Event Detection In Broadcast Soccer Video By Detecting Replays

Shivani Arbat, Shashi Kumari Sinha, Shikha, Beena Khade

Abstract: Research on methods for detection and acknowledgement of events in video is assembling attention from various enthusiasts working in field of image processing. However, various techniques and methodologies had been proposed in order to meet such requirements. In this paper, we has proposed the efficacious method for event detection in soccer game broadcasted video and comprehending aspects which have been proposed to detect event and classify them in order to generate highlights. Also we provide an overview of the applications, categorizing them according to event detection and classification of domains.

Index Terms: event detection, event classification, highlights extraction, soccer video, video indexing, detect replays, logo detection

1 Introduction

ALL the time the multimedia creation such as movie, sports, news, video surveillance from several domains has led to immense creation of data. Among the generated videos, digital sport videos have become more and more pervasive [1, 2]. Soccer is one of the most popular team sports in the world that due to its wide distribution and large global audiences attracted increasing researches in the field of soccer video semantic analysis in the last decade. Moreover, the distribution of sports video across the Internet further increases the need for automatic video analysis, such as quick browsing, video summarization, detecting and recording interesting highlights for later review. Soccer game lasts for at least 90 min usually contains few semantic events in which viewer are interested in. Not only, Soccer, but also if we consider other sports like Cricket, American Football [5] consequently have dull events. For most of individuals the summarized video has efficaciousness, rather than watching the full length video. Sports highlights can be generically composed of interesting events that may capture the user's attention. Although generic highlights are sufficiently effective for casual video skimming, domain-specific (or classified) highlights will support more personalized applications and user queries such as "show me all soccer goal events." Most sports broadcasters distinguish key events by inserting editing effects such as slow-motion replay scenes and superimposed text display [3].Various techniques and frameworks have been proposed in order to achieve the semantic event detection. As one of the most fundamental components for video information management systems, the main functionality of event detection is to extract event and exploit their relationships inside large scale collections.

- *Shivani Arbat is currently pursuing bachelors degree program in Information Technology engineering in University of Pune, India, PH-+919921979069. E-mail: shivani.arbat@gmail.com*
- *Beena Khade received the B.E. degree in Information Technology from the University of Pune (UoP) in 2007, and the M.E. degree in Information Technology from the University Of Pune (UoP) in 2011. E-mail: khadebeena@gmail.com*

The basic procedure for the existing event detection methods can be generally divided into two main steps — 1) generating video content representation, where the video properties are extracted from raw sequence, and 2) decision making process for detection. In the second step, different kinds of data mining or statistical methods can be applied as detectors with preselected training examples. However, achieving accurate and robust detection is very difficult and challenging task. Video information has many unique characteristics such as Rich Semantics, Large Volume, High Dimensionality, and Complex Internal Structure [4].Therefore, developing multimodal techniques to integrate different kinds of information seamlessly is central of importance for effective knowledge discovery and information retrieval. Besides, hierarchical classification of events builds more accurate results. This paper reviews the development in event detection in the soccer game video. It discusses several major research issues such as soccer video classification, ball/player tracking, game highlight extraction, etc. It also demonstrates the potential applications of the techniques defined for event detection. Sport is an ever-green field and attracts big spending each year [7]. The requirement and expectation of users have grown significantly since digital media is popular. The aim of this paper is to provide an insight and lead the discussion of event detection done precisely.

2 Applications

In this section, we present a few application areas that will highlight the potential impact of event detection and event classification systems.

2.1 Content-Based Video Analysis

Video has become a part of our everyday life. With video sharing websites experiencing relentless growth, it has become necessary to develop efficient indexing and storage schemes to improve user experience. This requires learning of patterns from raw video and summarizing a video based on its content. Content-based video summarization has been gaining renewed interest with corresponding advances in content-based image retrieval. Summarization and retrieval of consumer content such as sports videos is one of the most commercially viable applications of this technology [9].

2.2 Animation and Synthesis

The gaming and animation industry rely on synthesizing realistic humans and human motion [11]. Motion synthesis finds wide use in the gaming industry where the requirement is to produce a large variety of motions with some compromise on the quality. The movie industry on the other hand has traditionally relied more on human animators to provide high-quality animation. However, this trend is fast changing [10]. With improvements in algorithms and hardware, much more realistic motion synthesis is now possible. A related application is learning in simulated environments. Examples of this include training of new players in soccer game with simulated subjects.

