Views Detection from Cricket Video using Low Level Features

Hetal Chudasama
Assistant Professor, MBICT, New VVNagar., Gujarat, India
and
Dr. N M Patel
Associate Professor, BVM, VVNagar.,Gujarat, India

Detection of various views from the cricket video is a fundamental and useful step in cricket video summarization. In this paper, we propose an approach to detect various views from the cricket video. Detected views are Ground view, Pitch view, Boundary view, Close-Up view, Crowd view, Fielders gathering view and Sky view. Detection of specific view requires extraction of appropriate lower level features from the frame. Result analysis shows the accuracy of various algorithms used for view detection.

Keywords: view detection, lower level features, hue histogram

1. INTRODUCTION

Now a day, a huge volume of digital data is generated, uploaded and accessed on the multimedia by the users. Volume of this data increases every day with a very significant rate. These data is accessible in the form of image, video, audio or text.

Among various types of videos available like sports, news, entertainment, educational; sport video possess maximum mass of viewers. Sport video plays an important role in entertainment part of our daily life and has a wide range of audiences. Cricket, Football, Tennis, baseball, soccer and wrestling are among the variety of sports being played and viewed by the public in each and every country. Every country has one of the sports which rule as a favourite sport in the heart of their people. Indian would love to play and watch cricket more than any other sports. Almost each and every cricket match being played, it would be available on multimedia for their fans to watch later on. But, viewers may not be interested to watch entire match either because to watch entire match may be time consuming or users might be willing to watch only interested part from the match. As, much portion of the video dont contain meaningful event. Due to long duration of playing hours of cricket match, user may have to go through entire video until they get event of their interest. It might be very time consuming and tedious task. So, from the users perspective, it is necessary to develop a system to automatically analyse the sports video and generate summary of video according to ones interest so that they do not need to go through entire match. This need of the viewers, require to generate video summarization containing only interested events from the match like wicket falls and boundaries [Vijayakumar [2012]].

The major task of video summarization is to detect all such interested events from the cricket match video. At the base of the design of any such event detection algorithm is identification and classification of various views which constitute any sports event and an entire sports video. For cricket match, such views can be Field view, Crowd view, Close Up view and few other.

Temporal decomposition of any video will lead to its structural units like scene, shot, sequence and frame. Video is collection of various types of scenes. Scene is a collection of shots having the same context. Shot can be defined as unbroken sequence of frames captured by a camera. Frame is nothing but a 2−D image. Any randomly selected frame from the shot can be considered as key frame [Tang et al. [2011]]. Every key frame generates a particular view in the video. So, video must be decomposed into frames and need to be classified into various categories. It needs extraction of appropriate features from the frames.

Design of any such system which can analyse a sport will begin with gathering of information regarding characteristics of the various views available in that sport. Each and every view possesses some specific characteristics which differentiate it from other views. All such views which contain meaningful information needs to be identified [Jayanth and Srinivasa [2014]]. Characteristics of views can be represented by various types of visual features. Selection of visual features must be done carefully for proper representation of different views. Computational complexity of view detection algorithm is high, as visual features from the large number of frames needs to be extracted and analyzed to further detect various views from video.

The organization of paper is as follow. Section-2 represents work done related to view detection. Section-3 describes proposed methods. Section-4 shows experimental results. Section-5 concludes the paper with the idea of future work.

2. RELATED WORK

In cricket video, every shot can be represented by the key frame. Once all the key frames are extracted from the video, it can be classified into the various categories like field view or non-field view. Field view can be again categories into ground view, pitch view and boundary view. Non-field view can be close-up view or crowd view. Close-up view is further categories into player from Batting team, player of Balling team or umpire. Crowd can be sub categorise into spectators or gathering of fielders on the ground [Kolekar et al. [2008]]. Researchers use different features for view detection. Any view detection system will have feature extraction as its fundamental step.

