The square $\hat{\sigma}_{0d}^2$ of the variance of noise is estimated based on N_s, which is corresponding to static background η:

$$\hat{\sigma}_{0d}^2 = \frac{1}{N_s}\sum_{m\in_n} G_j, \qquad \hat{m}_d = \frac{1}{N}\sum_{m\in\Omega_m} G_T.$$

where N_s is the number of the static background pixels of m windows.

When we estimate the first two frames of the sequence or we must reset the computing, we can't get the fourth-order recursive cumulants by the function (4). So we define the fourth-order recursive cumulants of single frame as follows:

$$J_{43k} = \frac{\sum_{m\in\Omega_m} G_T^4}{\sum_{m\in\Omega_m} G_T^2} - 3 \cdot \frac{1}{N} \cdot \sum G_T. \qquad \hat{E}_0\{G_T^2\} = \frac{1}{N}\sum_{m\in\Omega_m} G_T^2.$$

3.2 Moving object segmentation

Due to the non-Gaussian signal discrimination capabilities of the HOS, we use the recursive higher-order statistics to detect non-Gaussian structured inter-frame variations, such as those due to edges belonging to moving objects, against completely random components and noise effects. We assume H_0 to be the "still background" class and H_1, the "foreground or covered/discovered background" class, if

$$J_{424} < c \cdot (\hat{\sigma}_{0d}^2)^2 \in H_0, \quad J_{424} > c \cdot (\hat{\sigma}_{0d}^2)^2 \in H_1$$

where c is an experimental constant.

3.3 Decision of the local threshold

To automatically segment the moving object in static background, we must estimate the square $\hat{\sigma}_{0d}^2$ of the variance of noise on the static background η. By analyzing the shape of the difference image histogram, we give an adaptive method to decide the static background η [5]. The separation between the stationary regions and the moving regions occurs at the valley point with the largest slope change. Figure 3 illustrates an example of displaced frame difference of two frames.

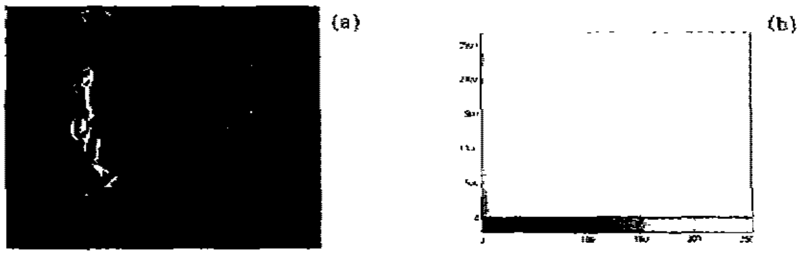

Fig. 3 (a) The displaced frame difference of two frames after thresholding, (b) histogram of the displaced frame difference of two frames
图 3 (a) 阈值化后的帧差图像,(b) 帧差图像的直方图

4 Regulation of Mask

The change detection methods have a major drawback. Unless moving objects contain sufficient texture, only occlustion areas are marked as changed while the interior of objects remains unchanged [2]. So we should use morphological operation to overcome the drawback. The method we use is morphological close-opening operation. The operations have two advantages. One is hole filling. The other is shape simplification.

5 Experimental Results

In this section, the results of our segmentation algorithms are given for the three MPEG-4 test sequence, Akiyo, news and hall monitor. Figures 4, 5, 6 illustrate the results.

6 Conclusion

In this paper, we develop a method for segmenting the moving objects from still back-

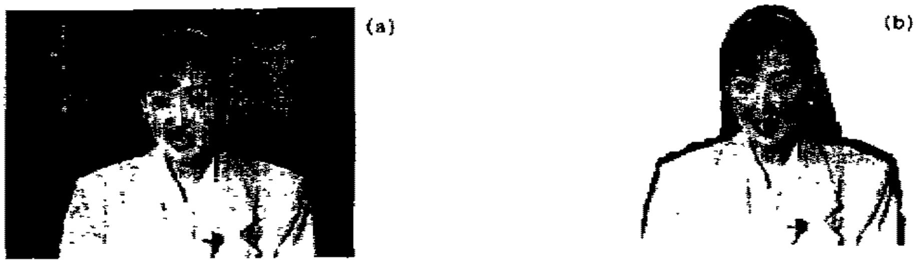

Fig. 4 Segmentation results on Akiyo sequence. (a) original, (b) segmentation result
图 4　Akiyo 图像序列的分割结果，(a) 原始图像，(b) 分割结果

