Review On Hybrid Approach Of Video Mining And Cassifiation For Cricket Highlight Extraction

[1]Amit M. Raval, [2] Prof. Kruti N. Khalpada,
[1]Student, [2]Assistant Professor,
[1]MECE, Atmiya Institute of Technology and Science (AITS), Rajkot, Gujarat

***Abstract* - This paper we audit every one of the strategies which can be useful to concentrate highlights from games video. There are a few techniques which take a shot at replay identification, subtitle location, occasion recognition, logo discovery and so on. All strategies have theirs upsides and downsides. We have quickly talked about this all techniques and attempt to dissect best coordinating strategy as per our application. Ordinarily highlight extraction is giving great results yet to decrease the length of separated video we can utilize order in it which can enhance the result.**

***Index Terms* - sports highlights, event detection, feature extraction, video summarization, classification**

I. INTRODUCTION

Regularly, sports recordings are fairly long, comprising of segments which are fascinating or energizing and segments which are exhausting, tasteless, and likely "a misuse of the viewer's profitable time."

Cricket is the most well-known game after soccer. It is the most loved game of Indians. Investigate work reported less in cricket correlation with different games. The explanation for that is cricket is long game correlation with soccer, b-ball, tennis, badminton. Option to that it is additionally intricate amusement and has different configurations like ODI, test cricket and as of late included T20.

The games video examination methods received so far can be comprehensively named occasion based or fervour demonstrate based. The occasion based video deliberation systems are generally in view of highlight based occasion discovery. Fervour demonstrate based highlight era procedures search for energy through elements, for example, movement action, cut thickness, sound vitality, shading following.

Occasion based highlights utilize more semantically important substance than the energy based highlights and its prosperity relies on the extravagance of the semantic ideas. Fervour show based strategies are bland in nature and require less space particular learning, however their execution changes from amusement to diversion.

Arrangement likewise utilized with the extraction. It is utilized to characterize the casings. Edges might be arranged into field-see, non-field see, shut everything down, and so forth. Grouping stresses the extraction video. It additionally lessens the extent of the video and evacuates the disconnected casings.

II. LITERATURE SURVEY

Different extraction techniques are: play and break occasion, question location, swarm cheer (sound-related), catchphrase spotting identification, replay discovery.

The point is to investigate cricket pictures that are to show up on TV and to concentrate subtitles [1]. It consequently creates highlights of diversion successions so that determinations of occasions can be found and played back. Energies levels are accumulated from the sound vitality and brief time zero crossing. Consider fig. 1. Inscription acknowledgment is done utilizing total of supreme contrast based subtitle acknowledgment display. Strategy diminishes manual preparing, empowers the era of customized highlight furthermore can be utilized for Content Based Video Retrieval. The approach appears to be compelling and around 80-85% exact in viable tests. It is important to give finish set of characters of the channel and earlier information of the inscription area is required.

A novel various leveled system and compelling calculations for cricket occasion discovery and grouping is given [2]. The proposed conspire performs top down video occasion location and order utilizing various leveled tree which maintains a strategic distance from shot recognition and grouping. In the chain of command, at level-1, sound elements are utilized to concentrate energy cuts from the cricket video. At level-2, fervour clasps are characterized into continuous and replay fragments. At level-3, the fragments are apportioned into field see and non-field see in light of overwhelming grass shading proportion. At level-4a, field view is characterized into pitch-see, long-view, and limit see utilizing movement veil. At level-4b, non-field view is grouped into close-up and swarm utilizing edge thickness include. At level-5a, close-ups are characterized into batsman, bowler/defender,

umpire utilizing shirt shading highlight. At level-5b, swarm section is grouped into onlooker and players' social event utilizing shading highlight.Consider fig. 2.

