AUTOMATIC CRICKET HIGHLIGHTS EXTRACTION USING EVENT DRIVEN AND EXCITATION BASED FEATURES

by

SATHISH KUMAR S 2015103530
SIVARANJANI P 2015103534
BELAL R 2015103007

A project report submitted to the

FACULTY OF INFORMATION AND COMMUNICATION ENGINEERING

in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

ANNA UNIVERSITY, CHENNAI – 25

MAY 2019

BONAFIDE CERTIFICATE

Certified that this project report titled **AUTOMATIC CRICKET HIGHLIGHTS EXTRACTION USING EVENT DRIVEN AND EXCITATION BASED FEATURES** is the *bonafide* work of **SATHISH KUMAR S (2015103530)**, **SIVARANJANI P (2015103534)** and **BELAL R (2015103007)** who carried out the project work under my supervision, for the fulfillment of the requirements for the award of the degree of Bachelor of Engineering in Computer Science and Engineering. Certified further that to the best of my knowledge, the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or an award was conferred on an earlier occasion on these or any other candidates.

Place: Chennai
Date:

Dr. P. Uma Maheswari
Associate Professor
Department of Computer Science and Engineering
Anna University, Chennai – 25

COUNTERSIGNED

Head of the Department,
Department of Computer Science and Engineering,
Anna University Chennai,
Chennai – 600025

ACKNOWLEDGEMENT

This work would not have been possible without the continued support in terms of discussions, encouragements and suggestions of many people.

First of all, we would like to express our deepest gratitude towards our project guide **Dr. P. Uma Maheswari** for her guidance, support, motivation and encouragement at every phase of the project. We owe her for harnessing our potential and bringing the best out of us. Without her support through every phase of the project, we could never have it to this extent.

We are grateful to **Dr. S. Valli**, Head of the Department, Computer Science and Engineering, Anna University, Chennai-25, for providing a platform and extending the facilities for our project.

We express our thanks to the panel of reviewers **Dr. S. Chitrakala, Dr. Angelin Gladston and Dr. S. Renugadevi** for their valuable suggestion and critical reviews throughout the course of our project.

We wholeheartedly thank our family, our parents, our friends for their unwavering support during our work.

Sathish Kumar S **Sivaranjani P** **Belal R**

ABSTRACT

Cricket is one of the most watched game in the world. People are attracted to watch interesting events of the match but creating such sports highlights manually is a tedious job. We propose a model to automatically generate highlights of the cricket match. Key events of the cricket match are fours, sixes, wickets. Here highlights of the cricket match is generated by recognizing excitation and event driven features. First, the replay and advertisements are detected and removed. Then event for each live frames are detected. Shot boundary is formed based on the event tags. Scores for each shot is recognized and compared to identify highlights. Excitement of commentators and crowd cheering is analyzed for detecting highlights. Both the highlights from scorecard extraction and excitation detection are combined for final highlights. The system is tested with different kinds of input to demonstrate the effectiveness. User study is also done to evaluate the system in real life scenarios.

Index Terms: OCR, CNN, Highlights, Scorecard.

ABSTRACT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

OCR Optical Character Recognition

CNN Convolutional Neural Network

BBL Big Bash League

CHAPTER 1

INTRODUCTION

1.1 PROBLEM DOMAIN

Computer vision is a field of computer science that works on enabling computers to see, identify and process images in the same way that human vision does, and then provide appropriate output. It is like imparting human intelligence and instincts to a computer. In reality though, it is a difficult task to enable computers to recognize images of different objects. Computer vision is closely linked with artificial intelligence, as the computer must interpret what it sees, and then perform appropriate analysis or act accordingly. Computer vision's goal is not only to see, but also process and provide useful results based on the observation. For example, a computer could create a 3D image from a 2D image, such as those in cars, and provide important data to the car and/or driver. For example, cars could be fitted with computer vision which would be able to identify and distinguish objects on and around the road such as traffic lights, pedestrians, traffic signs and so on, and act accordingly. The intelligent device could provide inputs to the driver or even make the car stop if there is a sudden obstacle on the road. Computer vision's aim is to enable computers to perform the same kind of tasks as humans with the same efficiency.

In this project, we have developed a system that would automate the cricket highlights extraction using both event and excitation based features. Cricket is the most watched game in the world. Thus considering

the time and labour consumption, we chose to develop this highlights extraction system.

1.2 PROBLEM STATEMENT

Nowadays digital video plays an important role in every days life and due to widely used low cost storage media, the volume of digital video tends to be very large and variety of available video data makes the search and retrieval of content a more and more difficult task. The amount of information generated in todays society is growing exponentially. Video images are voluminous, redundant and their overall contents cannot be captured at a glance. It is essential to help user to provide more compact, interesting video content with narrow bandwidth. In order to meet this need, video summarization is needed. A complete sports video whose duration is much longer like cricket is monotonous which most viewers dont prefer. Hence Sports Highlights extraction was brought to light. This work generates highlights of cricket match video.

1.3 PROBLEM DESCRIPTION

Given a cricket match video as input, the system should generate highlights. The highlights generated by the system should contain only the key events like 4s, 6s, wickets and milestones.

1.4 SCOPE

A cricket match has a minimum duration of 6hrs. People who has missed the match likes to watch only the key events of that match. This emphasizes the importance of highlights. Be it any sports, highlights are important for broadcasters in order to increase the viewership. Automatic highlights extraction is very less and hence this system.

1.5 CHALLENGES

1.5.1 Match Timing

The day time match will have high contrast and night match will have low contrast. This variations affects recognition of the scorecard.

1.5.2 Event Detection

Event detection for every frame in the cricket match is highly challenging. If not classified properly, it affects further classification in the highlights.

1.5.3 Replays and Advertisements

Key events like Wicket, Six, Four etc in Cricket Match are followed by replay of that event. This replay are the repetitive video shot of the actual event.