Various techniques have been developed in the literature for extraction of features like visual feature, audio feature or textual features. Generally, the visual feature based event detection involves low-level feature extraction and high-level concept detection. In a cricket video, low level features like color scheme, edges, texture, motion, caption text, audio properties like Zero crossing rate can be used for semantic information extraction for high level analysis. Extraction of these features will require generation of domain specific knowledge of the cricket. A robust statistical model has been developed for detecting the domain-specific views. The specified model combines domain-independent global color filtering method. It also uses domain-specific constraints on the spatio-temporal properties of the segmented regions. High-level events are detected after the view recognition [Zhong and Chang [2004]].

In [Money and Agius [2008]], Low-level features such as color, motion and textures are used for video processing. Generally the cricket matches always have mass spectators who show their response in the form of loud cheering and applause. Some of the events like wicket fall, boundaries can be easily detected using audio based features. These audio symbols can be detected by measuring audio level. In case of any exciting events, audio level will be normally higher than regular audio level [Baillie and Jose [2003]]. These exciting events can also be extracted from the cricket video using text processing scheme [Pradeep [2013]]. Ground view and pitch view can be extracted using dominant colour component analysis using HSV histogram. Close up view can be detected by identifying skin colour component. Crowd view can be separated with the help of high edge ratio [Kolekar et al. [2008]].Boundary view can be detected using motion detection, morphological analysis and colour detection [Harikrishna et al. [2011]].This method first estimates the probability density function of the color component of the input image, and then applies steepest ascent hill-climbing search to the probability density function to generate several clusters.

Dominant field color in soccer video can be described by mean value of each color components, which are computed around their respective histogram peaks [Ekin et al. [2003]].Technique for identification of ball in soccer video has been also proposed. Dynamic programming is applied to find the actual ball trajectory [Pallavi et al. [2008]]. Arnau Raventos calculated color descriptors like Color Layout, Color Structure and Dominant Color using the MPEG-7 Low Level Feature Extraction Library for automatic soccer highlight generation [Raventos et al. [2015]]. Video features like HSV color conversion, RGB histogram, 3D RGB histogram and audio feature like short-time energy were used for logo detection and view recognition from soccer video [Sigari et al. [2015]].Pitch shots, bowling and batting shots are extracted using RGB histogram technique and Replay and Non Replay shots are extracted using score card detection [Kumar and Puttaswamy [2015]].

Temporal template were build using audio based features like energy related features, phoneme-level features, information complexity features, prosodic features in [Rui et al. [2000]] to generate highlight from baseball programs.Baseball game videos are first segmented into scene shots, and video segment between two pitch shots were treated as the basic unit and name it Pitch Scene block (PSB). Audio analysis and shot changing rate in each PSB are used for highlight decision purpose [W [2011]]. Mid-level features like Close-up, Outfield, Audience, Zoom In, Pan Right, Excited Speech and Cheering were selected from three modalities: Baseball Shot, Camera Motion, Audio for high level event detection in baseball [Chen and Tsai [2014]]. A set of video segments were first chosen to construct the transition effect template. The candidate frames were compared with this template for searching the slow-motion video segments. In baseball videos, further the pitching view template were constructed to locate starting positions of the video segments of interest [Su et al. [2013]].

In football match, strategies like footballers position, their movement trajectory and motion pattern analysis can be done effectively by efficiently extracting field lines [Zhang [2015]].

Event detection and highlight extraction from basketball was done by extracting features like motion intensity, frame dynamics, event ratio, number of gamers, number of viewers chat and number of emotion symbols [Chu and Chou [2015]].

Audiovisual Features like Moving distance of the player, Relative position between the player and the court, Applause/cheer sound effects and Length of the play were extracted to detect events in tennis match [Tien et al. [2008]].

Zhong et al [Zhong and Chang [2001]] proposed a unified framework for scene detection and structure analysis by combining domain-specific knowledge with supervised machine learning methods. A verification step based on object and edge detection is used in order to obtain more reliable results.A generalized playfield segmentation method that is suitable for various types of sport videos has been proposed in [Pramod Sankar et al. [2006]]. Method based on radial basis decompositions of a color address space followed by spatially localized texture decompositions using Gabor wavelets in frequency space for scene detection in sports video [Kapela et al. [2014]].

