SEMI-AUTOMATED CRICKET BROADCAST HIGHLIGHT GENERATION

Sahil A. K. Ramlogan[1*], Akash Pooransingh[2] and Daniel J. Ringis[3]

[1,2]Faculty of Engineering, The University of the West Indies, Trinidad
[3]Faculty of Electronic & Electrical Engineering, Trinity College Dublin, Ireland
[1]Email: sahil.ak.ramlogan@gmail.com*
[2]Email: akash.pooransingh@sta.uwi.edu
[3]Email: ringisd@tcd.ie

Abstract*:* This paper provides an improvement to automated cricket highlight identification from full broadcasts. Further refinement is provided through semi- automated user verification. This visual result is synchronized with an automated text extraction process. The visual result was successfully extracted at a rate of 97.5% using the ORB function. With user validation a false detection rate of 0% was realized. The metadata from the commentary was then successfully extracted and combined with the visual results to allow for easy searching of the highlight video using any search field.

Keywords: *ORB, BRS, Automated Highlight Generation, Cricket Broadcasts.*

https://doi.org/10.47412/CIOE9513

1. Introduction

Professional sports is a big business that has grown rapidly over the last three decades. This growth has been mainly facilitated by the increased revenue from commercial deals with sponsors and various broadcasting networks. The biggest change in competitive professional sports this century is specifically the use of data science and analytics to win games. This increased investment in sport has also allowed for increased developments in technology. It allows coaches to coach more efficiently thanks to detailed training reports and instantaneous correction. In particular, video analysis technologies are key to building a competitive edge, more effective practices, scouting upcoming opponents, injury prevention, and breaking down game film [1]. Coaches can use video footage to show and correct mistakes instantly, motivate players by building up videos over time and showing them their progress, as well as prevent injuries by correcting bad habits in form and technique.

Using video analysis in coaching means that specific scenes must be picked out from match broadcasts. The development of T20 cricket into a multi-billion dollar industry has increased the importance of player statistics and analysis. Typically, this analysis includes the careful study of each player based on previous broadcast footage. Unfortunately, the key points from a single match broadcast occur within less than 10% of a four to five hour long video. Therefore, extracting these key points is time consuming and costly. With specific focus on cricket, T20 matches take approximately three hours to be completed. Cricket match footage would include sports action (the ball being delivered, played, fielded and returned) and time between deliveries, overs and innings that may include highlights or graphics.

This paper proposes a multimodal approach for smart highlight generation. The paper proposes an improved visual method for highlight detection for cricket from full broadcasts. The paper also proposes to incorporate independent text-based commentary to provide metadata input to allow easy searching of the video highlights.

2. Current Methods

Lipton [2] explores the use of Intelligent Video Surveillance (IVS). This allows computers to monitor real time feeds to detect suspicious activities and alert the relevant authorities. The basis of the IVS is known as an ObjectVideo. It watches video streams and extracts the descriptions of the objects within the frame. Thus, any changes in these descriptors allow the algorithm to detect crimes such as vandalism, theft, property intrusion, suspicious packages or suspicious loitering. This type of technology shows how automated video detection can be applied in the real world. However, Lipton [2] concludes that other infrastructures must be put into place to support the video detection system.

In a similar way, automatic highlight extraction from sports broadcast footage can be extracted using a combination of video analysis in conjunction with audio or textual analysis. Rui, Gupta and Acero [3] explored the use of only audio to extract video highlights. The paper used the excitement of the commentators' voices as well as the loudness of the crowd. The event that was used to initiate the audio analysis was the ball hit detection. It was used to refine the results due to the noisy nature of the audio samples for the crowd and the commentators. This method yielded approximately an 80% highlight detection rate when compared to what humans would detect as highlights.

Another automatic video extraction algorithm was developed by Sudhir, Lee and Jain [4]. This algorithm used video analysis to track tennis players positions and movement on the court. This data was then used to determine what events were taking place in the video and extract and store the data accordingly. The method used in this paper was one that was heavily reliant on tracking the tennis court lines and being able to estimate the angle of the camera with respect to the court. In cricket, however, the lines at the edge of the pitch are sometimes not as visible or they wear away during the game. Furthermore, due to the different sizes of stadia, the angle at which the camera is being used might be very different from one stadium to the next. Therefore, this is not a viable technique to be used in this project.