1.6 CONTRIBUTIONS

1. **Problem:** Given a cricket video match, generate highlights video. We propose a system that identifies the highlights in the input video.
2. **Problem:** Given a cricket video match with advertisements and replays,Identify these events.
   We propose a method that classifies the video as advertisements and replay or live frames.

CHAPTER 2

RELATED WORK

In [8] this approach automatically extracts highlights from sports videos based on multimodal sport-independent excitement measures, including audio analysis from the spectators and the commentator, and visual analysis of the players. Based on that, the system for auto-curation of golf and tennis highlight packages, which was demonstrated in three major golf and tennis tournaments in 2017. This method fuses information from the players reactions (action recognition such as high-fives and fist pumps), players expressions (aggressive, tense, smiling and neutral), spectators (crowd cheering), commentator (tone of the voice and word analysis) and game analytics to determine the most interesting moments of a game. It also identifies the start and end frames of key shot highlights with the players name and the hole number using OCR. Excitement score is computed based on audio tone analysis and speech to text analysis. Players expression are considered as valuable because detecting neutral face helps rejecting false positives.

In [4] here the replays are detected and extracted to form highlights. A replay segment is considered as a clip sandwiched in gradual transitions and absence of score-caption. The proposed method is robust to broadcasters variation, sports category, score-caption design, camera variations, replay speed, and logo design, size, and placement. The proposed algorithm does not rely on logo template recognition for replay detection, which makes it computationally efficient. The proposed system consists of 2 steps ie gradual transition and scorecard detection.

In gradual transition, candidate replay segment is done. The start and the end frame of gradual transition is detected and extracted. Score-card detection is done by converting the videos into frames then into grayscale, converting them into binary image,finally given to OCR for score recognition. Effectiveness of the proposed method is evaluated on a diverse set of real world videos. Experimental results indicate that the proposed system achieves average detection accuracy rate $>94\%$. It has been observed that under severe uneven illumination, performance of the proposed system degrades marginally.

In [6] BBN based framework is used for annotating the exciting clips with high-level semantic concept-labels. These annotated exciting clips are selected based on their importance degree for generating highlights for sports video sequences. Events are detected and classified using hierarchical tree. Concepts are collection of events. Some low events which are not important are removed from the concepts. Therefore the concept size is reduced. Automatic highlights for soccer video sequences and compared the performance with BBC highlights. The system is applicable to other types of sports videos with similar game structure such as basketball, volleyball, baseball, and cricket. There are other relevant low-level features such as camera motion, parallel line detection, caption detection which might provide complementary information and may help to improve performance of proposed approach by increasing number of events. The proposed framework recognizes energizing clasps in light of sound components and after that arranges the individual scenes inside the clasp into occasions. A probabilistic Bayesian conviction arrange in view of watched occasions is utilized to appoint semantic idea marks to the energizing clasps in soccer video groupings. The named clasps are chosen by level of significance to incorporate into the highlights.

In [9] approach, the model considers both event-based and excitement-based features to recognize and clip important events in a cricket match. Replays, audio intensity, player celebration, and playeld scenarios are examples of cues used to capture such events. The top four events are milestones, wickets, fours and sixes are recognized using event based approach. Replay are detected by CNN and SVM. Score-card detection is done using OCR. By using detected scorecard, score is extracted. Playfield scenarios like batting, bowling, umpire, non striker are detected using CNN and SVM. Using loudness as audio feature key events like dropped catches are detected. The audio level intensity is used to detect the excitement in the video. Player celebration are considered as milestones and they are detected using SVM,CNN. The system is evaluated by comparing the generated highlights with the official highlights. The subjective evaluation or user study is also done.

In [2]this approach, the highlights is extracted based on audio-motion integrated cues. The likelihood models measure the likeliness of low-level audio features and motion features to a set of predened audio types and motion categories, respectively. Our experiments show that using the proposed likelihood representation is more robust than using low-level audio/motion features to extract the highlight. With the proposed likelihood models, we then construct an integrated feature representation by symmetrically fusing the audio and motion likelihood models. Finally,we employ a hidden Markov model (HMM) to model and detect the transition of the integrated representation for highlight segments.

In [5], 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. At level-5b, swarm section is grouped into onlooker and players' social event utilizing shading highlight.

In [7] , it automatically generates highlights of game sequences so that selections of events can be located and played back. Excitements levels are gathered from the audio energy and short time zero crossing. Caption recognition is carried out using sum of absolute difference based caption recognition model. Method reduces manual processing, enables the generation of personalized highlight and also can be used for Content Based Video Retrieval. The approach seems effective and around 80-85% accurate in practical tests. It is necessary to give complete set of characters of the channel and prior knowledge of the caption location is required.

In [1], the algorithm uses textual information extraction method. This textual information is extracted from each frame by first detecting score bar, then converting the text on this score into a sentence like structure based on OCR. Once textual information is at our disposal, difference in score and wickets is detected to get information about 4, 6 or fall of a wicket. This forms the basis of event detection. Now those frames are included in which event has occurred and combined together to generate highlights.

A generic method for sports video highlight selection is presented in [3]. Processing begins where the video is divided into short seg-

ments 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.

In [10], the sports summarization is done using highlights and play breaks. Combining the audio and visual features provides more accurate results. This method first detects the highlights using whistle sound and excitement of the crowd and commentators and then text display is re checks whether it is a highlight. These annotations are evaluated manually. The highlights extracted using excitement is stored in the database (only the start and the end frames position are stored). The benefit of this method of integrating play- break and highlight scenes, combines the strengths of both the methods. Hierarchical structure is used to organize play, break and highlight scenes. Text detection is detecting the scorecard in the video of the match. Scorecard contains player name, score. For example goals in the soccer. Using excitement is a good measure to identify the highlights since the excitement in the crowd is sustained during goal celebration and commentators are also excited about the goal. This is done for swimming and soccer. When whistle , excitement, and text detection used, 85% to 100% of the highlights can be detected.