Fig. 1. Caption style in cricket video

Fig. 2. Replay detction frames

The proposed framework in [3] recognizes energizing clasps in light of sound components and after that arranges the individual scenes inside the clasp into occasions, for example, replay, player, official, observer, and players gathering. A probabilistic Bayesian conviction arrange in view of watched occasions is utilized to appoint semantic idea marks to the energizing clasps, for example, objectives, spares, yellow-cards, red-cards, and kicks in soccer video groupings. The named clasps are chosen by level of significance to incorporate into the highlights. The proposed framework chooses 100% premium idea which gives the great result. The proposed approach not just be stretched out to other sort of games, additionally to different sorts of recordings, for example, news, motion pictures for video rundown applications. Shut everything down pack casings are appeared in fig. 3.

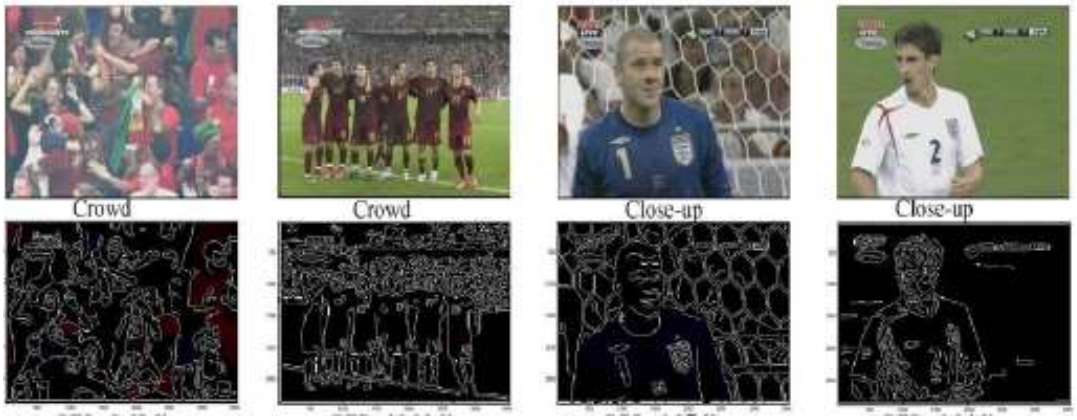

Fig. 3. Close up and crowd frames.

A generic method for sports video highlight selection is presented in [4]. Processing begins where the video is divided into short segments and several multi-modal features are extracted from each video segment. Excitability is computed based on the likelihood of the features lying in certain regions of their probability density functions that are exciting and rare. The proposed measure is used to rank order the partitioned segment stream to compress the overall video sequence and produce a contiguous set

of highlights. The video is first segmented into small blocks for feature extraction. Several features (scalar parameters) are extracted from each segment that is modelled to be generally proportional to the excitement level of the given segment. The multimodal events/features used for excitability measure: (1) slow motion replay, (2) camera motion activity, (3) scene cut density, (4) commentators' speech in high and (5) low excitement levels, and (6) audio energy.

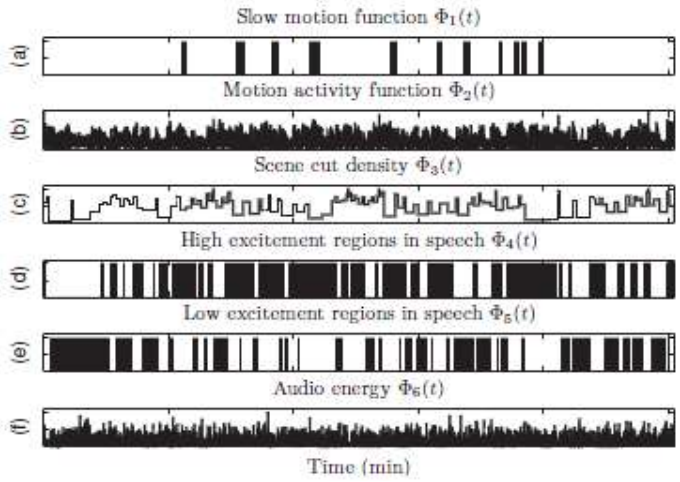

Fig. 4. A time-line view of the detected events/feature functions

In [5] 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 is proposed. The aim 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. Logo transition frames are shown in fig. 5. All frames from the video are analysed 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. It is difficult to extract distinguish motion features to represent slow-motion pattern. Analysis of intra-shot and inter-shot is not used.