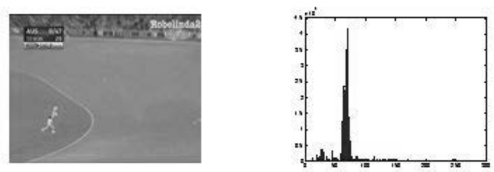
So, all sports analysis system design begins with the characterization of the views available and the identification of those views that contain the most of the information [Kokaram et al. [2006]].

3. PROPOSED METHOD

This section represents algorithms to detect different views like ground view, pitch view, boundary view, close up view, crowd view, fielders gathering view and sky view. Specific view can be categorized into appropriate category by selecting appropriate low level features. Selection of these features plays an important role in improving the accuracy of classification.

3.1 Field View

Here, color is considered as lower level feature for detection of Field view. Field view can be differentiated from other view by considering the dominant green color of grass. RGB image needs to convert to HSV image. Hue component of this HSV image is used to generate 256 bins Hue Histogram. It has been observed through various testing on field view images that peak amongst 256 bins occurs between bin $k = 60$ to $k = 75$. It indicates that, maximum pixels in the image are due to contribution of green color of the grass. This shows high probability of availability of grass in the image. Range of bins generating green color will be considered to cover the green color of grass in the image. Total number of pixels in this range will generate total number of grass pixels. GPR (Gp/n) is total number of grass pixels in the image to total number of pixels in the image. So, GPR will be high for views containing field image and less for the non-field images.



A. Field-View Image B. Hue-Histogram of Field-View Image

Figure 1 Field view image and its Hue Histogram

Figure 1A shows one of the images which belongs to the field view. We can see that majority of the pixels belongs to the grass of the field. Figure 1B shows its hue histogram. It shows that bins are concentrated between bins 60 and 65, which belong to green color. And, peak is at bin number 63 with $GPR = 0.895$. So, histogram indicates that this image belongs to the field view.

Algorithm 1 to detect Field View:

(1) Convert RGB frame into HSV frame format.
(2) Generate 256 bins Hue Histogram.
(3) Check the bin P with the highest peak.
(4) If the peak P is in the range of 60 to 75, Compute Grass Pixels by considering bins in the range of $P - 5$ to $P + 5$.
(5) Compute GPR(Grass Pixel Ratio) as ratio of Grass Pixels to Total pixels in the frame.
(6) Image can be consider as Field view, if $GPR > Th1$.

3.2 Pitch View

Here, color is considered as lower level feature for detection of pitch view. Pitch view can be differentiated from other view by considering the dominant soil colour of pitch. RGB image needs to convert to HSV image. Hue component of this HSV image is used to generate 256 bins Hue Histogram. It has been observed through various testing on pitch view images that peak amongst 256 bins occurs between bin $k = 25$ to $k = 35$. It indicates that, maximum pixels in the image are due to contribution of soil color of the pitch. This shows high probability of availability of pitch in the image. Range of bins generating soil color will be considered to cover the soil color of pitch in the image. Total number of pixels in this range will generate total number of pitch pixels. PPR (Pp/n) is total number of pitch pixels in the image to total number of pixels in the image. So PPR will be high for views containing pitch image and less for the non-pitch images.

A. Pitch-View Image B. Hue-Histogram of Pitch-View Image

Figure 2 Pitch view image and its Hue Histogram

Figure 2A shows one of the images which belongs to the pitch view. We can see that majority of the pixels belongs to the soil of the pitch. Figure 2B shows its hue histogram. It shows that bins are concentrated between bins 25 and 35, which belong to soil color. And, peak is at bin number 32 with $PPR = 0.773$. So, histogram indicates that this image belongs to the Pitch view.

Algorithm 2 to detect Pitch View:

(1) Convert RGB frame into HSV frame format.
(2) Generate 256 bins Hue Histogram.
(3) Check the bin P with the highest peak.
(4) If the peak P is in the range of 25 to 35, Compute Pitch Pixels by considering bins in the range of $P-5$ to $P+5$.
(5) Compute PPR(Pitch Pixel Ratio) as ratio of Pitch Pixels to Total pixels in the frame.

