MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

12—4 Moving object dectetion from MPEG coded picture

Shinichi MITSUMOTO, Hiroaki YUASA, Heitou ZEN *
Tokyo University of Mercantile Marine

Abstract

A proposed method extracts moving objects from MPEG bit stream directly, i.e. without video decoding process. Along with rapid progress of multimedia technology, motion picture is increasingly handled on computer these days. Accordingly the MPEG format becomes a worldwide standard, and is accepted in many fields, namely broadcasting, communication, networking and archiving. At the same time, it's so important how to extract the "contents" of a picture, in those fields. Moving objects detection is an important in image contents analysis. We present the proposal for motion analysis by directly manipulating MPEG bit stream. Our method is efficiency and feasibility for many applications such as video surveillance, ITS, video indexing etc. In this paper, we experiment with method: detection and matching for target.

1 Introduction

A motion analysis method based on direct manipulation of MPEG bit stream is proposed in this paper. Along with rapid progress of multimedia technology, motion picture is increasingly handled on computer these days. MPEG came into wide use as a standard for video data compression. And MPEG date is accepted in many fields, namely broadcast, communication, networking and archiving. At the same time, it's so important how to extract the "contents" of a picture, in those fields. Motion analysis plays an important role in many areas such as remote surveillance, video indexing, and so on[1]. In vided indexing, cut detction and text regions detection is used to handle video efficiently[6],[7]. In remote surveillance, moving object detection is neccesary for tracking[3].

Traditionaly, detection moving object is usually accepted by Optical Flow, to pixel information in frame[4].

As above, MPEG data is utilized in various case. We aim at two characteristic MPEG encoding method, i.e. Motion Compensation (MC) and DCT. Those appear the direct information of motion and texture.

We propose a method based on direct manupulation from MPEG bit stream, i.e. MC and DCT,to detection of moving objects. Our method is effecient and feastibillity without video decoding process. It's suitable to be imperimented on such a simple machine as PCs.

In the following sections, we explain the method for detection of moving objects, and the experimental results with some video sequences.

2 Detction moving object

MPEG exploits the both temporal and space redundancies, intrinsic to moving picture. As for the temporal redundancy, the scheme called MC based on block matching algorithm is employed. The motion vectors (MV) are calculated for each MB (Macro Block consisted of two by two blocks, i.e. 16 by 16 pixels). The MPEG motion picture has 3 types of frames by MC. Frame I is intra encoded frame. Frame P is Predictive frame. And frame B is Bidirectionally predictive Frame. There are MV in P and B frame. On the other hands, DCT cofficients appear the information of texture and color in block, consitsted of 8 by 8 pixels. Fig.1 shows MPEG data construction.

The proposed method forms an initial target region by extracting and grouping the MVs included in a MPEG bit stream. Though the motion vectors are not calculated to every block in MPEG, it works well as the segmentation method for a target tracking. As for a tracking, a correspondence of each target between frames are determined by their texture estimated from DCT coefficients that also extracted directly from MPEG code.

2.1 Initial segmentation

It's well known that a static region tend to produce the random MVs in MPEG, whose magnitudes are relatively small. So, elimination of noisy MVs is done firstly. Among the rest, the grouping is done using the MV's similarity. The criteria employed here is based on the similarity of the MV's magnitude and direction.

- Step.1 : When Motion Vector means $|MV| \leq 1$, those are noise or background, removing from moving region
- Step.2 : Calculaing x,y angle from MV eqnation(1)

*Address: 2-1-6 Etchujima, Koto-ku, Tokyo 135-8533 Japan. E-mail: smitu,hyuasa,zen@ipc.tosho-u.ac.jp

