Enriching Feature Extraction Using A-priory Algorithm for Cricket Video

Deepali Bhawarthi*, Prof Shriniwas Gadage**

*(GHRCEM, Department of Computer Science, Pune University, Pune)

** (GHRCEM, Department of Computer Science, Pune University, Pune)

ABSTRACT

Techniques for rapid video browsing, video retrieval become essential as video material became more and more available and subtle. Video browsing is very important functionality of multimedia systems which offer the user efficient way to view relevant information from large amount of video material. On the other hand, video retrieval enables the user to search for particular video segment based on some description. Video summary is a process of presenting an abstract of entire video within a shorter period of time. It also aims to provide extraction of silent features, while preserving the sequence of events, as of the original video. It can also allow us to extract the required information when we are not interested in the whole unit due to time constraint. The purpose of the video event detection evolves mainly due to viewing time constraints. It help us to assess the relevance or value of information within a shorter period of time while decision making. It also plays prime role where the resources like storage, communication bandwidth and power are limited. The goal of this paper is to outline the advance of the mining concepts for cricket highlight generation which is an effort towards summarization. Here we have proposed one aspect of cricket event generation for wicket fall which can be extended towards events like boundaries and sixes

***Keywords* - Event detection, video browsing, video retrieval, summarization, multimedia systems.**

1. INTRODUCTION

Video is the collection of continuous frames which is displayed at some specific rate of 25 fps. Video itself contains a huge amount of data and complexity that makes the analysis very difficult. The huge amount of data that is produced by digitizing sports videos demands a process of data filtration and reduction. With remarkable development in multimedia systems, many sports applications came into birth. Video mining aims at extracting semantic structures from a video and thus enable summarization, browsing and indexing of the video content [5]. Video event analysis and recognition is a critical task in many applications such as detection of sports highlights, incident detection in surveillance video, indexing human-computer interaction [5]. Due to the long duration of the video, it is quite difficult process to index some particular event of the video. In recent years sports video analysis has become a widely research area of digital video processing because of its huge viewer ship and commercial importance [1], [2]. It has achieved a huge importance due to following points : (1)All national and international news broadcasts contain specific regular segments devoted to sports.(2) Consumers themselves are increasingly acting as generators of sport media as it becomes easier for parents and coaches to record school sports games, for instance.(3)From a broadcaster's point of view, efficient archiving of media facilitates later reuse in the creation of highlights packages or specialized DVDs.(4) Coaches and athletes are using the medium more and more to measure and correct technique, and to analyze team and individual performances. Cricket is a popular international game with a large viewer ship. Television broadcasters like Neo cricket, ESPN, Star Sports have huge databases of cricket videos.

The lengths of these matches are 4-5 days and thus extracting meaningful events which are of interest to the viewer is very important. One has to go through each frame to find out some specific event from it. Summarization is thus critical to help the user to extract the visual information that he requires without having to watch the whole video. Summarization of cricket videos is very important because of its long duration making manual highlights generation, tedious less explored area compared to other sports like soccer huge viewer ship. Because of the enormous difference in sports videos, sport Specific methods show successful results and thus constitute the majority of work. Some of the genre specific researches have been done in soccer (football) [3], [4] tennis [5], cricket [6], basketball [7], volleyball [8], etc. less work is observed for genre-independent studies [8], [9].

In viewer ship and fan following cricket is next to soccer [1]. Major cricket playing nations are India, Australia, Pakistan, South Africa, Zimbabwe, Bangladesh, Sri Lanka, New Zealand, and England. In spite of its huge viewer ship cricket has not obtained its share in the research community [6], [10]. Cricket video analysis is far more challenging because of the complexities of game in itself. Cricket is a game of variable factors as compared to other famous sports such as soccer, basketball, tennis etc. Some of these are its various formats test series which is duration of 4-5 days, one days and the new popular addition -T20-20, dynamic playing Conditions (field area, pitches etc), Day and day/night matches (cause illumination related problem) and duration[10]. Even the latest version of cricket i.e.T20-20 is played for approximately 3 hours which is greater than soccer which is played approximately for 90 minutes and hockey which is played approximately for 70 minutes.