Fig. 5. Logo transition frames

An algorithm to detect semantic concepts from cricket video is proposed in [6]. The proposed scheme works in two parts. In first part a top-down event detection and classification is performed using hierarchical tree. In second part, higher level concept is identified by applying A-Priori algorithm. High level concept mining from given video is two levels hierarchical process. At first level, meaningful events are identified associated with video using low level features. And at second level, higher level concepts are recognized based on previous level results. For cricket, events like, real time video, reply, pitch view, field view, close up classification, crowd classification are extracted in first level. Collections of such events are used later on for extracting the concepts like wicket fall (balled out, catch out, run out, stumped out etc.), boundary (fours and sixes), run milestones (Half century or century) etc. There is always a semantic gap between low level feature and high level concepts which needs to be filled up. The proposed approach is tested on T20 matches as they offer lots of events in short duration as compared to ODI's.

A novel semantics-based content analysis system for reliable media highlight extraction using Dynamic Bayesian Network (DBN) is proposed in [7]. It extracts the low-level evidences and then converts the input video to high-level semantic meaning. Specific domains contain rich spatial and temporal transitional structures that help the transformation process. A robust audio-visual low-level evidence extraction scheme is introduced. Based on DBNs, soccer events such as goal event, corner kick, penalty kick event, and card event can be found. Given a video in specific domain, the proposed system can extract the low-level evidence and interpret the input video in terms of high-level semantic. The low features like close-up view, camera motion,

audience region, audio, gate, replay, board, referee. Training can be categorized into two kinds: qualitative (structural training) and quantitative training (parameter training). Qualitative training concerns the network structure of the model and quantitative training determines the specific conditional probabilities.

In [8], a reliable logo and replay detecting approach is proposed. It contains two main stages: first, a logo transition template is unsupervised learned, a key frame (K-frame) and a set of pixels that describes logo object (logo pixels, L-pixels) accurately are also extracted; second, the learned information are used jointly to detect logos and replays in the video. A logo transition usually contains 10-30 frames, describes a flying or varying object(s). In this paper, a novel unsupervised learning and detection approach for logos and replays cross different sports videos are proposed. It first learns a logo template based on the motion characteristics; then learns a colour representation (Lpixels)of a key frame of the logo. In the detection stage, the colour representation is used to filter out frames that not lie in a logo; and the logo template is used to verify the true logos .Finally pair these found logos to get the replay clips. It uses a sequence template to model the transition of the logo object. Motion feature is used. The whole detection procedure is totally automatically. The approach is not applicable in case when multiple types of logo object occur in a same video.

In [9], a novel approach for detecting highlights in sports videos is proposed. The videos are temporally decomposed into a series of events based on an unsupervised event discovery and detection framework. The framework solely depends on easy-to-extract low-level visual features such as colour histogram (CH) or histogram of oriented gradients (HOG), which can potentially be generalized to different sports. The unigram and bigram statistics of the detected events are then used to provide a compact representation of the video. The effectiveness of the proposed representation is demonstrated on cricket video classification: Highlight *vs.* Non-Highlight for individual video clips (7000 training and 7000 test instances). The proposed approach consists of six stages: feature extraction, event discovery, event model refinement, event detection, video clip representation, and highlight clip detection. Feature extraction is used to pre-process all videos; event discovery and model refinement stages construct the event vocabulary; event detection and video clip representation transcribe the video; and finally feed into the training of the highlights detection classifier. A low equal error rate of 12.1% using event statistics based on CH and HOG features is achieved.

A novel hierarchical framework for soccer (football) video classification is proposed in [10]. The proposed scheme performs a top-down video scene classification which avoids shot clustering. This improves the classification accuracy and also maintains the temporal order of shots. The hierarchy, at level-I, audio features are used, to extract potentially interesting clips from the video. At level2, it is classified into field view and non-field view using feature of dominant grass colour ratio. At level-3a, it classifies field view into three kinds of views using motion-mask. At level-3b, it classifies non-field view into close-up and crowd using skin colour information. At level-4,it classifies close-ups into the four frequently occurring classes such as player of team-A, player of team-B, goalkeeper of team-A, goalkeeper of team-B using jersey colour information. The proposed hierarchical semantic framework for event classification can be readily generalized to other sports domains as well as other types of video.