CHAPTER 3

REQUIREMENTS ANALYSIS

3.1 FUNCTIONAL REQUIREMENTS

The system outputs the highlights video for a given Cricket input video. The output video should adhere to the following requirements:

- The output video should contain all the key events such as wickets, boundaries and other excitement events
- The system must be able to detect the presence and absence of scorecard in the video
- The output video should not contain replays/advertisements
- The system must be optimized for time and space complexities
- The system must be able to classify the input video and annotate with its events

3.2 NON FUNCTIONAL REQUIREMENTS

3.2.1 User Interface

There must be a simple and easy to use user interface. The intermediate results such as video classification, OCR results are printed in the CSV file for easy perception. The output video clips are concatenated and stored in a new output file by using movie editing tools.

3.2.2 Hardware

A system capable of running a python program with a minimum of 2GB RAM is required.

3.2.3 Software

- Operating System: Windows
- Programming Language: Python
- Tools: Opencv, Keras, Moviepy, Pytesseract.

3.2.4 Performance

The system must be optimized, reliable, consistent and available all the time.

3.3 CONSTRAINTS AND ASSUMPTIONS

3.3.1 Constraints

- The system would work only for the BBL Cricket match videos.
- More than 2500 images are generated manually for CNN classification in order to attain better results.
- The output video will work effectively only when BBL cricket match video
- If the input video doesn't contain the key events as per the precoded criteria No output video will be generated.
- The Accuracy of the system is calculated only by the analysis of confusion matrix.

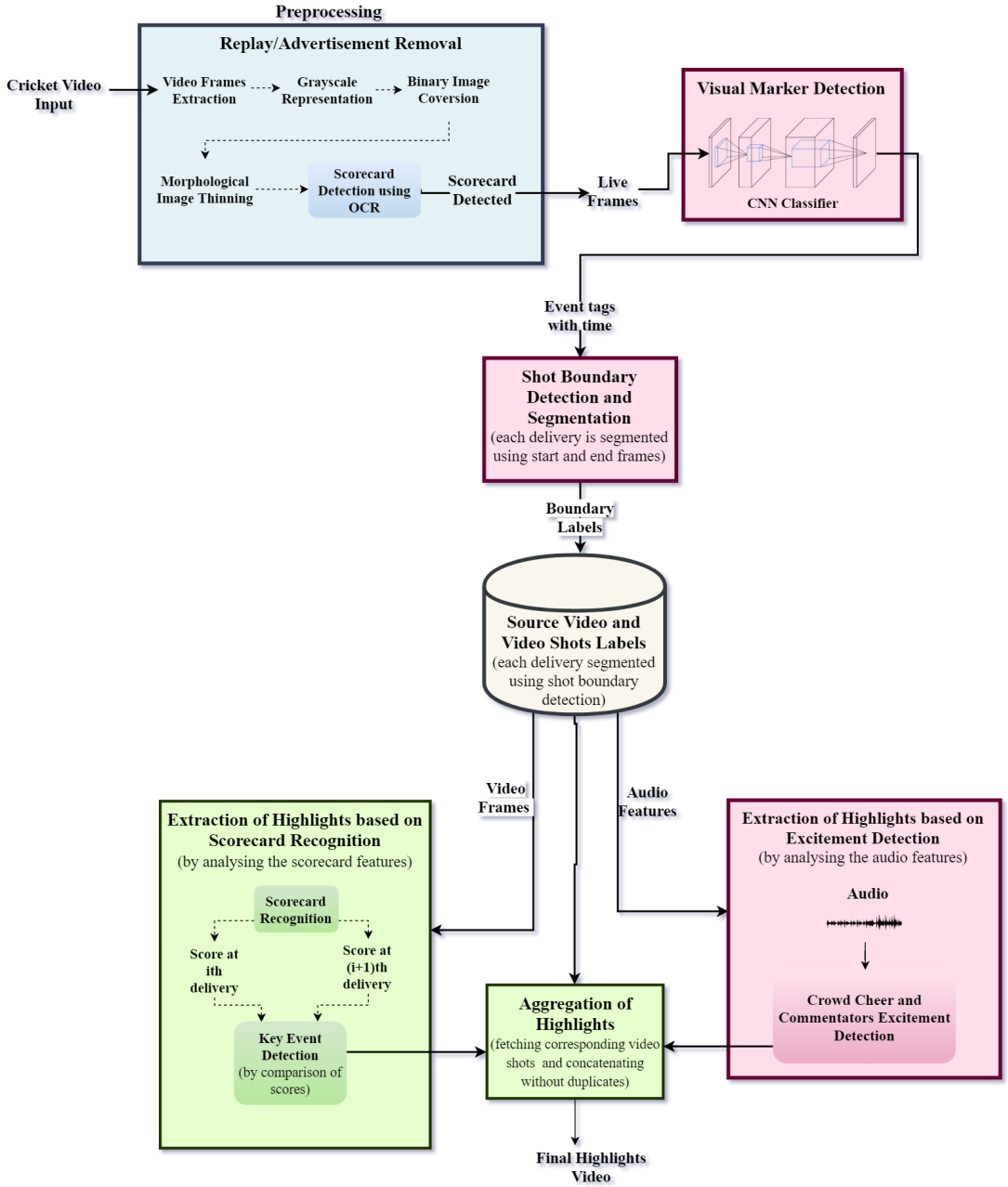
3.3.2 Assumptions

- The Fours/Sixes, wickets and excitement shots are the only highlights
- All the key events are detected only by analyzing the scorecard and the audio.
- OCR detects and recognizes all the scorecards effectively
- The Scorecard will display the exact score of that particular delivery without any errors.

CHAPTER 4

SYSTEM DESIGN

4.1 ARCHITECTURE



4.2 MODULE DESIGN

4.2.1 Preprocessing

Replays are the repetitive video shots of the key events in the cricket match. Advertisements are unrelated video shots for highlights generation. Therefore, Replays and advertisements are the unwanted shots needed to be removed. This module is used for detection and removal of replays and advertisements by checking the scorecard in the video frames. Absence and presence of the scorecard are used to detect replay and live frames, respectively. Replays and advertisements are categorized as frames without the scorecard. The source video is converted into frames. In order to detect the scorecard by OCR, the frames are converted to gray scale image then to binary image. OCR recognizes the scorecard from the binary image. The frames with the scorecard is live frames which is given to visual marker detection along with their time.

