

Video Shot segmentation algorithm based on SURF

Shujing Pan^{1, a}, Shutao Sun^{1, a}, Lifang Yang^{1, a} and Fengfeng Duan^{1, a}

¹School of Communication University of China, Beijing 100000, China;

^asjpan0124@163.com

Keywords: video retrieval, video shot segmentation, SURF.

Abstract. Video Shot segmentation is the key technology in content-based video retrieval and browsing, and which will directly affect the results of video retrieval. In view of the problems that the traditional shot segmentation algorithm is complex, the feature of video frame is not ideal, and the segmentation accuracy is low, this paper proposes a shot segmentation algorithm based on SURF (Speeded up Robust Features). This algorithm obtains the boundary of shots by computing the SURF features of video frames and calculating the feature matching rate. The experimental results show that: this algorithm can greatly reduce the amount of computation and data, improve the efficiency of the algorithm. And this algorithm has a great improvement in the accuracy of shot segmentation compared with shot segmentation algorithm based on color histogram.

1. Introduction

With the development of multimedia technology, video resources are becoming more and more abundant. In order to better manage and use video resources, it is a good method to construct content-based video retrieval system. In order to construct the system, we first need to restructure the video data into hierarchical structure data, which can be defined as the story, the scene, the camera and the frame according to the top-down structure. It requires complex technical support to transform a video stream to structure data accurately, and shot boundary detection is based on the analysis of the video stream. So it is a very important research topic to detect the shot boundary accurately and quickly. Shot segmentation is to find the boundaries of the video data, and analyze the difference of each frame of the video sequence.

In the past decade, many algorithms have been proposed to solve the problem of shot segmentation in video sequences, which have their own advantages and disadvantages [1, 2]. The classic method to determine the shot boundary is to select the underlying features of the image and calculate the frame difference between adjacent frames in the video. However, these algorithms are generally of high complexity, and the detection effect is not very ideal. In view of this situation, and taking into account the SURF [3] algorithm is small amount of calculation, fast calculation speed, good robustness for the translation, scaling, noise, gray changing, this paper proposes a method based on SURF.

2. Correlational Research

Video data is actually a set of static digital image sequences, and these digital images will be broadcast in the form of electrical signals orderly. The video data can be divided into four levels: video layer, scene layer, shot layer, video frame. As shown in Figure 1.

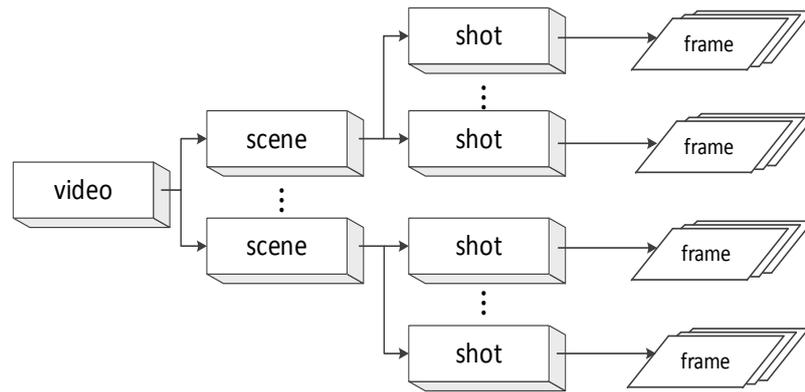


Fig. 1 Video structure

Shot segmentation is usually divided into the shot segmentation based on the compressed domain and the non-compressed domain [4-5]: The shot detection algorithm based on the compressed domain does not need to carry out the video decompression, but the compressed domain has less information of features; The basic principle of shot detection based on non-compressed domain is to extract the color, texture, shape and other basic information from the video frame directly. The shot segmentation algorithms based on non-compressed domain can be divided into Pixel Difference Method, Histogram Comparison Method, Clustering Method, Edge -based Method, Block Matching Method [6].

When the lens is converted, the conversion of the video frame between the conversions will occur a large range of pixel changes. Pixel difference method is to calculate the pixel difference data, monitor the shot boundary according to these data; Histogram comparison method is a common method in the field of shot boundary detection, which can be calculated by many kinds of histograms, including gray histogram, luminance histogram, RGB histogram, HSV color histogram, and so on. The method detects the shot boundary by calculating the difference between two adjacent frames and comparing the threshold; Clustering method is a non-supervised learning method, which is classified by the investigation of the elements in the sample set. The basic principle of the algorithm is that the video frames are divided into two types, the shot boundary and the non-shot boundary according to the feature similarity degree; Edge -based method is mainly characterized with calculating the edge of the object contour information in lens.