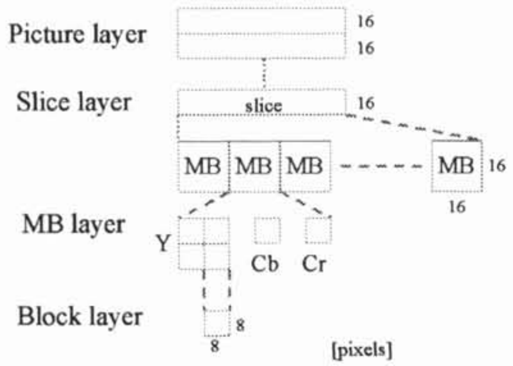

Figure 1: MPEG data constraction

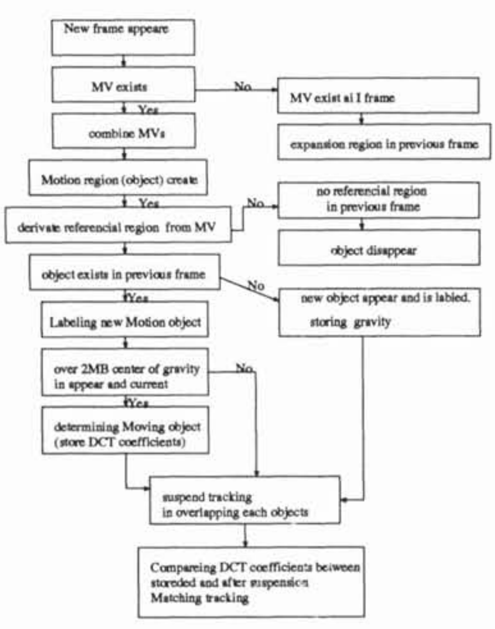

Figure 2: tracking flowchart

$$\theta = 180/\pi \tan^{-1} \frac{Y_{vector}}{X_{vector}} \tag{1}$$

- Step.3 : Comparing direction (from Step.2) within around 3 MB adress, we determine similar direction as same label. (fig3)

$$\theta - T \leq \theta' \leq \theta + T \tag{2}$$

(T is thresholding ,$\theta'$is leading angle)

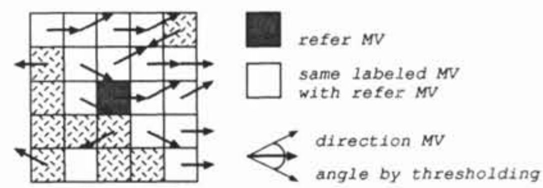

Figure 3: labeling with angle

- Step.4: In case of number of MV ,i.e. same label, within 3, determine those as noise(or background).

2.2 Motion estimation (tracking)

The tracking between adjacent frames is done at this stage. Each target region, a grouped MBs in Step 1, has it's own averaged MV. Correspondences are determined using the similarity of these averaged MV of each target region. If no appropriate region is found, the correspondence is suspended and the target is marked for later process.

2.3 Matching of suspended target

There is some reason why the tracking is suspended, namely occlusion, appearant or physical collision etc. The effort is made for some period to recover the tracking using DCT coefficients included in MPEG stream . The normalized distance predefined in the principal component space formed by available DCT coefficients, is used to matching of suspended target.

2.4 Comparing DCT coffcients

After suspended target divide into same number of targets again, we comper DCT cofficients from both previous and after target MBs. However DCT cofficients of P picture is differencial value from previous P or I picture cofficients, the way to adding difference intra cofficients resolve this problem.

$$\text{contraction}: g_{i,j} = \begin{cases} \text{background} & : \text{MB(i,j) or 8-neighbor MV is in background} \\ \text{moving region} & : \text{others} \end{cases}$$

$$\text{expansion}: g_{i,j} = \begin{cases} \text{moving region} & : \text{MB(i,j) or 8-neighbor MV is in moving region} \\ \text{background} & : \text{others} \end{cases}$$

On the other hands, the case of estmating moving object from MB mean almost edge area, in the inside of those area no cofficients. It is difficalt to matching target, because the parts of edge appers both background and moving object. Using contraction and expansion to MV adress, it is able to complete inside of target. However these proses cause moving object over 3 by 3 to match tracking by using DCT coffcients (consisted of 10 luminance average, 10 luminance standard deviation, and 10 chronominance). Redusing dimesion of DCT cofficients is easy to make cluster.

2.5 Matching of moving object by using DCT cofficients

Comparing cluster between previous and current DCT cofficients, We calculate percetntage of clusters, and we calculate difference percentage between previous and current. In case of the small difference, we judge that there are same object. On the other hands, in case of large difference, we consider those as new object.