Sankar, Pandey and Jawahar [5] segmented videos based on the online commentary for the match as well as shot cut detection. The shot cut detection was used to detect different shots and the probability whether the shot was a ball event or not was calculated based on the information that was contained in the corresponding line of commentary. The detection accuracy of this method proved to be 82.44%. The scenes were then annotated using the commentary by synching the balls to the corresponding line of commentary.

Feature matching is used for many computer vision solutions, such as object recognition [6]. Current methods are computationally costly to calculate the key points and descriptors. This paper compares the SIFT and SURF algorithms to the (Oriented FAST Rotated BRIEF) ORB algorithm. This algorithm is rotation invariant and resistant to noise. The ORB as suggested by its name is a combination of the two descriptors known as Fast and Brief. The fast method is efficient and finds reasonable corner key points. However, the orientation component is non-existent. The brief method uses simple tests between pixels. It computes the descriptors efficiently but performs poorly with rotation. After the tests were conducted, the ORB was found to be two orders of magnitude faster than SIFT, performing just as accurately.

Ringis and Pooransingh [7] investigated the use of the ORB method to automatically detect the Bowler Run-up Sequence (BRS). This sequence is considered to be the most significant broadcast view since it determines the start of a delivery. The method required a training frame to be provided for each inning. Analysis was done on a frame by frame basis. This method proved to be the least computationally expensive for BRS detection. The results from this paper was a BRS detection rate of 98%, with a false detection rate of 6%. The BRS detection percentage as well as the low false detection rate, makes this method ideal for BRS identification and video extraction.

3. Ingestion and Highlight Generation for Cricket Broadcasts

Cricket broadcasts consists of a large production with over 50 camera views placed around the ground. However, the main view for highlights are considered to be the Bowler Run up Sequence(BRS)as described in[8].This view of the broadcast is contained in less than 1/16th of the entire broadcast and comprises the main sequence of play that is essential for recording (video and text). Ringis and Pooransingh [7] showed how visual cues only using the ORB method [6] could quickly train and recognize highlights of cricket matches. The method in [7] gave a 98% detection rate. This paper improves on that result by considering multiple candidates for ORB training as compared to the minimal training done in [7].

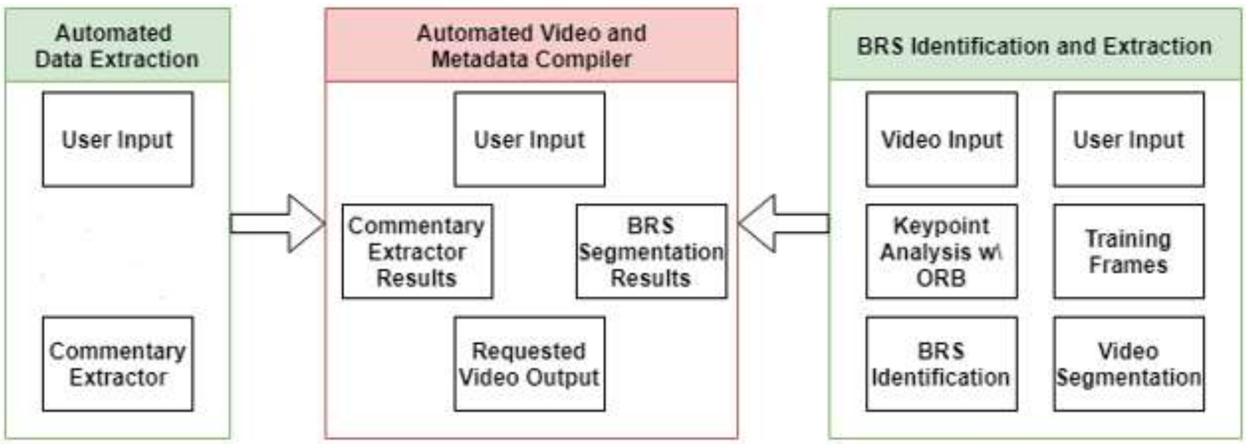

Figure 17: Ingestion and Highlight Generation Overview. The Automated Data Extraction and BRS Identification and Extraction modules feed into the Automated Video and Metadata Compiler

Text based commentary provide a summary of each of the visual event. Usually for cricket matches, this is recorded independently and broadcast live to users without the reliance of a video feed. Popular platforms include espncricinfo.com and cricbuzz.com. Text commentary provides valuable metadata that could be used in the annotation of visual highlights [5]. This paper provides a text-based ingestion combined with the improved ORB method for visual highlight detection for a robust and searchable highlight generation system as shown in Fig. 1. The ingestion and highlight generation system proposed here consists of three main components, Text data extraction, BRS identification, Video and metadata compiler.