2. RELATED WORK

There various types of sports structure depending upon which user can use this classification for highlight generation. Two broad classes of sports can be identified as court/table sports: e.g., badminton, tennis, table tennis, snooker/pool, volleyball, and basketball and field sports**:** e.g., soccer, football, baseball, cricket, and hockey[12]. In sports like tennis, badminton, cricket, and snooker, there is one camera view that dominates the broadcast (e.g., the full court in tennis) and contains almost all the semantically meaningful elements required to analyse the game. In sports like football, soccer, baseball, hockey, basketball, and volleyball, there is no such well-defined camera view that dominates the course of play [12].With the growth of digital video contents there is need to prune the recorded video contents to extract only the events, which are far more important then rest of the video .As video is collection of continuous frames displayed at normal rate of 25 frames per second, it occupies much more space as compared to image. so there is a need to define a continuity metric for the video in such a way that it is insensitive to gradual changes in camera parameters, lighting, and physical scene content, easy to compute and discriminate enough to be useful. The simplest way to do that is to extract one or more scalar or vector features from each frame and to define distance functions on the feature domain. Alternatively the features themselves can be used either for clustering the frames into shots or for detecting shot transition patterns or to index the frames to be used for later review.

3. Proposed Approach

In this paper we propose a key frame detection approach for minimizing the computation time which is crucial as the amount of data is huge. We have concentrated our work on T20-20 matches as they offer lots of events in shorter duration of time as compared to ODI's. The same work can be extended to ODIs as well as across multiple types of videos like news, movies, etc**.** Assuming temporal decomposition of video into structural units as clip, scenes, shots and frames forms the basis of most of the research in sports video processing. Based on single set of fixed or smoothly varying camera parameters like close-up, crowd, spectators, etc a group of sequential frames often form shot. Scene is a collection of related shots. A series of related scenes form a sequence and a part of the sequence is called as clip.

A video is composed of different story units such as shots, scenes, clips, and sequences arranged according to some logical structure defined by the screen play. In our work, we extract the events in the form of clips and after analysis assign a descriptive label to each clip.

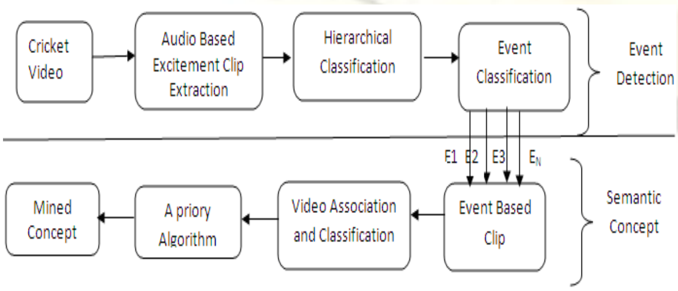

Fig. 1 System Architecture design

Figure 2 describes the complete flow of our proposed work. in first part key frames are detected, which reduces the processing time. From key frames, using image processing algorithms, events are detected. In second part, based on these events, video segment clips are segmented and processed separately. Labels are associated with frames and these labels are used as an input to a-priory algorithm to find out meaningful concept