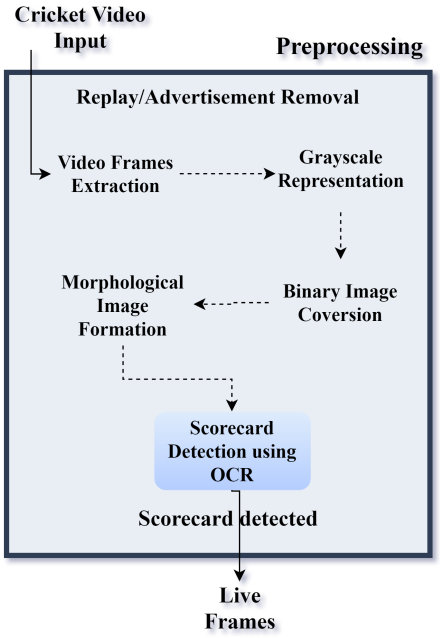


Figure 4.1 Preprocessing

```
Input:Source Input Video
Output:Live Frames
frames[] := generate_frames(Source video)
  for each frame in frames[] do
    image := grayscale(frame)
    image := binary_conversion(image)
    text := OpticalCharacterRecognition(image)
    if is_scorecard(text): then
      mark_as_live(frame)
    else
       mark_as_replay(frame)
    end if
  end for=0
```

Algorithm 1: Preprocessing

4.2.2 Visual Marker Detection

Events detection is important for shot boundary detection and segmentation. This module is detects events like batting, bowling, crowd, interviews, commentators etc. from the live frames using trained CNN classifier. The CNN is trained for bowling, commentators, interview, field view, crowd, stump view, umpire. The trained model is loaded to predict these classes. The CNN classifier predicts the events for each live frames. The event tag along with the time is given to the shot boundary detection and segmentation.

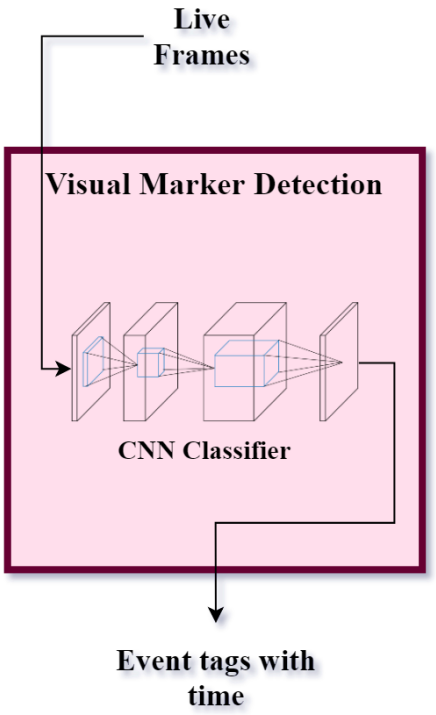


Figure 4.2 Visual Marker Detection

Input:Live frames

Output:Event tags with time

model :=load_trained_cnn_classifier()

for `each frame in frames[]` **do**

`event_tag :=model.classify(frame)`

`time := fetch_frame_time(frame)`

`out_to_CSV_file(event_tag,time)`

end for=0

Algorithm 2: Visual Marker Detection

4.2.3 Shot Boundary Detection and Segmentation

Unwanted events like crowd view, commentators, interviews, etc. are removed. Start/End frames are detected and segmented as video shots. The bowling event is considered as the start of the delivery. Every delivery is till the start of the next delivery. The frames with unwanted

events are removed. Using start and end of the delivery, shots are identified. The labels containing start and end time of the delivery is stored for further use.

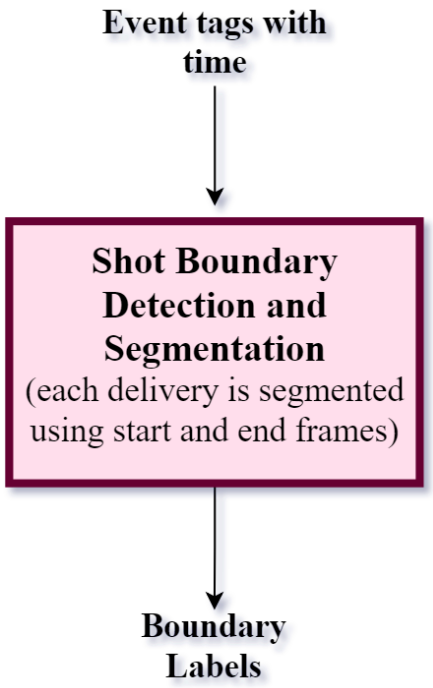


Figure 4.3 Shot Boundary Detection

```
Input:Event tags
Output:Segmented delivery labels
```
$i \leftarrow 1$

for each event in event[] **do**
if is_bowling(event) **then**
`start[i] := time(event)`
`out.write(start)`
else if event in $commentators, interview, crowd, umpire_signal$
then
`end[i] := time(event)`
`out.write(end)`
end if
end for=0

Algorithm 3: Shot Boundary Detection

4.2.4 Extraction based on Scorecard recognition

The difference of the score from the first frame of current shot and the first frame of the next shot is checked. If it is 4, 6 then it is added to highlights or if the wicket is increased then it is added to highlights. The score of the start frame of the delivery and the next delivery is recognized. Using the scores, key events are checked. If the score differ by more than 4, it is taken as highlights. If wicket increases, it is also considered as the highlights. Corresponding delivery labels are given as the output to the next module.