3.3 Boundary View

Here, color and motion are considered as lower level features for detection of boundary view. We can observe in the cricket video that players and crowd are objects which are entitled to movement. There will be a continuous motion when camera moves toward the crowd. Players will be also in constant movement, weather it is batsmen or baller or fielders. So we can utilize this motion information available in continuous frames of video to classify a specific frame as a Boundary view. By taking difference of continuous frames we can identify the pixel as field pixel or motion pixel. In the resultant image, background color is the color of field. We can observe that, in a boundary image, majority of time, audience will be available in upper part and field will be available in lower part. Divide the image into four blocks. So, upper two blocks will have movement of the audience. Lower two blocks will have green color of the grass and will not have movement. Motion can be detected in the image by taking difference of two consecutive frames. By comparing number of field pixels and motion pixels in upper two blocks with that of lower two blocks, image can be classified as the boundary image.

A. Boundary view image B. 4 Blocks view C. Motion detection

Figure 3 Boundary View, it's 4-blocks view and motion view

Figure 3A shows image from the boundary view. Figure 3B shows image 3A divided into 4 equal size blocks. It shows that upper two blocks contain audience and lower two blocks contain field with grass. Figure 3C shows corresponding motion view. We can see that majority of motion is present only in upper two blocks and lower blocks with least movement. Lower two blocks have majority of the field portion. So, these two blocks hue histogram will have concentration of bins at the green color of the grass.

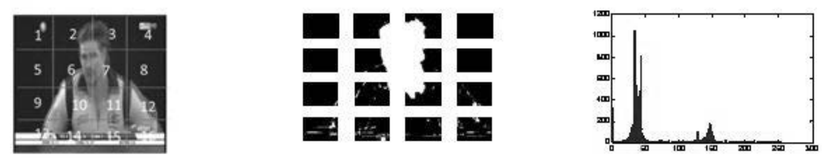
Algorithm 3 to detect Boundary View:

(1) Convert RGB image into Binary image frame.
(2) Take the difference of two consecutive frames to compute motion pixels for each frame.
(3) Divide the image into four equal sized blocks.
(4) Compute MPR(Motion Pixels Ratio) as total number of motion pixels to total no of pixels in the frame, for Blocks $B1$, $B2$ and $B3$.
(5) Calculate GPR(Grass Pixel Ration) for Blocks $B2$, $B3$ and $B4$, as per Algorithm 1.
(6) If MPR(Block $B1$ and Block $B2$) $> Th3$ AND GPR(Block $B3$ and Block $B4$) $> Th4$ OR MPR(Block $B1$ and Block $B2$ and Block $B3$) $> Th3$ AND GPR(Block $B4h4$ OR MPR(Block $B1$ and Block $B2$ AND Block $B4$) $> Th3$ AND GPR(Block $B3$)$> Th4$, Image is considered as Boundary view.

3.4 Close-Up View

Here, we consider color as lower level feature for detection of close up view. If the image is close-up view, we have to find out the location of the face of the person. Generally, during close-up, camera will be zooming out on face of the person. It indicates that face of the person will be available in the centre of the image. Image is divided into 16 equal size block image as shown in Figure $4A$. We can observe that face component of the close-up will appear mostly amongst Block2, Block3, Block6, Block7, Block10 and Block11. Skin pixel can be detected by converting image into ycbcr image and detecting threshold for each of three components, y, cb and cr. If y, cb and cr value of the pixel is above specific threshold, pixel can be considered as skin pixel. Total number of pixel belongs to the skin color to total number of pixels in the image will be calculated as Skin Pixel Ration (SPR).So, skin pixel ratio for above mentioned blocks is calculated. If this ration in any of the specified block is above specific threshold, image can be considered as close-up image. If image is detected as Close-up, it can be again classify as fielder, batsman or umpire.