3 Experiments

We experimented with some MPEG data to confirm the effctiveness of our method. We use MPEG2 encorded picture by MP@ML format. GOP is constract five frames, in this case No B frame(namely I,P,P,P,P). In any case, there is no camera motion(e.g. the egomotion).

3.1 Tracking pedestrains

Fig.4 shows a tracking results for pedestrians passed each other. Each target is tracked correctly shown at the left row. The right hand pedestrian moving toward left side, passed behind the trunk of a tree in a beginning of the sequence and the tracking is also succeeded.

Fig.5 shows the sequence is changeable brightness between righet side and left. The left side is brighter than the left. And pededstrians pass each other. In this case, tracking is also succeeded.

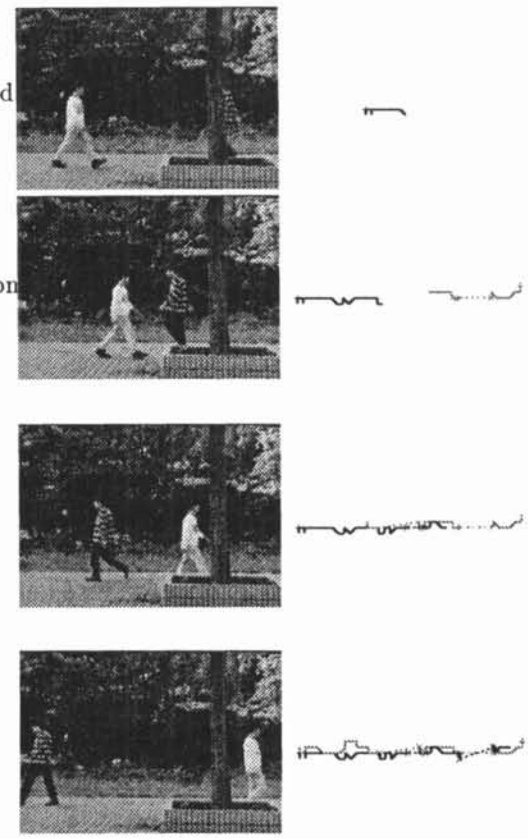

Figure 4: Tracking results (Peds. Passed each other)

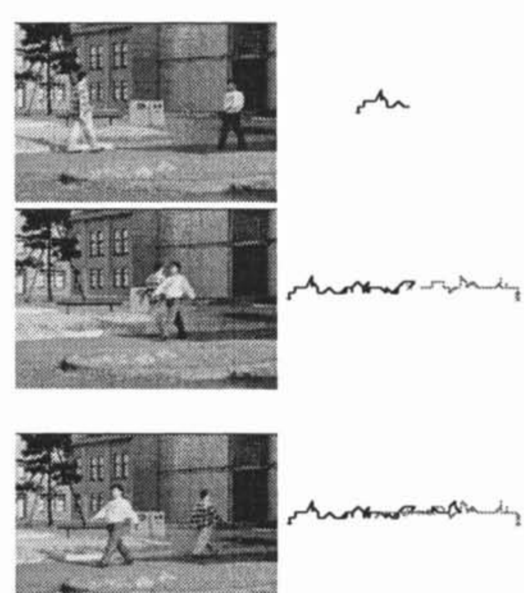

Figure 5: Tracking results (Blightness changeable)

3.2 Tracking Vehicles

Fig.6 shows another result applied to vehicles on a run. In a jammed traffic scene, vehicles are often occluded each other. Accurate vehicle tracking is very helpful at ITS, as for early detection of an accident or AVI (Automatic Vehicle Identification) system. Though some improvements must be considered, the proposed method has potential profitability to ITS environment, that MPEG is launched to remote surveillance.