3.1 Automated Text Extraction

Cricbuzz.com offers a repository of text-based commentary for specific cricket matches broadcast live on that site. The commentary from this website is copied to a text file to allow for further processing and extraction of data. Based on patterns found within the commentary, this allowed the relevant data to be extracted and stored, as discussed in section 4.1.

3.2 Semi-Automated BRS Identification

This paper seeks to improve on the ORB method for BRS identification in [7] where the ORB method was used for fast training for BRS detection. In this paper, a more measured approach was taken to training to improve the detection rate and reduce the false detection rate. In [7], only a single frame was used for training the BRS per inning. Upon analysis of the data, there is significant variation in the BRS features depending on the bowler and batsman type within the frame.

A single frame of the BRS was used for training for each of six (6) types of variation as follows:

- Right arm fast bowler over the wicket to right hand batsman
- Right arm fast bowler over the wicket to left hand batsman
- Left arm fast bowler over the wicket to right hand batsman
- Left arm fast bowler over the wicket to left hand batsman
- Spin bowler over the wicket with respect to a right hand batsman
- Spin bowler around the wicket with respect to a right hand batsman

Figure 2 shows an example of the variation of training frames. False detections would cause a failure for the combination of video and extracted text-based metadata. To eliminate the usage of false detection results, the user would verify each candidate BRS from an automated prompt.

Figure 18: Sample of Training Frame with a left arm fast bowler over the wicket to a left hand batsman

3.3 Video and Metadata Compiler

This process synchronizes the results of the automated text extraction to the video based semi-automated BRS identification. Once the sequence of video BRS highlights matches the text sequence, each video highlight can be encoded with metadata. This metadata comprises the Match, Video and Innings ID which combine to make a unique Storage ID, as well as information about the batsman, bowler and event.

4. Results

Testing was done on played matches of the Indian Premiere League (IPL) 2014 T20 tournament. Text and video broadcast footage were sourced from independent repositories.

4.1 Automated Text Extraction

Text-based commentary for matches can be found on many cricket websites such as Cricbuzz.com. This commentary was stored in a text file. The text within the commentary contained more information than was necessary for the database that was constructed. Therefore, an algorithm was set up, using the Regular

Expressions library, to extract the information using familiar patterns found in the text. Below is an example of some lines extracted from the commentary text file.

17.1 Zaheer to Kallis, 1 run, length delivery outside off stump, Kallis drives it to deep extra-cover.

16.6 Malinga to M Pandey, out Bowled!! Malinga gets his revenge! Not quite a perfect yorker this, Pandey tries to squeezes it out, misses the ball and the off-pole goes cartwheeling. Anyway, Pandey has done his part in getting his side to a good position for a big score! M Pandey b Malinga 64(53) [4s-6 6s-2]

The pattern of the commentary forms the basis of generating the metadata to be linked to the video based BRS result.

4.2 Semi-Automated BRS Identification

The match broadcast footage included the first 20 games of the IPL 2014 T20 tournament. The videos were then edited into two separate videos first innings and second innings. The training frames for each innings in each game are different since there is a change in uniform colour for the bowling and batting teams. A single frame of the BRS was selected for training for each of the six (6) types of variation as described in the previous section. Table 1 shows the results of a single inning of a match using the semi-automated BRS identification as compared to the method described in [7]. The implementation of the method in [7] provided a low BRS detection rate of 46.7% and a high false detection rate of 37% for the sample dataset of 121 deliveries. The proposed use of six frames for training showed a marked improvement in the BRS detection rate of 97.5% with a false detection rate of 33.9%. The false detection rate was reduced to 0% with user verification while maintaining a BRS detection rate of 97.5%.

Table 3: Comparison of BRS Identification [7] Versus Semiautomated BRS Identification

	Single Frame BRS	**Six Frame BRS**	**Six Frame BRS with Verification**
Correct Matches	58	121	121
False Matches	32	62	0
Deliveries Missed	66	3	3
Matches Found	86	183	121
BRS Detection Rate	46.8%	97.5%	97.5%
False Match Rate	37%	33.9%	0%

4.3 Video and Metadata Compiler

The improved BRS identification was synchronized with the results of the automated text extraction. The extracted text information was stored in a two-dimensional array, the first index indicating the video ID and the second index indicating delivery, bowler, batsman and event for the corresponding video ID. The video ID was linked directly to the BRS results. A table was created using MySQL and linked to the C# program to allow access to a library of functions that could be used to connect to the MySQL server to run SQL commands. This component was also designed to allow a user to search for information from one or more columns in the table. After the parameters for the search were entered, the requirements were changed to a SQL command to retrieve the information. This information was then displayed to the user in table form with the video ID clickable for immediate playback. Figure 3 shows a sample of the video and metadata compiler.