If a specific frame is detected as a Close-Up view, the player can belongs to Team A or Team



A. 16-block Close-Up View B. 16-Block view of skin image C. Jersey color Histogram

Figure 4 Close-Up view with 16 blocks, it's skin view and Jersey color histogram

B. In cricket match, jersey color of all the player of a same team will be similar. So, once the close-up frame is detected, we need to check the jersey color. From this jersey color, Country of the player can be identified. In Close-up view, we need to check the specific blocks which has Skin Pixel Ratio(SPR) greater than specific threshold. We have to check only those selected blocks to check the jersey color.

Figure $4A$ shows $16-$block close-up view. Figure $4B$ shows skin view of the corresponding image. White pixel indicates presence of skin pixels and black color pixels represents absence of skin pixel. We can see that, skin pixels are available in Blocks $2, 3, 6, 7$ and 11. Figure $4C$ shows hue histogram of block No. 15, as per condition mentioned in Table 1, to detect jersey color of the player [Kolekar et al. [2008]].

Table 1 provides the information about block numbers to be checked to identify jersey color when a specific condition about SPR of specific blocks is satisfied. This table gives the idea about conditions and results to be checked [Kolekar et al. [2008]]. Check the condition of the Table1 using skin block information. If condition is satised, result will give the block number for checking jersey color of the player to identify his team.

Algorithm 4 to detect Close-Up View:

(1) Convert RGB image into YCbCr image frame.

(2) Divide the image into 16 equal sized blocks.

(3) For each pixel, if Y, Cb and Cr component is in a specific range, pixel can be consider as Skin Pixel.

(4) Calculate Skin Pixels for Blocks $2, 3, 6, 7, 10$ and 11.

(5) Compute SPR(Skin Pixel ratio) as ratio of Skin Pixels to total number of pixels in those blocks.

(6) If SPR of any of the 6 blocks specified in step $4 > Th5$, it can be considered as Close Up frame.

Table I. Conditions for block selection for checking jersey color of the player [9]

	Condition	Result (Block Number for jersey color checking)
1	(SP3,SP4, SP5 and SP6) > Th5	Block14and Block15
2	(SP3 and SP5) > Th5	Block14
3	(SP4 and SP6) > Th5	Block15
4	(SP3 and SP4) > Th5	Block10and Block11
5	(SP5 and SP6) > Th5	Block14and Block15
6	SP6 > Th5	Block15
7	SP5 > Th5	Block14
8	SP4 > Th5	Block11
9	SP3 > Th5	Block10

3.5 Crowd View

Edge feature will be helpful for detection of Crowd view. The crowd view is very complicated image as it will have large number of objects in terms of spectators. It can be detected by presence of large number of edges in the image. Edges can be detected by applying canny edge detector on the image frame. So, the presence of crowd in a frame can be detected by performing canny edge detection on the given image .So, crowd image will have more edges compared to other frames. We therefore calculate the Edge Pixel Ratio (EPR) to detect availability of edges in the image. EPR (Ep/n) is total number of edge pixels in the image to total number of pixels in the image. So EPR will be high for crowd image compared to non-crowd image of the video. Crowd can be later on classified as Audience, gathering of Team-A on the ground or gathering of Team-B on the ground.

Figure 5 shows Crowd-view image, Pitch-view image, Field-view image and its corresponding

Figure 5 Crowd-view, Pitch-view, Field-view and its edge detected views

edge detected view. It can be seen that edge detected crowd-view has maximum number of edge pixels. So, it will have high value of EPR.

Algorithm 5 to detect Crowd View:

(1) Convert RGB image into binary image frame.
(2) Apply canny edge detector on the image.
(3) Calculate total edge pixels in the frame.
(4) Compute EPR(Edge Pixel Ratio) as ratio of edge pixels to Total number of pixels in the frame.
(5) If $EPR > Th6$, frame is considered as Crowd View.

3.6 Fielders Gathering View

Fielders will always gather on the field, after every exciting event. Most of the fielders gathering frame will have field as the background. So, to improve the classification accuracy, bins representing dominant green colour for the field can be set to zero. Maximum connected component will be detected from the image. That component represents fielders on the field. Fielders gathering can be further classified into gathering of Team-A player or gathering of Team-B player. This classification can be done with the help of jersey colour. As, the jersey colour of all the players of same team will be similar. So, from the jersey colour of the fielders, we can classify them into players of Team-A or Team-B.