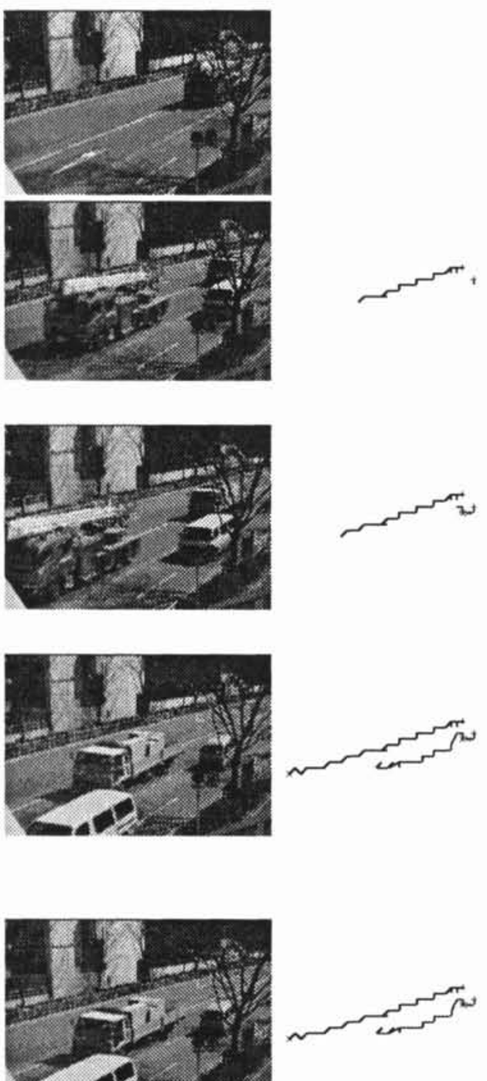

Figure 6: Tracking results (Vehicles)

4 Conclusion

In this paper, we propose method detects moving objects from MPEG bit streem directly. Our approach, using MV and DCT, is effective to match tracking target without video decoding process. However, reducing dimension costs much time, difficalutly to use in case of Remote surveillance. And our method also is not suited to egomotion, in case of video indexing. The images exploited as video indexing, consist not only fixed egomotion, but chageable.

After this, we will cope with these problems as processing in real time and detecting moving objects with egomotion.

References

[1] Michal Irani, and P.Anandan, "A Unified Apporoach to Moving Object Detection in 2D and 3D Scenes,"IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no.6, pp.577-589, Jun. 1998.

[2] Eiichi Iwanari, and Yasuo Ariki, "Scene Clustering and Cut Detection in Motion Images by DCT components,"Technical report of IEICE Vol.93 PRU No.119 1994 (in Japanese).

[3] Shigeki Nagaya, Takafumi Miyake, Takehiro Fujita, Wataru Ito, and Hirotada Ueda,"Moving Object Detction by Time-Correlation-Based Judgement Method,"Proc. IEICE(Institute of Electronics, Information and Communication Engineers), vol.J79-D-II, no.4, pp.568-576, Apr. 1996 (in japanese).

[4] Atsushi Nagai, Yoshinori Kuno, and Yoshiaki Shirai, "Detction of Moving Objects Changing Background," Proc. IEICE(Institute of Electronics, Information and Communication Engineers), vol.J80-D-II, no.5, pp.1086-1095, May. 1996 (in japanese).

[5] Akio Yoneyama, Yasuyuki Nakajima, Hiromasa Yanagihara, and Masaru Sugano,"Moving Object Detction from MPEG Video Steam," Proc. IEICE(Institute of Electronics, Information and Communication Engineers), vol.J81-D-II, no.8, pp.1776-1786, Aug. 1998(in japanese).

[6] Yasuhiro Niikura, Akihiro Akutsu, Yukinobu Taniguchi, Takashi Sattou, Yoshinobu Tonomura, and Hiroshi Hamada, "Shot Change Detction from MPEG Encoded Video Data," Proc. IEICE(Institute of Electronics, Information and Communication Engineers), vol.J81-D-II, no.8, pp.1838-1846, Aug. 1998(in japanese).

[7] Takashi Sattou, Yasuhiro Niikura, Yukinobu Taniguchi, Akihiro Akutsu, Yoshinobu Tonomura, and Hiroshi Hamada, "An Efficient Method for Detecting Text Regions from MPEG Encoded Video," Proc. IEICE(Institute of Electronics, Information and Communication Engineers), vol. J81-D-II, no.8, pp.1847-155, Aug. 1998(in japanese).