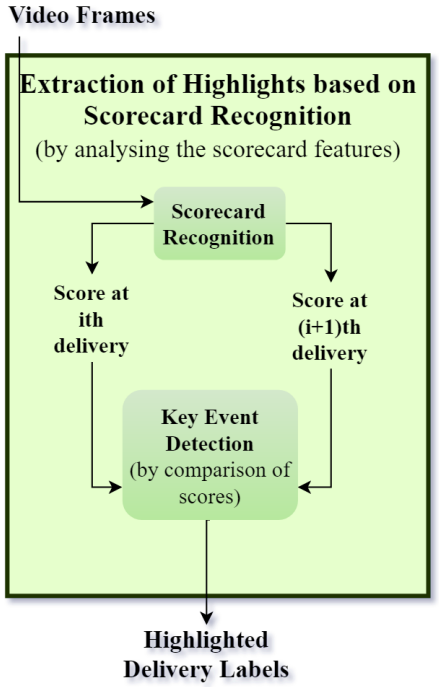


Figure 4.4 Scorecard recognition

```
Input:video frames of each segmented deliveries
Output:delivery labels that contains the
highlights
function ISKEYEVENT(i,highlights)
```

0: **function** ISKEYEVENT(*i*,*highlights*)
if get_score(get_frames(i)) - get_score(get_frames(i+1)) >=4 **then**
highlights[i]=1
else if get_wicket(get_frames(i)) - get_wicket(get_frames(i+1)) >=1
then
highlights[i]=1
else
highlights[i]=0
end if
$i \leftarrow 1$
highlights[total_deliveries]
while $i < total\_deliveries$ **do**
if $isKeyEvent(i, highlights)$ **then**
end if
end while=0

Algorithm 4: Scorecard recognition

4.2.5 Extraction based on Excitement Detection

This module is used for extracting highlights based on excitement using crowd cheer and commentators excitement from audio. The audio energy of the whole is analyzed and computed to get the threshold value then the audio energy of each shots is analyzed and computed, checked with the threshold to identify the highlights. Corresponding labels of the shots are considered as the highlights and given to the next module.

4.2.6 Aggregation of Highlights

Highlighted video labels based on scorecard and excitement is combined and corresponding video shots are fetched from the database.

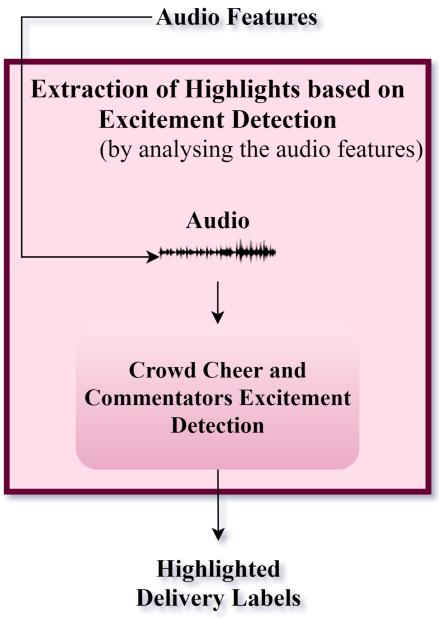


Figure 4.5 Excitement Detection

These video shots are then concatenated to generate highlights. Highlighted video labels based on scorecard and excitement are combined without duplicates. Corresponding video clips are fetched from the database using the combined labels. These clips are concatenated to form final highlights.

$$
\begin{aligned}
\text{Labels1[]} &:= \mathit{HighlightLabels\_scorecard} \\
\text{Labels2[]} &:= \mathit{HighlightLabels\_excitement} \\
\text{Result[]} &:= \text{Label1[]} \cup \mathit{Label2}[] \\
\text{ResultVideo[]} &:= \text{FetchVideoFromDatabase(Result[])} \\
\text{FinalVideo} &:= \text{ConcatenateVideo(ResultVideo)}
\end{aligned}
$$

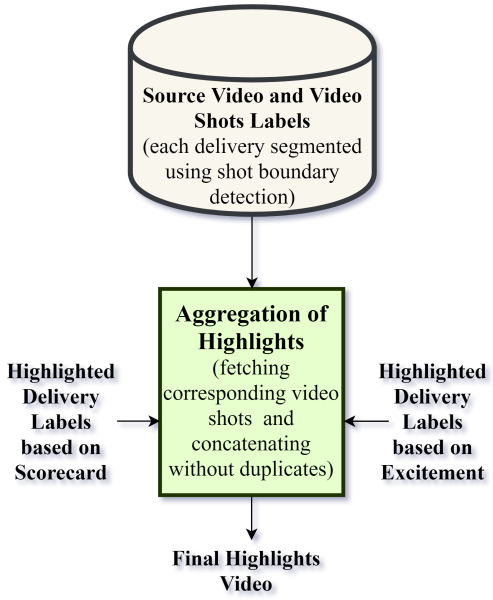


Figure 4.6 Aggregation of Highlights

CHAPTER 5

SYSTEM DEVELOPMENT

The system consists of replay/advertisements removal, CNN classifier, Shot boundary detection, Highlights generator. The overview of the entire system is

$input \leftarrow cricketmatch$
`live_frames := ReplayRemoval(`$input$`)`
`Event[] := CNN_classifier(live_frames)`
`shot[] := Shot_detector(Event)`
`Highlights :=` $Highlights_Extractor(shots[])$

5.1 PROTOTYPES OF THE MODULES

1. **Replay/Advertisement removal** Input video is checked for scorecard. Frames with scorecard is live frames and given as output.
2. **Visual Marker Detection** Live frames are given to CNN classifier for event tagging. Event tags are given as output.
3. **Shot boundary detection** With event tags, Shot boundary is detected and segmented. Shots are given as output.
4. **Extraction of highlights based on scorecard recognition** Scores in 2 consecutive shots are compared and highlights are identified.
5. **Aggregation of highlights** Both the highlights are combined to form final highlights.

6. **Extraction of highlights based on Excitement Detection** The video shots are analysed for the presence of Excitement and then classified as highlights.