Literature [3] first proposed SURF algorithm, which is a fast and robust feature extraction algorithm based on the SIFT (Scale Invariant Feature Transform) algorithm. Under the influence of image rotation, translation, scaling and noise, the SURF algorithm has good robustness. In addition, it also improves the matching time by adding an integral image and a box filter. In the literature [7, 8], the performance evaluation of some typical local feature algorithms is carried out. The results show that the SURF algorithm is the most robust local feature algorithm. In view of the advantages of SURF algorithm, this paper introduces the SURF into the research of shot segmentation.

3. Research on SURF

As mentioned in the second section of the lens segmentation methods, shot segmentation is to find lens boundary which is a video frame changed obvious. In this paper, we propose a method of shot segmentation. Firstly, the video frames are registered. The second step is to determine the boundary of the shot, and then a shot is divided according to the results.

In this paper, we use SURF algorithm for image registration. The SURF algorithm is a fast and robust feature extraction algorithm based on SIFT. In 1999, Lowe D G proposed an image local feature descriptor operator (SIFT) [9]. It is a strong local feature of the image, which has a certain robustness to the image rotation, translation, scaling and noise. The algorithm was improved in 2004. However, the SIFT algorithm is large amount of calculation, and the registration time is longer, for these problems, the SURF algorithm is born. At present, many scholars at home and abroad have

carried on the research to SURF algorithm, and try to combine the concrete application. In short, the SURF feature detection algorithm is an accelerated version of the SIFT algorithm.

The calculation steps of SURF feature detection algorithm are as follows:

(1) Scale space extreme value detection

For the detection of extreme value, we need to establish a scale space. The Surf algorithm carries on the convolution computation using the box filter box and the original image. In this process, the amount of calculation has nothing to do with the image size because of the integral image's characteristic. The SURF algorithm can structure a multi-scale spatial function by changing the size of the box filter, using different sizes of the filter with the original image in different directions. After the establishment of the multi-scale space, the algorithm needs to use the fast Hessian matrix to detect the extreme points of each image. In the scale space and image space, the points of the extreme points are interpolated, and the points of the low contrast are selected [6].

(2) Characteristic point direction determination

In order to keep the SURF feature to the rotation, a main direction is needed for each feature point.

(3) Feature vector generation

A SURF feature point is composed of 4 seed points, and each seed point has 8 directions.

4. Shot Segmentation Algorithm based on SURF

In this paper, the shot segmentation algorithm uses the SURF features as the features of the current video frame, then uses the feature difference between video frames as the basis for the detection of shot boundaries. Because of the similarity of video frames in a shot, it is not necessary to calculate the SURF features of each video frame. By the way, calculating the SURF features of each frame will increase the amount of computation. So the algorithm uses the method of interval frame. The algorithm first sets the frame interval as N , then extracts the current frame SURF feature points and calculates the SURF features, and finally gets the shot jump boundary through the matching rate between the frames.

4.1 Algorithm principle

The SURF algorithm is used to extract the feature points from the current frame and the previous frame. Matching rate is the ratio of the number of matching points and the total number of feature points. The matching rate is used to determine whether the standard of switching the lens, such as the change of the matching rate is judged to be shot boundary.

4.2 Algorithm description

Step 1 Set the interval frame as N .

Step 2 Scale the current frame into a certain height (default 216) of gray scale images.

Step 3 Extract SURF feature points and calculate the SURF features.

Step 4 Matching the SURF feature points of the interval frame. If the matching distance is less than 0.3, it is considered to find a pair of matching points.

Step 5 Calculate the matching rate of interval frame. If the matching rate of the frame is low, the frame is considered as the video shot boundary.

In this paper, the SURF features are used to represent the video information, so the accuracy of video segmentation is greatly improved. The surf features are computed by the method of interval frame, and the amount of calculation of the algorithm is reduced.

5. Experiment And Result Analysis

The experiments are implemented using `c++` programming language on Microsoft visual studio 2010. All the experiments use 10 Chinese opera video clips for testing to certificate the accuracy of this algorithm. The frame numbers and the actual shot numbers of these video clips are shown in table 1.

Table 1 Experimental video data

Chinese opera clips	Frame numbers	Actual shot numbers
clip 1	9336	35
clip 2	10256	46
clip 3	8692	41
clip 4	9573	43
clip 5	9049	40

5.1 Interval frame experiment

The surf feature points are extracted from each frame of the video data. The workload is too large, and the similarity of each video frame is high. So there is no need to extract features from each frame. This paper chooses the method that extracts SURF feature from every N frames. In order to find the best number of interval frame, this paper selects five video clips for testing. Each clip carries on shot segmentation experiments as 5 frames and 10 frames, 20 frames, 30 frames, 50 frames, 100 frames. Experimental results are shown in Figure 2.