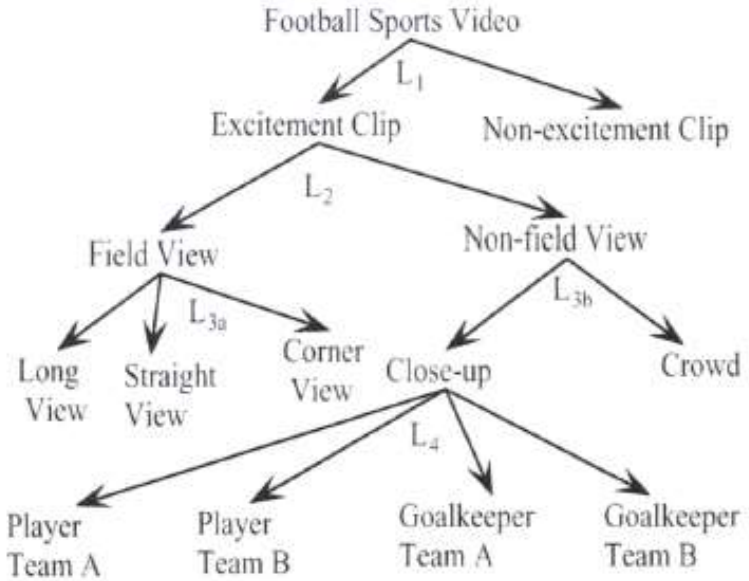

Fig. 10 tree diagram for hierarchical framework.

In [11], a visual content based algorithms to automate the extraction of video frames with the cricket pitch in focus. As a pre-processing step, first select a subset of frames with a view of the cricket field, of which the cricket pitch forms a part. This filtering process reduces the search space by eliminating frames that contain a view of the audience, close-up shots of specific players, advertisements, etc. The subset of frames containing the cricket field is then subject to statistical modelling of the grayscale (brightness) histogram (SMoG). Since SMoG does not utilize colour or domain specific information such as the region in the frame where the pitch is expected to be located, an alternative algorithm: component quantization based region of interest extraction (CQRE) for the extraction of pitch frames is proposed. The SMoG-CQRE combination for pitch frame classification yields an average accuracy of 98:6% in the best case (a high resolution video with good contrast) and an average accuracy of 87:9% in the worst case (a low resolution video with poor contrast). Since, the extraction of pitch frames forms the first step in

analysing the important events in a match; a post processing step, viz., an algorithm to detect players in the extracted pitch frames is proposed. The method presented in this paper to extract frames with a view of the cricket pitch is carried out in three phases. The first phase comprises a pre-processing step, which acts as a coarse filter. The field-frames include both pitch frames and extraneous information such as a view of the field near the boundary. In the second phase of the method, field-frames are further processed to separate frames with a view of the pitch ("pitch frames") from non-pitch frames using the following pitch detection algorithms: SMoG, CQRE and a combination of SMoG and CQRE methods. This second phase forms the crux of the proposed work. Finally, the third phase comprises a post-processing step, which uses the pitch frames to localize key players in the field of view.

In [12] a composite feature combining Optical flow analysis along with camera view analysis to model the type of shots played is presented. The work first presents an improved camera shot analysis based on learning parameters from a small supervision set. This splits the broadcast video into shots which are combined into balls and, the segment where the batsman is playing the stroke is identified. The approach works at three levels: a) classifying the type of view b) detecting the shots where effective play happens, and c) identifying the direction in which stroke is played. work can be combined with these to develop a fully automated cricket commentary system, which can provide a detailed analysis on a ball by-ball basis. The work has potential to be integrated with other requirements - such as video summarization for mobile devices, or identifying commercial break insertion points. The general approach of combining optical flow with other features already available may also be useful in other spatially coherent games such as baseball. After that optical flow analysis is used to determine the direction of the stroke with an accuracy of 80 percent.