5.2 DEPLOYMENT DETAILS

The deployment of the system requires opencv, tesseract, tensorflow and keras.Any python3.x can used to deploy the system successfully. Anaconda is preferred to deplay the system in case of windows.

CHAPTER 6

RESULTS AND DISCUSSION

6.1 DATASET FOR TESTING

The input to the system consists of cricket matches from BBL. The input consists of various types like match with/without highlights, with-/without replays. Each module is separately tested.

6.2 OUTPUT OBTAINED IN VARIOUS STAGES

This section shows the results obtained during the module testing.

6.2.1 Preprocessing

5_BBL_Trim_preprocessing.csv

No.	time	scorecard present	scorecard absent
1			
2	0s-18s	1	0
3			
4	19s-35s	0	1
5			
6	36s-73s	1	0
7			
8	74s-95s	0	1
9			
10	96s-131s	1	0

Figure 6.1 Output of preprocessing

6.2.2 Visual Marker Detection

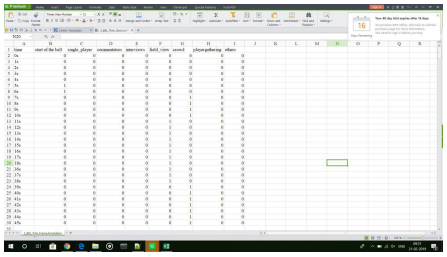
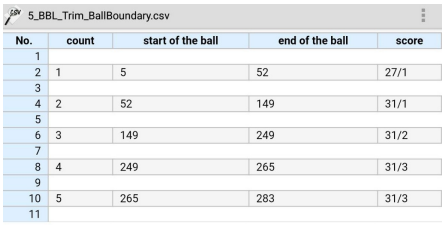


Figure 6.2 Output of Visual Marker Detection

6.2.3 Shot Boundary Detection and Segmentation



5_BBL_Trim_BallBoundary.csv

No.	count	start of the ball	end of the ball	score
1				
2	1	5	52	27/1
3				
4	2	52	149	31/1
5				
6	3	149	249	31/2
7				
8	4	249	265	31/3
9				
10	5	265	283	31/3
11				

Figure 6.3 Output of Shot Boundary Detection and Segmentation

6.2.4 Extraction of Highlights based on scorecard recognition

5_BBL_Trim_ScorecardHighlights.csv

No.	count	start of the ball	end of the ball	type
1				
2	1	5	19	Boundary
3				
4	2	52	74	Wicket
5				
6	3	149	167	Wicket

Figure 6.4 Output of Extraction of Highlights based on scorecard recognition

6.2.5 Extraction of Highlights based on excitement detection

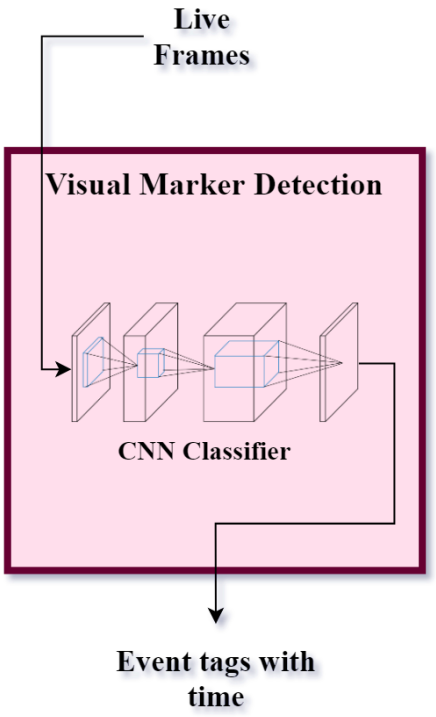


Figure 6.5 Output of Extraction of Highlights based on excitement detection

6.2.6 Aggregation of highlights

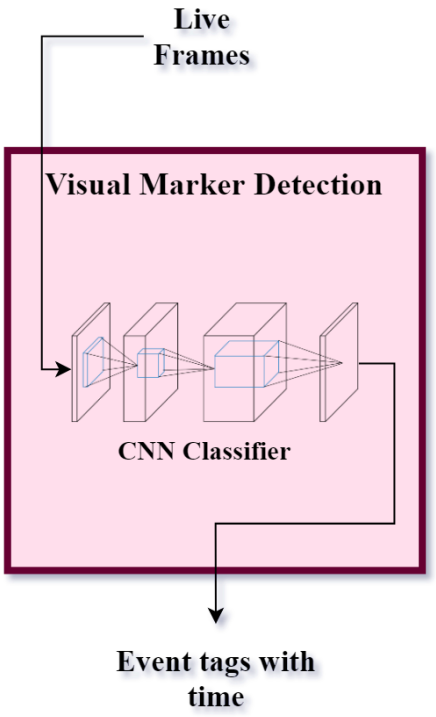


Figure 6.6 Output of Aggregation of highlights

6.3 RESULTS

The system is developed using python 3x. and keras. Big Bash League match videos are used to test the system. Day and Night matches are selected to test whether the system can support matches with variations in the contrast. System is trained using 2000 images for event detection in each frame.

6.4 EVALUATION

The performance evaluation can be divided into two major categories, namely:

1. Objective evaluation
2. Subjective evaluation

Objective evaluation criterion relies on metrics such as precision, recall, accuracy, false alarm rates, etc. to measure the performance of the system.Following are the some evaluation metrics:

- **True Positive (TP):** It refers to the positive samples that are correctly labeled by the classifier.
- **True Negative (TN):** It refers to the negative samples that are correctly labeled by the classifier.
- **False Positive (FP):** It refers to the negative samples that are incorrectly labeled as positive by the classifier.
- **False Negative (FN):** It refers to the positive samples that are incorrectly labeled as negative by the classifier.
- **Precision Rate (PR):** It is a ratio of number of correctly labeled events (or frames), TP, to the total number of events detected.

$$Precision = \frac{TP}{TP + FP}$$

- **Recall Rate (RR):** It is a ratio of true detection rate with respect

to the actual events (frames) in the video.

$$Recall = \frac{TP}{TP+FN}$$

- **Error Rate (ER):** It is a ratio of the miss labeled events (both false positives and false negatives) to the total number of events examined.

$$ErrorRate = \frac{FP+FN}{TP+TN+FP+FN}$$

- **Accuracy Rate (AR):** It is a ratio of the correctly labeled events (both true positives and true negatives) to the total number of events.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

6.4.1 Preprocessing

6.1 contains result of the preprocessing module. In preprocessing module, live frames and Replay/Ads frames are detected and classified.