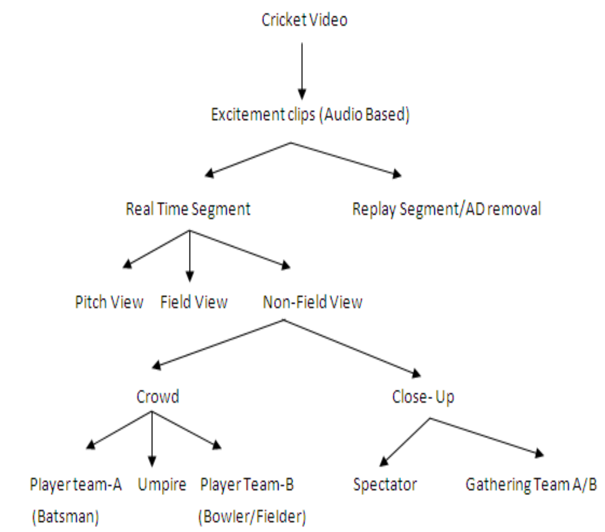

Fig. 2 Overall flow of the system

4. Classification and Detection

Event is defined as scenes in the video with some semantic meaning (i.e. labels from a semantic hierarchy) attached to it based on the leaf nodes shown Figure 2. Events are extracted as the leaf nodes of the level-2 to level-5 of hierarchical tree. The events are Real, Replay, Pitch View, Field view, Non-Field view, crowd, close up, Gathering of player of team A, Gathering of player of team B. The low-level feature extraction for the event detection and labeling is discussed in the following subsections.

4.1 Level 1 key frame extraction

We have observed that during the exciting events spectator's cheer and commentator's speech becomes louder. Based on this observation, we have used short-time audio energy feature for extracting excitement clip. The short-time is considered as the number of audio samples corresponding to one video frames. a particular video frame is considered as an excitement frame if its audio excitement exceeds a certain threshold. Any video frame sequences have certain properties like some frame sequences are either same or have a little difference between them. Thus using classifiers to classify all such frames is computationally inefficient. Suppose a sequence of same (s) and different (d) frames are:

s s s d s s s

For such a sequence of frames there would be a spike in the HHD plot with a rising edge and falling edge .we have to consider only the rising edges to decide on key and non key frames.

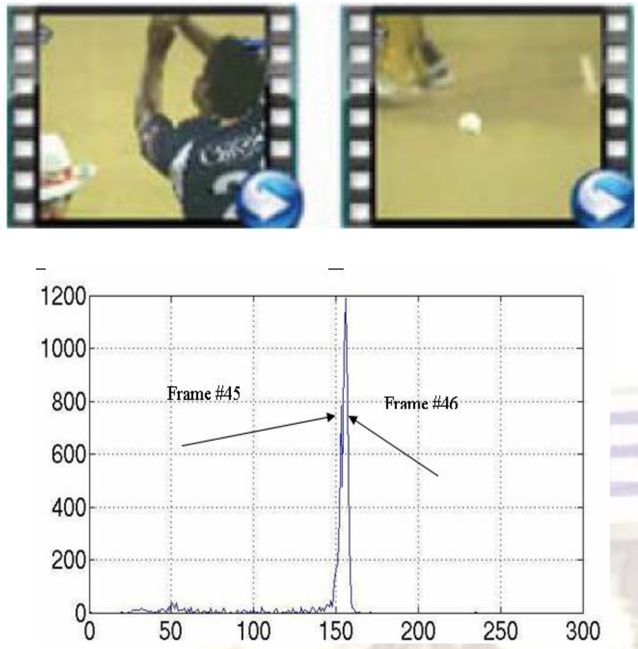

Fig. 3 key frame detection by hue histogram difference

4.2 Level 2 Reply / Real Time Segment Detection

Motion vector [6] and replay structures [12] are used to detect replays from sports video. But these methods are not robust enough to be suitable for various kinds of sports video replay detection because replays in different sports video are various and compiled in different manners and can hardly be represented by such simple features. Therefore the recent approach is to detect the accompanying logo effect of the replays in sports videos to acquire the replay segmentations. [13], [14]. It is been commonly observed that a replay segment is always sandwiched between two logo transitions or flying graphics which last for 8-15 frames as shown in figure 5. We have used the logo color property to detect logo frames or flying graphics and thus detect replay and live segment

4.3 Level 3 Pitch view/Field view

A pitch in a cricket field can be easily distinguished by its color, using this property we calculate the Dominant Soil Pixel Ratio (DSPR) to decide on the pitch view and non pitch view. We plot 256-bin histogram of the hue component of the frames. We pick up the peaks of hue histogram of these images. As the soil color of pitches in a cricket field change from venue to venue change in bins is also possible so it is important to compute the histogram first then observe the bins. A field view image can be distinguished by its grass color as in, using this property we calculate the Dominant Green Pixel Ratio (DGPR) [15] to decide the field view and non field view images as shown in the figure.