3 General Overview

In recent years, researches for discovering a solution to automate sport video semantic analysis have dedicated a lot of attention and panoptic research efforts have been devoted toward it. Semantic analysis of sport video due to varieties in different broadcasters, existence massive volume of data and semantic gap between low level features and high level semantics have been become a challenging problem in machine vision and pattern recognition techniques[8].Therefore , occupying this gap by combining several precise features from video can be a suitable consequence. A generic event detection system can be viewed as proceeding from a sequence of frames to a higher level interpretation of audio visual feature in a series of steps. The major steps involved are the following:

1) Input video or sequence of frames;
2) Extraction of brief low-level features;
3) Midlevel event descriptions from low-level features;
4) High-level semantic interpretations from audio-visual features.

(a) (b) (c)

Fig.1. Primary camera views in: (a) soccer, (b) AFL, and (c) basketball. [3].

This section we will discuss some relevant aspects of item 2 i.e. low-level feature extraction and item 3 and 4 will be form the subject of discussion in further sections. Event: An Event has been referred as a dynamic component: an event is “something happening at a given time and in a given location” [13]. Thus Event detection systems are built on event of interest, categorizing them as interesting or non-interesting. In this context we are referring events as “goal”, “save”, “corner kick”, “free kick”, etc., all events in a soccer game. To detect an event in video various methodologies have been proposed based on low-level feature extraction or high-level extraction, replays, tracking objects, tracking players, etc,. Besides, respective models are proposed like Hidden Markov Model, etc,. Basically, in any broadcasted video highlights are being replayed in the replay section. Mostly replays contains the events in game like saves , goals ,corner kicks , free kicks , card events , etc,. Viewers are interested in these events only. However, some of part is also being repeated twice or even more times respective of the importance of the event. Thus these replay section can better work for the highlights generation phenomenon by detecting them directly for further analysis. While detecting events in the video or views in particular video in particular frame then in spite of analyzing and reading each and every frame in video these shots can also be found in replay sections. Thus studying extracted replay sections nearly less than half of the length of video makes it easier and efficient to study events precisely.

4 Our Proposed Approach

All literature surveyed above could address the challenges and pervasive importance of event detection. However, most of them are heavily based on manual made rules and some unique patterns and predefined structures of audio/visual/textual features for detection and differentiation

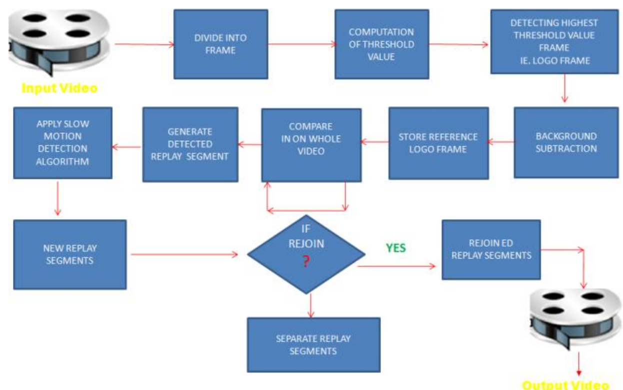

Fig -2: Our Proposed Approach

Events from each other. They consider crisp feature vectors for events and train their systems with some predefined models. But due to different presentation styles of broadcasters, loose structure of soccer games and a lot of motion and view changes, considering just crisp patterns arose out of domain knowledge would not provide necessitous accuracy during classification or event detection. Above figure depicts the flow from analyzing each and every frame of video to the output product of generated highlight video. The aim of this paper is to propose a method to minimize the amount of manual supervision in considering which set of features are related to an event and to provide a flexible system being able to tackle sequences exclude a regular pattern. Also, we aim at utilizing as little domain knowledge as possible to make the framework easily adapted to other sports with minimum adjustment. The overall diagram of the proposed method is shown in Fig. 2. As the figure shows, the method consists of only one phase. Here first the video will be divided into frames. Video we are considering is 25fps. Then each and every frame from the video is analyzed for calculating the threshold value. Threshold value v is the smallest detectable sensation on calculating the brightness of the frame. This value v is the primary measure to evaluate whether the frame is logo frame or any normal frame. Thus after setting the threshold value v, value vi (brightness for each frame) is been compared. This threshold value sets the template logo frame. If vi passes the threshold value then the frame is considered as non-logo frame or else if logo frame found then the timing of the frame in mm:ss is stored in file. This file is used for further evaluation of replay segments and to detect the presence of logo frame.

5 Logo Detection

5.1 Logo Sample frame detection

In a logo-transition, there is an image frame that has sudden increase in brightness of the frame. The frame having transition not necessarily need to be logo frame or any frame having any cup icon. Usually, the logo is highlighted and located at the middle part of a frame. Also the sudden increase in the brightness of the frame can be captured. With these prior knowledge's, we can extract the logo-samples threshold values. The algorithm for logo sample frame detection is as:

1) Compute frame-to-frame difference.
2) Check the frame difference (brightness difference). Proceed to step 3) when the difference exceeds a threshold, otherwise go back to step 1) for next frame.
3) Count the number of consecutive frame-differences that all exceed the frame-difference-threshold until encounter several consecutive frame-differences that drop brightness count below this threshold. If the counter exceeds a certain threshold, a wipe transition can be determined.