		Predicted		
		Live Frames	Replay/Ad Frames	Accuracy
Actual	Live Frames	148	2	99
	Replay/Ad Frames	0	60	100

Table 6.1 Preprocessing

6.4.2 Visual marker detection

6.2 shows the result of visual marker detection where each live frame's event is detected. Here 1 represents start of the ball, 2 represents single player, 3 represents commentators, 4 represents interviews, 5 represents field view, 6 represents crowd, 7 represents player gathering, 8 represents other frames.

Actual \ Predicted	1	2	3	4	5	6	7	8	Accuracy
1	45	0	0	9	0	0	0	6	83
2	0	24	3	0	3	0	0	0	91
3	0	6	18	0	3	3	0	0	88
4	9	0	0	18	0	3	0	0	88
5	0	0	6	0	21	3	0	0	87
6	0	3	0	0	6	21	0	0	90
7	0	6	18	0	3	3	0	0	88
8	9	0	0	0	0	0	0	7	0

Table 6.2 Visual marker detection

6.4.3 Shot boundary detection

6.3 shows the result of shot boundary detection where shots are created using the event tag of the corresponding frames. Here each number represents its corresponding delivery's frames.

Actual \ Predicted	1	2	3	4	5	6	7	8	others	Ads	Accuracy
1	30	0	0	6	0	0	0	4	0	0	83
2	0	16	2	0	2	0	0	0	0	0	91
3	0	4	12	0	2	2	0	0	0	0	88
4	6	0	0	12	0	2	0	0	0	0	88
5	0	0	4	0	14	2	0	0	0	0	87
6	0	2	0	0	4	14	0	0	0	0	90
7	0	4	12	0	2	2	0	0	0	0	88
8	0	0	0	0	0	0	0	14	0	0	0
others	0	4	12	0	2	2	0	0	0	0	88
Ads	0	0	0	0	0	0	0	14	0	0	0

Table 6.3 Shot boundary detection

6.4.4 Scorecard recognition

6.4 shows the result of scorecard recognition where highlights shots are created using score present in the scorecard. Here each number represents its corresponding delivery's frames of the highlights.

		Predicted							
		1	2	3	4	5	6	Others	Accuracy
Actual	1	15	0	0	0	0	0	3	0
	2	0	10	0	0	0	0	2	0
	3	0	0	20	0	0	0	4	0
	4	0	0	0	21	0	0	1	0
	5	0	0	0	0	17	0	0	0
	6	0	0	0	0	0	11	3	0
	Others	1	1	0	3	2	1	205	0

Table 6.4 Scorecard recognition

CHAPTER 7

CONCLUSIONS

7.1 CONTRIBUTIONS

The proposed model generates highlights based on the scorecard and excitement features. The replays and advertisements are considered to be undesirable hence they are detected and removed. Key events like 4s, 6s, wicket are identified and categorized as highlights. Highlights based on the crowd cheering and excitement of the commentators are also generated. In order to have well defined boundaries, Events like bowling, commentators, interviews, field view are detected and shot boundaries are detected and segmented.

7.2 FUTURE WORK

The tasks which needs further exploration are as follows:

- The proposed system is designed for 20 over match. The system can also be generalized for 50 over match. This can be challenging because of its longer duration.
- Separate highlight video for each key event can be done.
- Different broadcasters will have scorecard in different formats.Instead of specifying the scorecard's position, detection of the scorecard can be automated.
- Parallel processing can be done to improve the execution time.

APPENDIX A

Cricket

Cricket is a bat-and-ball game played between two teams of eleven players on a field at the centre of which is a 20-metre (22-yard) pitch with a wicket at each end, each comprising two bails balanced on three stumps. Cricket is played with two teams (say A and B) normally of 11 players a side, one being the batting team while the other one is the fielding team. It is generally played on field with the main playing surface being called a pitch. (For details and dimensions of the pitch, wickets and creases, click the tab 'Playing Surface' on the right).

Team A will bat first and try to score as many runs as possible while the second team, team B, will bowl and field to make it as hard as possible for the batting team (A) to score these runs and to get them out . Once team A are all out or otherwise their batting is determined closed as per the laws, the teams then swap over. So team B will bat to try and beat the score (number of runs scored) set by team A. Team A will bowl and field and try and restrict Team B from beating their score / getting them all out before they do.

Cricket is a game for all - adults, young people, children, men and women, girls and boys. They play cricket all over the world - on the street, on the beach, in the local park, wherever they can find a place to play. Above all they have fun doing so!

A.1 MAIN ASPECTS OF PLAYING THE GAME

Thus, broadly summarized, there are 6 key elements of cricket: Batting Bowling Fielding Catching Wicket Keeping Scoring runs Thanks to our cricketing friends in America, we can bring a good introductory compilation of the key aspects of the game through some video footage.

A.2 FORMATS OF CRICKET

There are various formats ranging from Twenty20, played over a few hours with each team batting for a single innings of 20 overs, to Test matches, played over five days with unlimited overs and the teams each batting for two innings of unlimited length. Traditionally cricketers play in all-white kit, but in limited overs cricket they wear club or team colours. In addition to the basic kit, some players wear protective gear to prevent injury caused by the ball, which is a hard, solid spheroid made of compressed leather with a slightly raised sewn seam enclosing a cork core which is layered with tightly wound string.