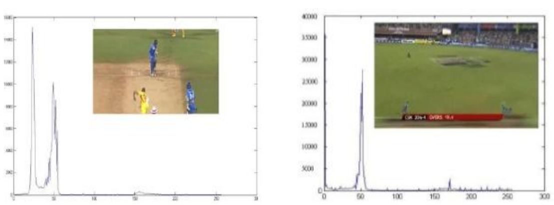

Fig. 4 Pitch view and field view

4.4 Level 4 Close up / Crowd Detection

At this level, we try to detect close up and crowd frames from the non pitch / non field view Frames detected at level 3. From our observations we have seen that close up or crowd frames are shown frequently whenever an exciting event occurs such as when a wicket falls, close up of batsman and bowler, then view of spectators and the player's gathering of fielding team are certainly shown. The edge detection is performed by finding the maximum gradient value of a pixel from its neighboring pixels. If the maximum value of gradient satisfy the threshold than the pixel is classified as an edge pixel.

Fig. 5 (a, b) Gathering of players,(c) Spectators ,(d) Edge detection.

5. SEMANTIC CONCEPT MINING

As we have determined the frame labels using low level features as described above now we can find out the concept associated with the set of events. After first level processing, we have labeled frames with the events like pitch view, field view, player of team A, player of team B, Umpire, Spectators and so on. Table 1 shows the label associated with individual such events

Table 1 Label associated with events

Sr. No	Label	Associated Event
1	Vp	Pitch View
2	Vf	Field View
3	Pa	Player of Team A
4	Pb	Player of Team B
5	U	U Umpire
6	Ga	Gathering of Players of Team A
7	S	Spectators
8	R	Reply

It's very obvious that during any excitement event camera will move over different parts of the ground and so various events are formed. By proper association of these events, we can extract out the concept like wicket fall. If there is a wicket fall of team A, than player gathering of team B, spectators, Close up of player of team A, umpire etc events will be detected repeatedly in very short span of video. Entire video segment is divided in to real time small segments; because there is no meaning to process the reply scene as it is already processed in real time clip. Key frames are extracted and reply segments are skipped from further processing. Remained real time segments are applied as input and key frames are classified. Using table 1, we associate specific label to each classified frame. For the robustness, we have kept the threshold of consecutive 25 frames. It means, if frames with same class label does not have count greater than 25 in single shot, than it would not be assigned any label for further processing. On label sequence, we are applying A-Priori algorithm to find out meaningful concepts like wicket fall or boundary and some more. Figure 5 show the flow of complete concept mining process.

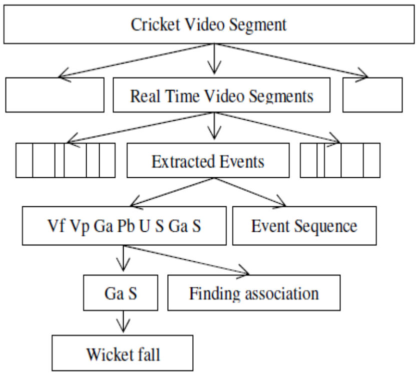

Fig 5. Typical example of wicket fall

6. Conclusion

Implementing video event detection using A-priory algorithm is an effort in making utilization of sports video more unique and natural for users. With a minimum hardware resource an attempt to provide a hassle-free attractive solution for most important event detection is made in a more natural way. The use of classifiers can exhibit better detection and classifications which can reduce the processing time thus save lot of CPU usage. This can create a new scope to generate customized and automatic cricket video highlights. This approach can not only be extended to other sports, but also to other type of videos such as news, movies, etc. for video summarization applications

References

[1] Mahesh Kumar H. Kolekar, Kannappan Palaniappan "Semantic Concept Mining Based on Hierarchical Event Detection for Soccer Video Indexing", journal of multimedia, vol. 4, no. 5, October 2009.