Figure 6A shows view of fielders gathering on the field. Figure 6B shows its corresponding max-

A. Fielders' Gathering view   B. Maximum Connected component   C. jersey color histogram

Figure 6 Fielders view, it's maximum connected component and jersey color histogram

imum connected component. Where, connected component represents fielders on the ground. Figure 6C represents hue histogram of jersey color of the fielders. Using this jersey color information, we can identify team of the players.

Algorithm 6 to detect Gathering View:

(1) In a RGB image, bins representing dominant green color for the field can be set to zero.
(2) Find maximum connected component of the frame.
(3) If maximum connected component of the frame $> Th7$, frame is considered as Fielders View.

3.7 Sky View

When ball goes high in the air, it can be observed that ball will remain in the air at least for a while. At that time, camera man will track the ball and background of the ball will be the sky. During night match sky colour will be obviously black and during day match generally it will be sky blue or white. Sky pixel Ratio (SPR) will be calculated as total number of sky pixels to the total number of pixels in the frame. So, if SPR is greater than threshold value, it can be detected as the sky view.

Figure 7A shows sky view during night match and Figure 7B shows sky view during day match.

A. Sky-View in Night Match   B. Sky-view in day Match

Figure 7 Sky view in night match and day match

Figure 7A shows sky view during night match and Figure 7B shows sky view during day match.

Algorithm 7 to detect Sky View:

(1) In a RGB image, for every pixel, if value of R, G and B component is in a specific range , it is considered as sky pixel.
(2) Compute total sky pixels in the frame.

(3) Compute SPR(Sky Pixel Ratio) as ratio of sky pixels to Total number of pixels in the frame.
(4) If SPR $> Th8$, frame is considered as Sky View.

Table II gives the description of various views detected from the cricket video.

Table II. Cricket match view description

Cricket video view description		
	View	Description
1	Field View	It represents generally fielding by the fielders during match. Catch may also include field view.
2	Pitch View	It shows balling of the baller, strokes by the batsman, on pitch run, umpire decision etc.
3	Boundary View	Boundary might be displayed during Four, Six, Catch or when fielders try to prevent Four or Six. Sometimes a view of nearby audience may also include boundaries.
4	Close-Up View	This view represents close-up of a Batsman, Baller, Fielder or Umpire. Sometimes any personality form the Audience may also appear as close-up.
5	Crowd View	An exciting audience may be displayed after an interesting event like Catch, Four, Six etc.
6	Fielders' Gathering View	During wicket fall, fielders will gather on the ground to celebrate the event.
7	Sky View	When batsmen hit Six, ball will go high in the air and at that time sky view will be displayed.

4. EXPERIMENTAL RESULTS

4.1 Dataset

We have analysed our algorithms with different videos at 25 fps. One of the matches is One day match, 1st inning highlight for Australia VS Srilanka with 8000 frames that runs for approximately 10 minutes. It is a day match. Another match considered is T20, ICC World CupTwenty 20 South Africa 2007, 2nd semi-final India vs Australia with 66,000 frames that runs for approximately 1 hour and it is a night match.

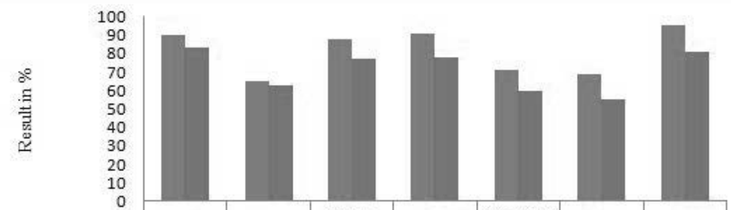
4.2 Implementation Details

Our algorithm has been implemented in MATLAB version 2015a, on a system with windows 10, Intel Core i3 processor and 2GB RAM.