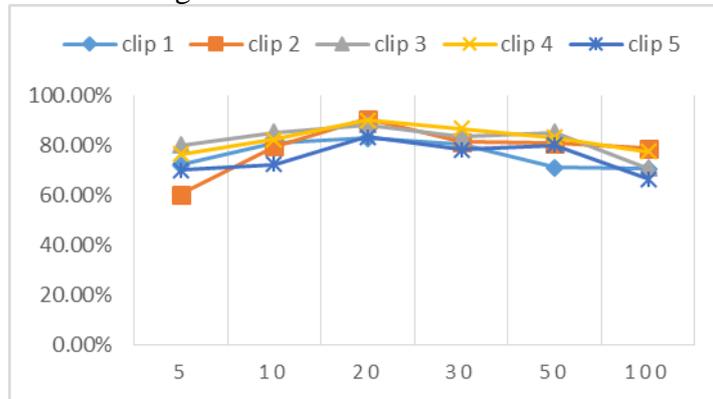


Fig. 2 Experiment Result

From Figure 2 it can be seen that the shot segmentation algorithm based on SURF features has a high accuracy. With the different frame interval, the accuracy of the shot segmentation will be changed. If frame interval is smaller, the result of segmentation will detect a lot of false detection lens, and the amount of calculation will increase; if the interval is large, the result will miss some shots. After comparison, the 20 frames can achieve very good shot segmentation results, and reduce the calculation amount.

5.2 Contrast experiment

Histogram comparison method is commonly used in the shot segmentation algorithm. In order to make a comparison between Histogram comparison method and the algorithm based on SURF of 20 frames, two experiments are carried. The experimental results are shown in Figure 3.

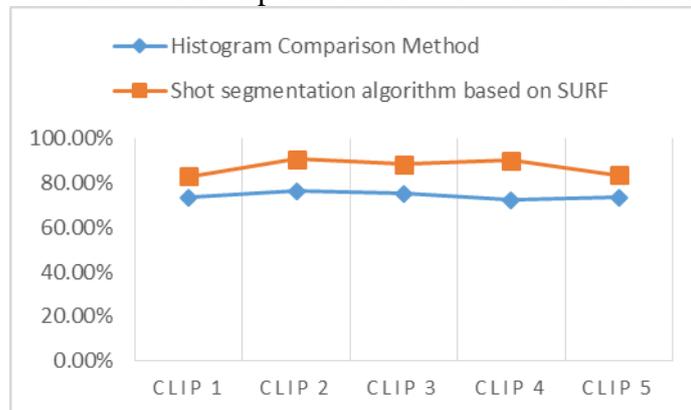


Fig. 3 Experiment Result

From Figure 3 it can be seen that the shot segmentation algorithm based on SURF has a great improvement compared with the Histogram comparison method. The accuracy rate of this algorithm has improved and the process is simple. The calculation amount of this algorithm is greatly reduced and it has very high practical value.

6. Conclusion

This paper analyzes the general principle of shot segmentation methods, and proposes an effective shot segmentation algorithm based on SURF. From the above analysis, it can be seen that the shot segmentation algorithm based on SURF can meet the video analysis based on content. Compared with the traditional algorithms, such as Histogram comparison method, this algorithm has a great improvement in accuracy. At the same time, this algorithm can reduce the computational complexity and data quantity by using the SURF feature of 20 interval frames.

7. Acknowledgment

This research was partially supported by The National Key Technology Research and Development Program of the Ministry of Science and Technology of China (No. 2012BAH37F02).

8. References

- [1] LIU Wen-ping, LI Ye-bai, ZHANG Chang-nian. Video-shot-edge detection technology: a survey [J]. Computer Engineering And Application, 2006, 42 (21): 17-20.
- [2] Quenot G M, Daniel Moraru, Laurent Besacier. Clips at trecvid: shot boundary detection and feature detection [C]. Gaithersburg, Marland, USA: Proceedings of the TRECVID 2003 Workshop, 2006: 35-40.
- [3] Bay H, Tuvtellars T, van Gool L. SURF: Speeded Up Robust Features [C] // Proceedings of the European Conference on Intelligent and Autonomous Systems, Toulouse, France, 2007: 2-12.
- [4] Ramin Zabih, Justin Miller, Kevin Mai. A Feature Based Algorithm for Detecting and Classifying Scene Breaks [C]. San Francisco, California, United States: Proceedings of the third ACM international conference on Multimedia, 1995, 5(9): 189 - 200.
- [5] Lefever S, incent V. Efficient and robust shot change detection [J]. Springer, 2007, 2(1): 23 - 34.
- [6] Tsamoura E, Mezaris V, Kompatasiaris I. Gradual transition detection using color coherence and other criteria in a video shot meta-segmentation framework. IEEE, 2008: 45-49.
- [7] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27 (10): 1615-1630.
- [8] Mikolajczyk K, Tuytelaars T, Schmid C, et al. A comparison of affine region detectors [J]. International Journal of Computer Vision, 2005 (6): 43-72.
- [9] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004 (2): 1-5.