Otherwise go back to step 1) for next frame.

Fig -3: Logo frame transition and sudden increase in brightness in last logo frame

5.2 Logo threshold value determination

Threshold value has to be set in order to compare with other frames. If the frame value vi exceeds the threshold value then that frame is considered and noted to be logo frame , if not then the comparison extends to next frame of the video. All these vi when outperformance of the value is been captured then timing of that particular frame is been noted in text file for further replay recognition.

6 Replay Detection

After logo detection, a video can be divided into segments with taking logos as boundaries. Ideally, a pair of logos determines a replay segment. But actually, there are false and missing detection. Actually, we can recognize replay segments according to logos as we are focusing over the brightness change of the frame and however we are not annotating the significant frame as only the logo frame. Therefore, replay recognition is indispensable.

6.1 Character Selection

There are four kinds characteristics considered here: Length of the segment (LS). Usually, a non-replay is longer than a replay segment. Count of different types of shots. We classify shots into four types: Long Shot, Median Shot, Close-up Shot and Out-field Shot [14]. CLS, CMS, CCS and COS are represent the count of long shots, median shots, close-up shots and out-field shots respectively. Motion vectors are extracted through block matching [15]. To smooth the motion vector field, we run the mean shift [16] procedure to remove small motion blobs and noises. Motion activity (MA) is defined as the mean of magnitude of motion vectors in a frame.Thus, the feature vector of a segment consists of eight components, i.e. <SL, CLS, CMS, CCS, COS, ALS, MMA, VMA>[17].

7 Result Analysis

To test and validate effectiveness in logo detection, the experiments are conducted on video frame dataset. There can be as many number of frames which are present in the running video. The video frames taken for analysis has 25 fps. The classification of the frames can be done in only two ways , either the frame is logo frame or the frame is non –logo frame. The experiment show that this technique enhances the detection of the logo frame making ease in detecting replays. In the Fig. 3 , the set of input frames are loaded for the frame classification. The below graph Fig. 4 represent the brightness counts considered from particular frame set from the video where the logo frame is supposed to be viewed. The steep increase at frame 409 is the logo frame. The sudden increase from frame 408 to frame 409 and then utmost fall in brightness count at frame 410 clues the presence of logo frame.

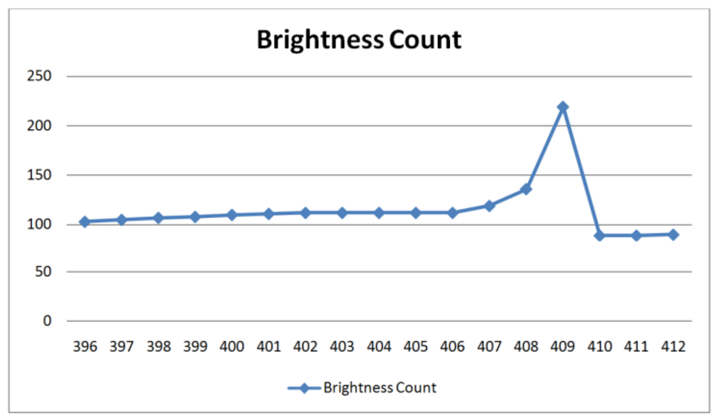

Fig.4-Graph Showing the variance in brightness of respective frames at time of logo occurrence.

Above graph depicts the increase in brightness count. However, the increase is kind of exponential and then the value count again stabilizes showing no uneven occurrence of high brightness frame. Thus we can resolve it as , the brightness count goes suddenly high and then shows steep decrease in its next successive frame. Graph in figure 5 depicts the brightness count when normal frames , where there is no logo occurrence. However in above figure we can learn that steep increase and decrease helps us in frame occurrence phenomenon , no such steep occurances in below figure.

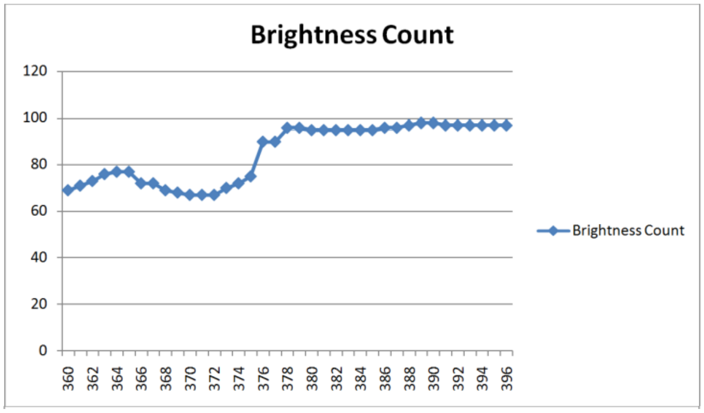

Fig.5-Graph Showing the variance in brightness of respective frames at normal frame occurrences in the video.