StorageID	MatchID	VideoID	Innings	Delivery	Bowler	Batsman	Event	Ground	Video_Address
IPL201400111	IPL2014001	1	1	0.1	Zaheer	Gambhir	0		D ECNG 3020 Raw Match FootageBRSBall1.avi
IPL2014001110	IPL2014001	10	1	1.3	Malinga	Gambhir	0		D ECNG 3020 Raw Match FootageBRSBall107.avi
IPL20140011100	IPL2014001	100	1	16.1	Malinga	M Pandey	0		D ECNG 3020 Raw Match FootageBRSBall78.avi
IPL20140011101	IPL2014001	101	1	16.2	Malinga	M Pandey	6		D ECNG 3020 Raw Match FootageBRSBall79.avi
IPL20140011102	IPL2014001	102	1	16.3	Malinga	M Pandey	0		D ECNG 3020 Raw Match FootageBRSBall8.avi
IPL20140011103	IPL2014001	103	1	16.4	Malinga	M Pandey	leg bye		D ECNG 3020 Raw Match FootageBRSBall80.avi
IPL20140011104	IPL2014001	104	1	16.5	Malinga	M Pandey	0		D ECNG 3020 Raw Match FootageBRSBall81.avi
IPL20140011105	IPL2014001	105	1	16.6	Malinga	M Pandey	out		D ECNG 3020 Raw Match FootageBRSBall82.avi
IPL20140011106	IPL2014001	106	1	17.1	Zaheer	Kallis	1		D ECNG 3020 Raw Match FootageBRSBall83.avi
IPL20140011107	IPL2014001	107	1	17.2	Zaheer	Uthappa	1		D ECNG 3020 Raw Match FootageBRSBall84.avi

Figure 19: Sample of the video and metadata compiler result

5. Conclusion

This paper provides an improvement to the BRS identification of video highlights as compared to [7] through a six-frame training step. Further refinement is provided through semi- automated user verification. This visual result is synchronized with an automated text extraction process that provides valuable metadata to allow for easy searching of the BRS video using any search field. This paper provides a state-of-the-art method for automated ingestion and smart highlight generation. This tool is valuable in the creation of a searchable player database that may allow for deeper analytics of player performance in the future.

References

[1] FraserCarson, "Utilizingvideotofacilitatereflectivepractice:Developing sports coaches," International Journal of Sports Science & Coaching, vol. 3, no. 3, pp. 381–390, 2008.

[2] A. J. Lipton, "Keynote: intelligent video as a force multiplier for crime detection and prevention," in The IEE International Symposium on Imaging for Crime Detection and Prevention, 2005. ICDP 2005., June 2005, pp. 151–156.

[3] Yong Rui, Anoop Gupta, and Alex Acero, "Automatically extracting highlights for tv baseball programs," in Proceedings of the Eighth ACM International Conference on Multimedia, New York, NY, USA, 2000, MULTIMEDIA '00, pp. 105–115, ACM.

[4] G. Sudhir, J. C. M. Lee, and A. K. Jain, "Automatic classification of tennis video for high-level content-based retrieval," in Proceedings 1998 IEEE International Workshop on Content-Based Access of Image and Video Database, Jan 1998, pp. 81–90.

[5] K. Pramod Sankar, Saurabh Pandey, and C. V. Jawahar, "Text driven temporalsegmentationofcricketvideos," inProceedings of the 5th Indian Conference on Computer Vision, Graphics and Image Processing, Berlin, Heidelberg, 2006, ICVGIP'06, pp. 433–444, Springer-Verlag.

[6] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in 2011 International Conference on Computer Vision, Nov 2011, pp. 2564–2571.

[7] D. Ringis and A. Pooransingh, "Automated highlight generation from cricket broadcasts using orb," in 2015 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM), Aug 2015, pp. 58–63.

[8] A Kokaram and P Delacourt, "A new global motion estimation algorithm and its application to retrieval in sports events," 02 2001, pp. 251 – 256.