MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

13—28

Abrupt Shot Boundary Detection from Video Sequence Using Motion Direction Histogram Feature

Ken-ichirou SUZUKI* Masaomi NAKAJIMA Hitoshi SAKANO Yasuo SAMBE
Sakuichi OHTSUKA
Research and Development Headquarters
NTT Data Corp.

Abstract

We propose the motion direction histogram (MDH) as a new feature for abrupt shot boundary detection in video sequences. Conventional methods for abrupt shot boundary detection use a threshold difference as a measure of the feature values between consecutive images. It is difficult for such conventional methods to detect shot boundaries in "busy" scenes, in which intensities change substantially from frame to frame, because these methods use features such as color or luminance difference histograms and motion vectors. The values of these kinds of features are significantly increased when a video sequence includes such intense motions. To overcome this problem, we have developed the MDH feature. We confirmed its effectiveness experimentally and obtained an accuracy of over 95 % for abrupt shot boundary detection.

1 Introduction

In this paper, we propose a motion direction histogram (MDH) as a new feature for abrupt shot boundary detection in video sequences. We also demonstrate the effectiveness of our method through experimental results.

Shot boundary detection in a video sequence is an important task in constructing indexed video databases, because dividing a video into shots is the first step towards automatic video indexing.

There are two kinds of shot boundaries in a video sequence: the abrupt shot boundary, and the gradual shot boundary. An abrupt shot boundary is the result of splicing two dissimilar shots together, and this transition occurs in a single frame. A gradual shot boundary, which can be attributed to effects such as fade-ins, fade-outs, wipes, and dissolves, occurs over multiple frames. In this paper, we address the abrupt shot boundary, which occupies the main part of a shot boundary. For automatic shot boundary detection, it is thus necessary to detect abrupt shot boundaries with high accuracy.

Conventional methods of abrupt shot boundary detection use a threshold difference between the feature values of successive images. Examples of these feature include the color difference histogram [1, 2], motion vector length [4], and edge difference [5, 6]. Santo, et al. obtained good accuracy by combining the results for three thresholds with a majority decision [7].

*Address: Kayaba-cho Tower Bldg., 1-21-2 Shinkawa, Chuo-ku, Tokyo 104-0033, Japan. E-mail: ksuzuki@rd.nttdata.co.jp

Because conventional methods use the most intense movement in the whole screen for boundary detection, they often fail to detect real shot boundaries. These intense movements, which occur for large moving objects or quick panning, affect the feature values so that shot boundaries are detected incorrectly. To overcome this problem, we propose using the motion direction histogram as a feature [8]. The MDH is based on the fact that the distribution of motion vector directions differs between an abrupt shot boundary and intense movement across the whole screen. We examined the effectiveness of this feature by evaluating it through detection experiments.

2 Abrupt Shot Boundary Detection Method Using Motion Direction Histogram

2.1 Motion direction histogram feature

We notice the directions of motion vectors between two adjacent frames because there is no similar macro block between consecutive frames at the shot boundary. In this case, it is consider that the directional distributions of motion vectors can be regarded as random, because motion vector search algorithm detects pseudo corresponding points. On the other hand, motion vectors for intense movement can have the same directions, because there is a similar macro block between consecutive frames in a shot with intense movement.

We have found that the directional distributions of motion vectors differ between an abrupt shot boundary and a shot with intense movement. This means that it is possible to prevent false detection of an abrupt shot boundary in a shot with intense movement. To achieve this, we propose the use of a motion direction histogram (MDH) to express the differences in the directional distributions of motion vectors. The computing process for our MDH feature is as follows:

1. Find motion vectors between consecutive frames;
2. Generate a directional histogram from these motion vectors;
3. Compute the sum of the absolute bin-wise differences between this histogram and the previous one.

We illustrate the concept of the MDH feature in Fig.1.

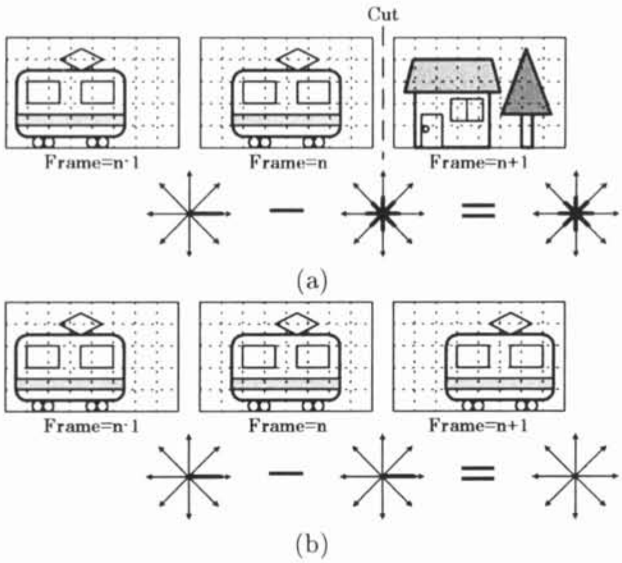

Figure 1: Concept of the motion direction histogram: (a) shot boundary, (b) moving object.