4.3 Result Analysis

Effectiveness of information retrieval techniques can be measured in terms of two parameters, recall and precision. Recall is the ratio of the number of events/clips/shots/frames detected correctly over actual number of events/clips/shots/frames in a given video clip. Precision is the ratio of the number of events/clips/shots/frames detected correctly over actual number of events/clips/shots/frames detected (correctly or incorrectly). Figure 8 shows effectiveness, in terms of precision and recall, of algorithms used to detect various views from cricket video using various low level features.

Our view detection methods work well and shows high rate of precision and recall for all views except fielders and Boundary view. In some of the cases, algorithm is not able to detect boundary view either due to very less audience nearby boundary area or lack of grass nearby boundary area. Fielders view is having less precision and recall because algorithm cannot detect fielders gathering in some of the cases. We have detected stand alone efficiency of our algorithm instead of comparing it with other scheme. Accuracy of view detection can be improved by adding up lower level features used for detection of specific view. We would like to use these detected views in our further work for shot classification and event detection.
Figure 9 shows few of the detected views from different categories.



	Field	Pitch	Close Up	Crowd	Boundar y	Fielders	Sky
■ Precision(%)	90	65	88	91	71	69	95
■ Recall(%)	83	63	77	78	60	55	81

Figure 8 precision and recall for view detection

Fig. 9 Detected views of different category

4.4 Challenges

First of all, processing of video is very time consuming. A video of even five minutes contains 8000-9000 frames. And, single inning of even T20 match can long for 2 hours. So, it is very time consuming to process an entire video for detecting various views from it. Second, Different match can be played for different time duration and at different cities of different countries. This diversity of place and time may lead to varied illumination conditions and texture of grass and soil. So, it becomes difficult to determine the hue range of grass and pitch in Hue Histogram to detect grass and pitch view from the match. Third, Players from various teams may also have different skin complexion. So, skin color must be properly detected to identify close up view from the match.

5. CONCLUSION AND FUTURE WORK

We have proposed an efficient method for cricket match view detection. It achieves competitive performance in precision and recall for each of the view categories. The proposed method can easily be extended to other sports also. It needs to have domain knowledge related to that sport. Our future work would be to improve the precision and recall for view detection system and to use these algorithms for event detection from cricket match.

REFERENCES

Baillie, M. and Jose, J. M. 2003. Audio-based event detection for sports video. In *CIVR*. Vol. 2003. Springer, 300–310.

Chen, H.-S. and Tsai, W.-J. 2014. A framework for video event classification by modeling temporal context of multimodal features using hmm. *Journal of Visual Communication and Image Representation 25,* 2, 285–295.

Chu, W.-T. and Chou, Y.-C. 2015. Event detection and highlight detection of broadcasted game videos. In *Proceedings of the 2nd Workshop on Computational Models of Social Interactions: Human-Computer-Media Communication*. ACM, 1–8.

Ekin, A., Tekalp, A. M., and Mehrotra, R. 2003. Automatic soccer video analysis and summarization. *IEEE Transactions on Image processing 12,* 7, 796–807.

Harikrishna, N., Satheesh, S., Sriram, S. D., and Easwarakumar, K. 2011. Temporal classification of events in cricket videos. In *Communications (NCC), 2011 National Conference on*. IEEE, 1–5.

Jayanth, S. B. and Srinivasa, G. 2014. Automated classification of cricket pitch frames in cricket video. *ELCVIA Electronic Letters on Computer Vision and Image Analysis 13,* 1.

Kapela, R., McGuinness, K., and OConnor, N. E. 2014. Real-time field sports scene classification using colour and frequency space decompositions. *Journal of Real-Time Image Processing*, 1–13.

Kokaram, A., Rea, N., Dahyot, R., Tekalp, M., Bouthemy, P., Gros, P., and Sezan, I. 2006. Browsing sports video: trends in sports-related indexing and retrieval work. *IEEE Signal Processing Magazine 23,* 2, 47–58.

Kolekar, M. H., Palaniappan, K., and Sengupta, S. 2008. Semantic event detection and classification in cricket video sequence. In *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*. IEEE, 382–389.

