IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 49-52
www.iosrjournals.org

Temporal Video Segmentation

Aslam Y. Suriya[1], Dr. S. B. Kishor[2]
[1]*(aslamsuriya@gmail.com, Research Scholar, Gondwana University, Gadchiroli,India)*
[2]*(s.b.kishor.spc@gmail.com,HOD, Dept. of Computer, S. P College, Chandrapur.India)*

Abstract: *The process of shot break detection is a fundamental component in automatic video indexing, editing and archiving.Temporal video segmentation is that the first step towards automatic annotation of digital video for browsing and retrieval. This article offers an outline of existing techniques for video segmentation that operate each uncompressed and compressed video stream. The segmentation task is accomplished through event detection in a frame-by-frame processing setup.*
Keywords -*Temporal video segmentation, Shot boundaries detection, Shot classification, Video segmentation, unsupervised video analysis.*

I. Introduction

Several temporal segmentation ways are developed for different varieties of videos.Hanjalic et al. [1] planned a technique for detection boundaries of logical story units inmovies. In their work, inter-shot similarity is computed supported block matching of the keyframes. Similar shots square measure connected, and also the segmentation method is performed by connecting theoverlapping links. Rasheed et al. [2] planned a two-pass algorithmic program for scene segmentationin feature films and television shows. Within the first pass, potential scene boundaries of the video square measureinitially detected supported the colour similarity constraint, Backward Shot Coherence (BSC).

Over-segmented scenes from the first pass square measure then incorporate within the second pass, supported the analysis of the motion content within the scenes. Sundaram et al. [3] used the audio-visualfeatures of the video in show scene segmentation. First, 2 varieties of scenes, audio scenesand video scenes, square measure detected one by one. Then, the correspondences between these 2sets of scenes square measure determined employing a time-constrained nearest-neighbor algorithmic program. Adams et al. [4] proposed the “tempo" for the segmentation of the movies. The “tempo" of a shot is a combination of the shot length and the motion content of shot. The dramatic story sections or events in the movie are detected by ending the zero-crossings of the “tempo" plot.

Recent advances in transmission compression technology, as well as the significant increase in computer performance and therefore the growth of net, have diode to the widespread use and accessibility of digital video. Applications like digital libraries,distance learning, video-on-demand, digital video broadcast, interactive TV, transmission data systems generate and use massive collections of video data. This has created a requirement for tools that may efficiently index, search, browse and retrieve relevant material. Consequently, many content based retrieval systems for organizing and managing video databases are recently projected.[5][6][7].

As shown in fig. 1 temporal video segmentationis the first step towards automatic annotation ofdigital video sequences. Its goal is to divide thevideo stream into a set of meaningful and manageablesegments (shots) that are used as basic elementsfor indexing. Each shot is then represented by selectingkey frames and indexed by extracting spatialand temporal features. The retrieval is based on the similarity between the feature vector of the queryand already stored video features.

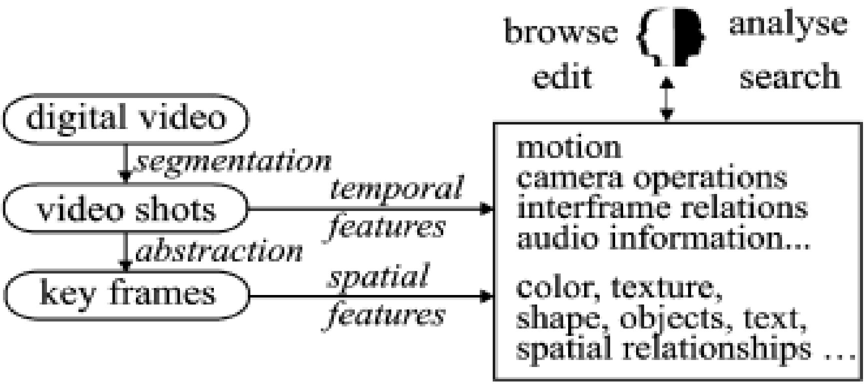

Fig.1. Content-based retrieval of video databases.

II. Temporal video segmentation

Early work focuses on cut detection,while more recent techniques deal with the harderproblem gradual transitions detection.

2.1. Temporal video segmentation inuncompressed domain

The majority of algorithms process uncompressedvideo. Usually, a similarity measure between successive images is defined. When twoimages are sufficiently dissimilar, there may bea cut. Gradualtransitions are found by usingcumulative difference measures and more sophisticatedthresholding schemes.Based on the metrics used to detect the differencebetween successive frames, the algorithms can bedivided broadly into three categories: pixel, block basedand histogram comparisons.

2.1.1. Pixel comparison

Pair-wise pixel comparison (also called templatematching) evaluates the differences in intensity orcolor values of corresponding pixels in two successiveframes.The simplest way is to calculate the absolute sumof pixel differences and compare it against a threshold [8]

2.1.2. Block-based comparison

In contrast to template matching that is based onglobal image characteristic (pixel by pixel differences),block-based approaches use local characteristicto increase the robustness to camera andobject movement. Each frame *I* is divided into*b* blocks that are compared with their correspondingblocks in i+*1*. Typically, the difference between*i*and *i*+1 is measured by

$$D(i , i+1) = \sum_{k=1}^{b} Ck\,DP(i, i+1, k)$$

Where*ck*is a predetermined coefficient for the block*k* and DP(*i*, *i*+1, *k*) is a partial match value betweenthe *k*th blocks in *i*and *i*+1 frames.