When an abrupt shot boundary occurs, the value of the MDH increases, because there is no continuity between frames. Then the directional distribution of the motion vectors changes from an inclined distribution to a random distribution. Therefore, the sum of the absolute bin-wise differences increases. This situation is shown in Fig.1(a). On the other hand, the value of the MDH becomes small in a shot with intense movement across the whole screen, because there is continuity between frames in this case. The directional distribution of the motion vectors does not change significantly, so the sum of the absolute bin-wise differences is small. This situation is illustrated in Fig.1(b). Thus, by applying the MDH, it is possible to distinguish between an abrupt shot boundary and a shot with intense movement throughout the screen.

2.2 Scheme for feature integration

Here we describe our scheme for integrating multiple features in shot boundary detection. Conventional methods determine their thresholds through trial and error from the viewpoint of expressing shot boundary features.

In contrast to these methods, we consider the shot boundary detection problem to be a dichotomy between shot boundaries and non-shot boundaries in a multidimensional feature space. Our task in shot boundary detection is thus to distinguish shot boundaries from other boundaries. In turn, we can consider this a pattern recognition problem. Therefore, we can apply the techniques of statistical pattern recognition.

3 Abrupt Shot Boundary Detection Experiment

We confirmed the effectiveness of our method through experiments. In these experiments, we used a discriminant method to integrate the MDH feature with three conventional features: the average of the color difference for each macro block, the color histogram difference, and the average length of the motion vectors. The computing processes for these features are as follows:

Table 1: Experimental video data.

Name	Shot Boundary	Minutes	Usage
Drama1	322	43	Learning
Drama2	317	43	Evaluation
Drama3	349	43	Evaluation
News1	241	45	Learning
News2	276	45	Evaluation
News3	126	25	Evaluation
Sports1	246	40	Learning
Sports2	221	40	Evaluation
Sports3	361	45	Evaluation

- Average color difference between macro blocks.
  1. Divide a frame into macro blocks (the block size is 16 x 16 pixels);
  2. Compute the average of the color values (red, green, blue) for each macro block;
  3. Compute the color difference between this frame and the previous frame for each macro block;
  4. Compute the average of the color differences.
- Color histogram difference.
  1. Divide a frame into macro blocks (16 x 16 pixels);
  2. Compute the average of the color values for each macro block;
  3. Generate a 64-bin color histogram for all macro blocks in the frame;
  4. Compute the sum of the absolute bin-wise differences between this histogram and the previous one.
- Average length of the motion vectors.
  1. Divide a frame into macro blocks (16 x 16 pixels);
  2. Find a motion vector between consecutive frames for each macro block (the search region is 16 pixels up, down, right, and left);
  3. Compute the length of each motion vector;
  4. Compute the average of these lengths.
- MDH Feature.
  1. Divide a frame into macro blocks (16 x 16 pixels);
  2. Find a motion vector between consecutive frames for each macro block (search region of 16 pixels in each direction);
  3. Generate a 32-bin directional histogram from these motion vectors;
  4. Compute the sum of the absolute bin-wise differences between this histogram and the previous one.

Table 2: Accuracy of shot boundary detection.

	Without MDH		With MDH	
Video	Recall	Precision	Recall	Precision
Drama	99.4 %	93.4 %	99.5 %	98.5 %
News	97.8 %	66.7 %	98.0 %	92.5 %
Sports	93.3 %	69.2 %	88.1 %	94.1 %
Total	96.8 %	76.7 %	95.2 %	95.5 %

Table 3: Accuracy of shot boundary detection with discrimination dictionary for each genre.

	Without MDH		With MDH	
Video	Recall	Precision	Recall	Precision
Drama	99.7 %	87.9 %	99.2 %	99.8 %
News	97.0 %	85.5 %	96.8 %	98.2 %
Sports	97.9 %	60.4 %	97.1 %	92.0 %
Total	98.4 %	75.4 %	97.9 %	96.6 %

In the experiments, we employed linear discriminant analysis, because it is simple and highly accurate. We prepared three sample videos from different genres: sports, news, and drama. The videos are listed in Table 1. Among the three sample videos, one was used for learning, while the other two were used for evaluation. We defined the notions of recall and precision as follows:

$$Recall = \frac{Correct}{Correct + Missed} \times 100, \quad (1)$$

$$Precision = \frac{Correct}{Correct + FalsePositive} \times 100. \quad (2)$$

We show our experimental results in Table 2. Our MDH feature achieved a total precision of 95 %.

For the sports video, however, the recall was decreased by adding the MDH feature. The features that enhance shot boundary detection are seemingly different for each genre. To understand this degrade for the sports video, we obtained scatter plots by principal component analysis for each video genre. The scatter plots are shown in Fig.2. We found that the shot boundaries area and non-shot boundaries area were quite near in the sports video. In other words, it is difficult to distinguish shot and non-shot boundaries properly for all genres with a single discriminant function.

We thus developed the following hypothesis: the detection accuracy can be improved by applying a different discriminant function for each genre. To test this hypothesis, we conducted further experiments by using three different discriminant functions, each trained for a specific genre. We show these experimental results in Table 3. With this improved approach, we could detect abrupt shot boundaries more correctly.