Kumar, P. and Puttaswamy, P. 2015. The extraction of events and replay in cricket video. In *Emerging Research in Electronics, Computer Science and Technology (ICERECT), 2015 International Conference on*. IEEE, 54–58.

Money, A. G. and Agius, H. 2008. Video summarisation: A conceptual framework and survey of the state of the art. *Journal of Visual Communication and Image Representation 19,* 2, 121–143.

Pallavi, V., Mukherjee, J., Majumdar, A. K., and Sural, S. 2008. Ball detection from broadcast soccer videos using static and dynamic features. *Journal of Visual Communication and Image Representation 19,* 7, 426–436.

Pradeep, K. 2013. Significant event detection in sports video using audio cues. *International Journal of Innovations in Engineering and Technology (IJIET) 3,* 1.

Pramod Sankar, K., Pandey, S., and Jawahar, C. 2006. Text driven temporal segmentation of cricket videos. *Computer Vision, Graphics and Image Processing*, 433–444.

Raventos, A., Quijada, R., Torres, L., and Tarrés, F. 2015. Automatic summarization of soccer highlights using audio-visual descriptors. *SpringerPlus 4,* 1, 301.

Rui, Y., Gupta, A., and Acero, A. 2000. Automatically extracting highlights for tv baseball programs. In *Proceedings of the eighth ACM international conference on Multimedia*. ACM, 105–115.

Sigari, M.-H., Soltanianzadeh, H., and Pourreza, H. R. 2015. Fast highlight detection and scoring for broadcast soccer video summarization using on-demand feature extraction and fuzzy inference. *International Journal of Computer Graphics 6*.

Su, P.-C., Lan, C.-H., Wu, C.-S., Zeng, Z.-X., and Chen, W.-Y. 2013. Transition effect detection for extracting highlights in baseball videos. *EURASIP Journal on Image and Video Processing 2013,* 1, 27.

Tang, H., Kwatra, V., Sargin, M. E., and Gargi, U. 2011. Detecting highlights in sports videos: Cricket as a test case. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*. IEEE, 1–6.

Tien, M.-C., Wang, Y.-T., Chou, C.-W., Hsieh, K.-Y., Chu, W.-T., and Wu, J.-L. 2008. Event detection in tennis matches based on video data mining. In *Multimedia and Expo, 2008 IEEE International Conference on*. IEEE, 1477–1480.

Vijayakumar, V. 2012. Event detection in cricket video based on visual and acoustic features. *Journal of Global Research in Computer Science 3,* 8, 26–29.

W, L. C. 2011. Baseball game highlight and event detection. *Visual Communications and Image Processing*, 1–4.

Zhang, S. 2015. Research on effective field lines detection and tracking algorithm in soccer videos. *International Journal of Multimedia and Ubiquitous Engineering 10,* 7, 75–84.

Zhong, D. and Chang, S.-F. 2001. Structure analysis of sports video using domain models. In *IEEE ICME.* Citeseer.

Zhong, D. and Chang, S.-F. 2004. Real-time view recognition and event detection for sports video. *Journal of Visual Communication and Image Representation 15,* 3, 330–347.

Hetal Chudasama received her Bachelor degree in Informaion Technology in 2002 from C U Shah college of engineering and Technology,Saurashtra University, India.She received her master degree in field of Computer engineering in 2009 from Sardar Patel University. At present, she is working as a Assistant Professor at Informatin Technology Department, MBICT,GTU, Anand, India.She is having acad emic experience of 11 years. Her area of research interest is in the field of Image Processing.

Dr Narendra M Patel received his BE degree in Electronics Engineering from M S University, Baroda in 1993 and ME degree from M S University, Baroda in 1997.He received PhD degree from SVNIT, Surat in 2012. He is currently working as Associate Professor in Computer Engineering Department, BVM Engineering College, V V Nagar, Gujarat. His research interests include Digital Image Processing, Real time operating systems, Distributed systems and Computer Graphics. He authored more than 50 papers which are published in Prominent international and conference proceedings. He has guided more than 50 ME dissertations in Computer engineering.