2.1.3. Histogram comparison

A step further towards reducing sensitivity tocamera and object movements can be done bycomparing the histograms of successive images.The idea behind histogram-based approaches isthat two frames with unchanging background andunchanging (although moving) objects will havelittle difference in their histograms. In addition,histograms are invariant to image rotation andchange slowly under the variations of viewing angleand scale [9]. As a disadvantage one can note thattwo images with similar histograms may have completelydifferent content.

2.2 *Temporal video segmentation in MPEGcompressed domain*

The previous approaches for video segmentationprocess uncompressed video. As nowadays video isincreasingly stored and moved in compressed format(e.g. MPEG), it is highly desirable to developmethods that can operate directly on the encodedstream. Working in the compressed domain offersthe following advantages. First, by not havingto perform decoding/re-encoding, computational complexity is reduced and savings on decompressiontime and decompression storage square measure obtained.

Second, operations square measure quicker as a result of thelower rate of compressed video. Last however notleast, the encoded video stream already containsa rich set of pre-computed options, like motionvectors (MVs) and block averages, that square measure appropriatefor temporal video segmentation.
Several algorithms for temporal video segmentationin the compressed domain are reported .According to the sort of knowledge used Fig. 2, they will be divided into six non-overlapping groups segmentation supported (1) DCTcoefficients; (2) DC terms; (3) DC terms, macroblock(MB) committal to writing mode and MVs; (4) DCT coefficients, MB committal to writing mode and MVs; (5) MB committal to writingmode and MVs and (6) MB committal to writing mode andbit-rate info. Before reviewing everyofthem, we have a tendency to gift a short description of the basicsof MPEG compression customary.

	Group					
Information used	1	2	3	4	5	6
DCT coefficients	✓			✓		
DC terms		✓	✓			
MB coding mode			✓	✓	✓	✓
MVs			✓	✓	✓	
Bit-rate						✓

Fig.2. Six groups of approaches for temporal video segmentation incompressed domain based on the information used

All of the previously mentioned algorithms have beendevised for shot cut detection only. The difference betweena frame pair during a gradual transition is much smaller thanthe difference that occurs during a shot cut. Lowering thethreshold to detect such small differences may result inmany false detections due to the differences caused bycamera and object motion. Zhang et al. proposed a twincomparison technique comparing the histogram differencewith two thresholds [10]. A lower threshold was used todetect small differences that occur for the duration of thegradual transition while a higher threshold was used in thedetection of shot cuts and gradual transitions. This methodcan fail when camera operations such as pans generate achange in the color distribution similar to that caused by agradual transition. To overcome this, they suggested analyzing the motion between frames to identify cameraoperations such as pans, tilts and zooms. Where this type ofmotion is identified the gradual transition is assumed to befalse to reduce the number of false positives. However, thismeans that gradual transitions containing object or cameramotions will not be detected.

III. Conclusions

Temporal video segmentation is the first step towards automatic annotation of digital video for browsing and retrieval. It is an active area of research gaining attention from several research communities including image processing, computer vision, pattern recognition and artificial intelligence. More than eight years of video segmentation research have resulted in a great variety of approaches. Early work focused on cut detection, while more recent techniques deal with gradual transition detection. The majority of algorithms process uncompressed video. Since the video is likely to be stored in compressed format, several algorithms which operate directly on the compressed video stream were reported.

References

[1]. A. Hanjalic, R.L. Lagendijk, and J. Biemond, \Automated High-Level Movie Segmentation for Advanced Video-Retrieval Systems", IEEE Transactions on Circuits andSystems for Video Technology, Vol. 9, Issue 4, 1999.

[2]. Z. Rasheed, M. Shah, “Scene Detection In Hollywood Movies and TV Shows", International Conference on Computer Vision and Pattern Recognition, 2003.

[3]. H. Sundaram and S.F. Chang, “Video Scene Segmentation Using Video and AudioFeatures", IEEE International Conference on Multimedia and Expo, 2000.

[4]. B. Adams, C. Dorai, S. Venkatesh, \Novel Approach to Determining Tempo and Dramatic Story Sections in Motion Pictures", International Conference on Image Processing, 2000.

[5]. S.-F. Chang, W. Chen, H.J. Meng, H. Sundaram, D. Zhong, VideoQ: an automated content based video search system using visual cues, in: Proceedings of ACM Multimedia Conference, Seattle, 1997.

[6]. W. Niblack, X. Zhu, J.L. Hafner, T. Breuer, D.B. Ponceleon, D. Petkovic, M.D. Flickner, E. Upfal, S.I. Nin, S. Sull, B.E. Dom, B.-L. Yeo, S. Srinivasan, D. Zivkovic, M. Penner, Updates to the QBIC system, in: Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases VI, Vol. 3312, 1997, pp. 150 - 161.

[7]. M. Smith, T. Kanade, Video skimming and characterization through the combination of image and language understanding, in: Proceedings of International Workshop Content-Based Access of Image and Video Databases (ICCV'98), Bombay, 1998.

[8]. T. Kikukawa, S. Kawafuchi, Development of an automatic summary editing system for the audio-visual resources, Trans. Electron. Inform. J75-A (1992) 204 - 212.

[9]. M.J. Swain, Interactive indexing into image databases, in: Proceedings of SPIE Conference on Storage and Retrieval in Image and Video Databases, 1993, pp. 173-187.

[10]. H. Zhang, A. Kankanhalli, S.W. Smoliar, Automatic partitioning of full-motion video, Multimedia Systems 1 (1) (1993) 10–28.