4 Conclusion

In this paper, we have proposed the motion direction histogram (MDH) as a new feature for abrupt shot boundary detection in video sequences. The MDH was designed to solve the problem of shots with intense movement being detected as abrupt shot boundaries. Our experimental results demonstrated that by applying discriminant analysis and integrating the MDH with conventional features, we could successfully reduce the false detection of shots with intense movement. In addition, we found that the shot boundary detection accuracy was different for video of different genres. By developing a separate discriminant function for each genre, we could further improve the accuracy of shot boundary detection. In the future, we will apply our proposed MDH feature and feature integration scheme to gradual shot boundary detection.

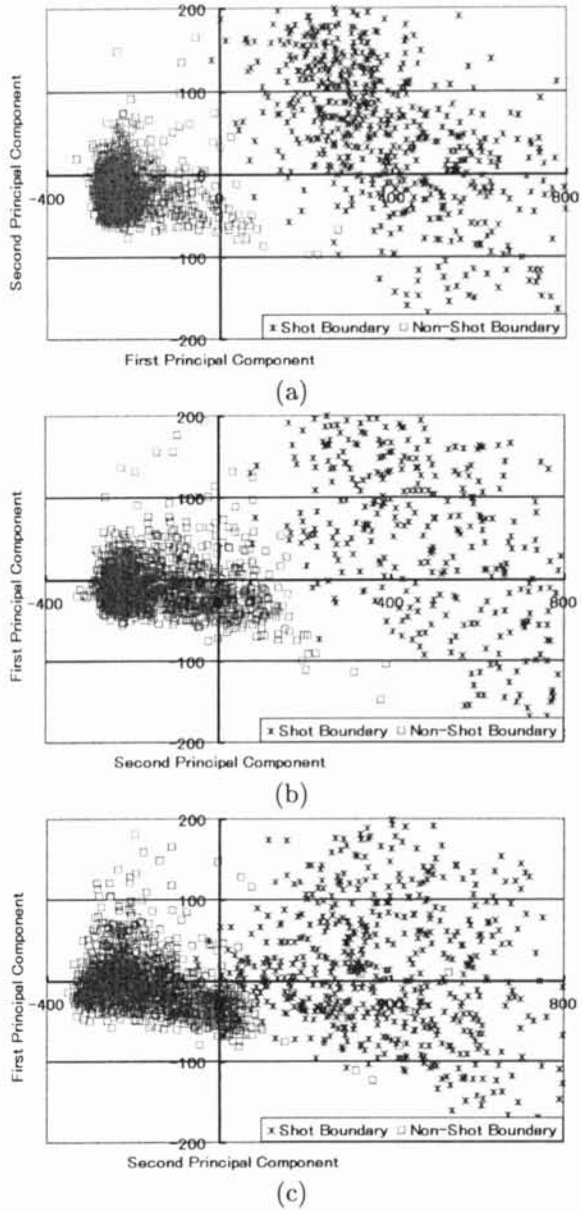

Figure 2: Scatter plots obtained by principal component analysis for each genre: (a) drama, (b) news, (c) sports.

References

[1] B. T. Truong, C. Dorai, S. Venkatesh, "New Enhancements to Cut, Fade, and Dissolve Detection Processes in Video Segmentation", ACM, ACM

Multimedia 2000, pp. 219–227, 2000.
[2] R. M. Ford, C. Robson, D. Temple, M. Gerlach, "Metrics for shot boundary detection in digital video sequences", Springer, Multimedia Systems, vol. 8, pp. 37–46, 2000.
[3] B. L. Yeo, B. Liu, "Rapid Scene Analysis on Compressed Video", IEEE, Trans. on Circuits and Systems for Video Technology, vol. 5, no. 6, pp. 533–544, December 1995.
[4] C. W. Ngo, T. C. Pong, H. J. Zhang, R. T. Chin, "Motion-based Video Representation for Scene Change Detection", Proc ICPR, vol. 1, pp. 827–830, September 2000.
[5] R. Zabih, J. Miller, K. Mai, "A feature-based algorithm for detecting and classifying production effects", Springer, Multimedia Systems, vol. 7, pp. 119–128, 1999.
[6] S. W. Lee, Y. M. Kim, S. W. Choi, "Fast Scene Change Detection using Direct Feature Extraction from MPEG Compressed Videos", IEEE, Trans. on Multimedia, vol. 2, no. 4, pp. 240–254, December 2000.
[7] M. D. Santo, G. Percannella, C. Sansone, R. Santoro, M. Vento, "Automatic Cut Detection in MPEG Movies: A Multi-expert Approach", T. Caelli, A. Amin, et al, Eds, "Structual, Syntactic, and Statistical Pattern Recognition", LNCS 2396, pp. 699–708, Springer, 2002.
[8] K. Suzuki, M. Nakajima, H. Sakano, "Scene Change Detection from Video Sequence Based on Discriminant Method", Technical Report of IEICE, PRMU2001-47, July 2001, (in Japanese).