In [13] a novel approach towards customized and automated generation of sports highlights from its extracted events and semantic concepts. A recorded sports video is first divided into slots, based on the game progress and for each slot, an importance-based concept and event selection is proposed to include those in the highlights. Our approach uses two levels of abstractions- one at the micro level, called events and the other at the macro level, called concepts. It creates a new scope to generate a customized and automatic cricket video highlight. 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.

In [14] a new framework meant for replay frames detection in cricket video is proposed. Framework is based on a block of score bar which is present in non-replay frame and not present in replay frame. Correlation is a signal matching technique used for obtaining similarity between two signals. Support Vector Machine (SVM) is a supervised machine learning technique meant for binary classification. A block of score bar from non-replay frame and replay frame are selected as template1 and template2 respectively. A training set of thousand frames containing five hundred non replay frames and five hundred replay frames used for training SVM. Feature vectors are calculated using correlation coefficient between templates and the training frames in the training set. Results have achieved an average recall of 96% and precision of 99%.

III. Conclusion

As reviewed, there are different methods of highlight extraction in sports videos and many techniques in classification of the extracted frames. Each methods and techniques have their merits and demerits.

References

[1] Kolekar, Maheshkumar H., and S. Sengupta. "Caption Content Analysis Based Automated Cricket Highlight Generation." *National Communications Conference (NCC), Mumbai*. 2008.

[2] Kolekar, Maheshkumar H., KannappanPalaniappan, and SomnathSengupta. "Semantic event detection and classification in cricket video sequence."*Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*. IEEE, 2008.

[3] Kolekar, Maheshkumar H., and SomnathSengupta. "Bayesian Network-Based Customized Highlight Generation for Broadcast Soccer Videos."

[4] Hasan, Taufiq, et al. "A multi-modal highlight extraction scheme for sports videos using an information-theoretic excitability measure." *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012.

[5] Arbat, Shivani, Shashi Kumari Sinha, and Beena Khade Shikha. "Event-Detection-In-Broadcast-Soccer-Video-By-Detecting-Replays." *international journal of scientific & technology research* 3.5 (2014): 282-285.

[6] Goyani, Mahesh, et al. "Wicket fall concept mining from cricket video using a-priori algorithm." *Proc Int J Multimedia Appl (IJMA)* 3.1 (2011).

[7] Chao, Chung-Yuan, Huang-Chia Shih, and Chung-Lin Huang. "Semantics-based highlight extraction of soccer program using DBN." *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*. Vol. 2. IEEE, 2005.

[8] Huang, Qiao, et al. "A reliable logo and replay detector for sports video."*Multimedia and Expo, 2007 IEEE International Conference on*. IEEE, 2007.
[9] Tang, Hao, et al. "Detecting highlights in sports videos: Cricket as a test case."*Multimedia and Expo (ICME), 2011 IEEE International Conference on*. IEEE, 2011.
[10] Kolekar, M. H., and K. Palaniappan. "A hierarchical framework for semantic scene classification in soccer sports video." *TENCON 2008-2008 IEEE Region 10 Conference*. IEEE, 2008.
[11] Jayanth, Sandesh Bananki, and GowriSrinivasa. "Automated classification of cricket pitch frames in cricket video." *Electronic Letters on Computer Vision and Image Analysis* 13.1 (2014).
[12] Kumar, Ashok, Javesh Garg, and AmitabhaMukerjee. "Cricket activity detection." *Image Processing, Applications and Systems Conference (IPAS), 2014 First International*. IEEE, 2014.
[13] Kolekar, Maheshkumar H., and SomnathSengupta. "Event-importance based customized and automatic cricket highlight generation." *Multimedia and Expo, 2006 IEEE International Conference on*. IEEE, 2006.
[14] Ravinder, M., and T. Venu Gopal. "Replay Frames Classification in a Cricket Video Using Correlation Features and SVM." *European Journal of Applied Sciences* 7.2 (2015): 92-97.