8 Conclusion

A reflexive and efficacious replay detection method is proposed in this paper. Avoiding complex and robustness direct analysis, our method is based on logo detection and replay recognition. Firstly, we detect some logo-transitions and extract some logo-samples from them during the video. Afterwards, the logo-template threshold value from these samples is been collected. Then, we use this template threshold to detect the other logos in the same video. After all logos are detected, the video can be divided into segments by taking logos as boundaries. Thus these detected replay segments can be expended for further analysis of the video or the game structure as soccer itself is a game with loose structure. A replay segment is usually played with a slow-motion pattern. Deeply analysis for motion is a natural and ultimate solution. But, it is difficult to extract distinguish motion features to represent slow-motion pattern. Apart from our method based on logo, analysis of intra-shot and inter-shot, i.e. characteristics of a shot itself and its context may be a praiseful method.

References

[1]. M.H. Hung, C.H. Hsieh, C.M. Kuo, J.S. Pan, Generalized playfield segmentation of sport videos using color features, Pattern Recognition Letters 32 (7) (2011) 987–1000.

[2]. M. H. Kolekar and S. Sengupta, “Event-importance Based Customized and Automatic Cricket Highlight Generation,” in IEEE Int. Conf. on Multimedia and Expo, pp. 1617–1620, 2006.

[3]. D. Tjondronegoro, Y.P.P. Chen, Knowledge-discounted event detection in sports video, IEEE Transactions on Systems, Man, and Cybernetics. Systems and Humans 40 (5) (2010) 1009–1024.

[4]. Jialie Shen, Dacheng Tao, and Xuelong Li, “Modality Mixture Projections for Semantic Video Event Detection”, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 18, NO. 11, NOVEMBER 2008

[5]. N. Babaguchi, Y. Kawai, T. Ogura, T. Kitahashi, Personalized abstraction of broadcasted American football video by highlight selection, IEEE Transactions on Multimedia 6 (4) (2004) 575–586.

[6]. M.H. Kolekar, K. Palaniappan, Semantic concept mining based on hierarchical event detection for soccer video indexing, Journal of Multimedia 4 (5) (2009) 298–312.

[7]. J. R. Wang and N. Parameswaran. Survey of sports video analysis: research issues and applications. In VIP ’05: Proceedings of the Pan-Sydney area work-shop on Visual information processing , pages 87–90,Darlinghurst, Australia, Australia, 2004. Australian Computer Society, Inc.

[8]. S. F. Chang, “The Holy Grail of content-based media analysis,” IEEE Multimedia Mag., vol. 9, no. 2, pp. 6–10, Apr. 2002.

[9]. D. A. Forsyth, O. Arikan, L. Ikemoto, J. O’Brien, and D. Ramanan, “Computational studies of human motion: Part 1, tracking and motion synthesis,” Found. Trends Comput. Graphics Vis., vol. 1, no. 2-3, pp.77–254, 2005.

[10]. BEN MEDLER, PhD STUDENT & BRIAN MAGERKO “Analytics of Play: Using Information Visualization and Gameplay Practices for Visualizing Video Game Data” PARSONS JOURNAL FOR INFORMATION MAPPING VOLUME III ISSUE 1, WINTER 2011.

[11]. David A. Sadlier and Noel E. O’Connor “Event Detection in Field Sports Video Using Audio–Visual Features and a Support Vector Machine”, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 15, NO. 10, OCTOBER 2005

[12]. Lamberto Ballan , Marco Bertini ,Alberto Del Bimbo ,Lorenzo Seidenari ,Giuseppe Serra, “Event Detection and Recognition for Semantic Annotation of Video”

[13]. Xiaofeng Tong, Qingshan Liu, Hanqing Lu and Hongliang Jin, “ Shot classification in sports videos,” International Conference on Signal Processing, Beijing, China, Aug, 2004.

[14]. URL ttp://www.intel.com/research/mrl/research/opencv

[15]. D. Comaniciu, and P. Meer, “Mean shift: a robust approach toward feature space analysis,” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, May, 2002, pp. 603-619.

[16]. Xiaofeng TONG, Hanqing LU, Qingshan LIU, Hongliang JIN “Replay Detection in Broadcasting Sports Video”