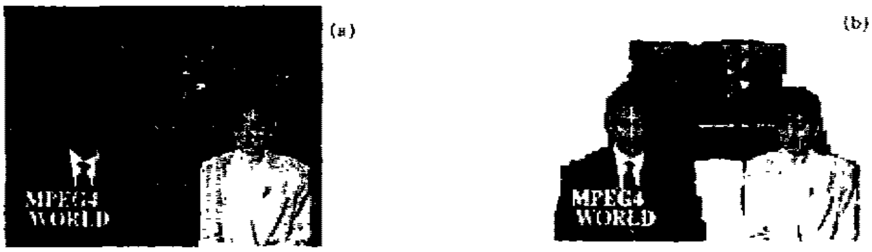

Fig. 5 Segmentation results on news sequence. (a) original, (b) segmentation result
图 5　News 图像序列的分割结果，(a) 原始图像，(b) 分割结果

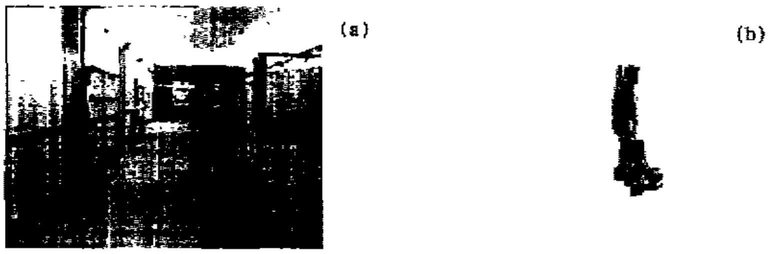

Fig. 6 Segmentation results on hall monitor sequence, (a) original, (b) segmentation result
图 6　Monitor 图像序列的分割结果，(a) 原始图像，(b) 分割结果

ground. The method is based on change detection. After the change detection, the recursive higher-order statistics is performed. At the end of the method, we use a morphological close-opening operation to overcome the drawback of change detection. Further research will be carried out on two topics in the future. First, the camera motion estimation and compensation should be investigated against the fixed camera model. Because in really video sequence, the background is not always still. Second, in recursive higher-order statistics step, the coefficient μ of function (5) is decided by the change of scene. If the scene in the image sequences changes rapidly, coefficient μ is chosen close to one, and close to zero otherwise. So we must design a scene change detector to decide the coefficient μ.

REFERENCES

1 Martin Bichsel. Segmenting simply connected moving objects in a static scene. *IEEE Transaction on PAMI*, 1994, **16**(11): 1138—1142
2 Thomas Meier, King N Ngan. Automatic segmentation of moving objects for video object plane generation, *IEEE Transaction on Circuits and Systems for Video Technology*, 1998, **8**(5): 525—538
3 Haddadi S, Fernandez C. Luminance and texture variation analysis for motion detection. *SPIE*, 1997, **3309**: 1008—1018
4 Michael Hötter, Robert Thomas. Image segmentation based on object oriented mapping parameter estimation, *Signal Processing* 1998, **15**: 315—334
5 Maylor K Leung, Yee-hong Yang. Human body motion segmentation in a complex scene, *Pattern Recognition*, 1987, **20**(1): 55—64
6 Neri A, Colonnese S. Adaptive segmentation of moving object versus background for video coding, *SPIE*, 1997, **3164**: 443—453
7 Sayrol E, Gasull A. Motion estimation using higher order statistics, *IEEE Transaction on Image Processing*, 1996, **5**(6): 1077—1083
8 Sayrol E, Gasull A. Motion estimation using higher order statistics, *IEEE Transaction on Image Processing*, 1996, **5**(6): 1077—1083
9 Amblard P O, Brossier J M. Adaptive estimation of the fourth-order cumulant of white stochastic process, *Signal Processing*, 1995, **42**(1): 37—44

(2)

343-350

运动目标和背景的自动分割*

詹劲峰 戚飞虎 赵雪春

（上海交通大学计算机科学与工程系，上海，200030）

摘要 本文提出一种固定背景下的运动目标自动分割技术．首先，提出了一种新的运动变化检测方案．然后利用递归高阶统计的方法从高斯噪声中提取运动变化．同传统的高阶统计方法相比，递归高阶统计由于利用前 n 帧的图像信息，所以能够更有效地抑制噪声，检测小目标，并且分割效果较好．

关键词 视频序列分割，递归高阶统计，形态滤波．

运动目标

* 国家自然科学基金（编号 69772002）资助项目

稿件收到日期 1999-05-23，修改稿收到日期 1999-07-02