[2] Changsheng Xu, Jinjun Wang, Hanqing Lu, Yifan Zhang , "A Novel Framework for Semantic Annotation and Personalized Retrieval of Sports Video", IEEE transactions on multimedia, vol. 10, no. 3, April 2008.

[3]N.Harikrishna,Sanjeev,Satheesh,S.Dinesh,Sriram,K.S. Easwarakumar,"Content Based Image Retrieval using Dominant Color Identification Based on Foreground Objects".

[4] Hu Min, Yang Shuangyuan, "Overview of content-based image retrieval with high-level semantics" 20IO 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE).

[5] Dr.N.Krishnan, M.Sheerin Banu and C.Callins Christiyana, "Content Based Image Retrieval using Dominant Colour Identification Based on Foreground Objects", International Conference on Computational Intelligence and Multimedia Applications , 2007 IEEE, DOI 10.1109/ICCIMA.2007.64

[6] Xiaoyun Wang, Jianfeng Zhou" An Improvement on the Model of Ontology-Based Semantic Similarity Computation", 2009 First International Workshop on Database Technology and Applications, 2009 IEEE, DOI 10.1109/DBTA.2009.17

[7] Yu Xiao Hong and Xu Jinhua," The Related Techniques of Content-based Image Retrieval", 2008 International Symposium on Computer Science and Computational Technology.

[8] A. Kokaram, N. Rea, R. Dahyot, M. Tekalp, P. Bouthemy,P. Gros, and I. Sezan, "Browsing sports video: trends in sports-related indexing and retrieval work," in IEEE Signal Processing Magazine, vol. 23, no. 2, pp. 47–58, 2006.

[9] M. H. Kolekar and S. Sengupta, "Semantic Indexing of News Video Sequences: A Multimodal Hierarchical Approach Based on Hidden Markov Model," in IEEE Int.Region 10 Conference (TENCON), 2005.

[10] R. Brunelli and O. Mitch, "Histograms Analysis for image Retrieval," Pattern Recognition, Vol.34, No.8, pp1625–1637, 2001.

[11] A. Kokaram, N. Rea, R. Dahyot, M. Tekalp, P. Bouthemy,P. Gros, and I. Sezan, "Browsing sports video: trends in sports-related indexing and retrieval work," in IEEE Signal Processing Magazine, vol. 23, no. 2, pp. 47–58, 2006

[12] C Arun K.S, Hema P Menon, "Content Based Medical Image Retrieval by Combining Rotation Invariant Contour let Features and Fourier Descriptors", International Journal of Recent Trends in Engineering, Vol 2, No. 2, November 2009

[13] Arun K.S , Hema P Menon, "Content Based Medical Image Retrieval by Combining Rotation Invariant Contour let Features and Fourier Descriptors" International Journal of Recent Trends in Engineering, Vol 2, No. 2, November 2009

[14] Dr.N.Krishnan, M.Sheerin Banu and C.Callins Christiyana, "Content Based Image Retrieval using Dominant Colour Identification Based on Foreground Objects", International Conference on Computational Intelligence and Multimedia Applications , 2007 IEEE, DOI 10.1109/ICCIMA.2007.64

[15] H.B.Kekre, Dhirendra Mishra, "DCT Sectorization for Feature Vector Generation in CBIR ", International Journal of Computer Applications (0975 – 8887) Volume 9– No.1, November 2010.

[16] Hu Min, Yang Shuangyuan, "Overview of content-based image retrieval with high-level semantics" 20IO 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)