APPENDIX B

Video Summarization

There have been tremendous needs of video processing applications to deal with abundantly available accessible videos. One of the research areas of interest is Video Summarization that aims creating summary of video to enable a quick browsing of a collection of large video database. It is also useful for allied video processing applications like video indexing, retrieval etc. Video Summarization is a process of creating presenting a meaningful abstract view of entire video within a short period of time. Mainly two types of video summarization techniques are available in the literature, viz. key frame based and video skimming. For key frame based video summarization, selection of key frames plays important role for effective, meaningful and efficient summarizing process. novel variant of video summarization, namely building a summary that depends on the particular aspect of a video the viewer focuses on. We refer to this as viewpoint. To infer what the desired viewpoint may be, we assume that several other videos are available, especially groups of videos, e.g., as folders on a persons phone or laptop. The semantic similarity between videos in a group vs. the dissimilarity between groups is used to produce viewpoint-specific summaries. For considering similarity as well as avoiding redundancy, output summary should be (A) diverse, (B) representative of videos in the same group, and (C) discriminative against videos in the different groups. To satisfy these requirements (A)-(C) simultaneously, we proposed a novel video summarization method from multiple groups of videos. Inspired by Fishers

discriminant criteria, it selects summary by optimizing the combination of three terms (a) inner-summary, (b) innergroup, and (c) between-group variances defined on the feature representation of summary, which can simply represent (A)-(C). Moreover, we developed a novel dataset to investigate how well the generated summary reflects the underlying viewpoint. Quantitative and qualitative experiments conducted on the da As the name implies, video summarization is a mechanism for generating a short summary of a video, which can either be a sequence of stationary images (key frames) or moving images (video skims). Video can be summarized by two different ways which are as follows.

B.1 KEY FRAME BASED VIDEO SUMMARIZATION

These are also called representative frames, R-frames, still-image abstracts or static storyboard, and a set consists of a collection of salient images extracted from the underlying video source [2]. Following are some of the challenges that should be taken care while implementing Key frame based algorithm 1. Redundancy: frames with minor difference are selected as key frame. 2. When there are various changes in content it is difficult to make clustering.

B.2 VIDEO SKIM BASED VIDEO SUMMARIZATION

This is also called a moving-image abstract, moving story board, or summary sequence [2]. The original video is segmented into various parts which is a video clip with shorter duration. Each segment is joined by either a cut or a gradual effect. The trailer of movie is the best example for video skimming.

APPENDIX C

Big Bash League

The Big Bash League (BBL) is an Australian professional Twenty20 cricket league, which was established in 2011 by Cricket Australia. The Big Bash League replaced the previous competition, the KFC Twenty20 Big Bash, and features eight city-based franchises instead of the six state teams which had participated previously. The competition has been sponsored by fast food chicken outlet KFC since its inception. It is one of the two T20 cricket, alongside the Indian Premier League, to feature amongst the Top 10 Most Attended Sport Leagues in the world. BBL matches are played in Australia during the southern hemisphere summer, in the months of December, January and February.

C.1 TOURNAMENT FORMAT

Ben Cutting of Brisbane Heat batting against Melbourne Stars in 2014 Since the inception of the BBL in 2011, the tournament has followed the same format every year except the inaugural season.[23] The first BBL season had 28 group stage matches, before expanding to 32 in the following season. Since the 201819 season, each team plays all other teams twice during a season, for a total of 56 regular season matches before the finals series..

In previous seasons of the tournament, the group stage matches were divided into eight rounds, with four matches played in each round. Each team played six other teams once during a season, and one team twice. This allowed for both Sydney and Melbourne (which have two

teams each) to play 2 derbies within a single season.Each team played eight group stage matches, four at home and four away, before the top four ranked teams progressed to the semi finals.In the 2017/18 Season) the format changed so that there would be 40 group stage matches with each team playing 10 matches before the semi finals.The season was held over a similar time-frame thus resulting in more doubleheaders (one game afternoon, one game night) and teams playing more regularly.

The final of the tournament is played at the home ground of the highest-ranked team. The only exception to this rule was 201415 season when the final was played at a neutral venue (Manuka Oval), due to the 2015 Cricket World Cup.

In the 201819 season, the league introduced a 'bat flip' (instead of a coin toss) to decide who would bat/bowl first.

C.2 CURRENT TEAMS

The competition features eight city-based franchises, instead of the six state-based teams which had previously competed in the KFC Twenty20 Big Bash. Each state's capital city features one team, with Sydney and Melbourne featuring two. The team names and colours for all teams were officially announced on 6 April 2011.The Melbourne Derby and Sydney Derby matches are some of the most heavily attended matches during the league and are widely anticipated by the fans.The Scorchers and Sixers have also developed a rivalry between them over the years and their matches attract good crowds and TV ratings

REFERENCES

[1] M. E. Anjum, S. F. Ali, M. T. Hassan, and M. Adnan, "Video summarization: Sports highlights generation", In *INMIC*, pp. 142–147, Dec 2013.

[2] Chih-Chieh Cheng and Chiou-Ting Hsu, "Fusion of audio and motion information on hmm-based highlight extraction for baseball games", *IEEE Transactions on Multimedia*, vol. 8, num. 3, pp. 585–599, June 2006.

[3] Taufiq Hasan, Hynek Bořil, Abhijeet Sangwan, and John H. L Hansen, "Multi-modal highlight generation for sports videos using an information-theoretic excitability measure", *EURASIP Journal on Advances in Signal Processing*, vol. 2013, num. 1, p. 173, Nov 2013.

[4] A. Javed, K. B. Bajwa, H. Malik, and A. Irtaza, "An efficient framework for automatic highlights generation from sports videos", *IEEE Signal Processing Letters*, vol. 23, num. 7, pp. 954–958, July 2016.

[5] M. H. Kolekar, K. Palaniappan, and S. Sengupta, "Semantic event detection and classification in cricket video sequence", In *2008 Sixth Indian Conference on Computer Vision, Graphics Image Processing*, pp. 382–389, Dec 2008.

[6] M. H. Kolekar and S. Sengupta, "Bayesian network-based customized highlight generation for broadcast soccer videos", *IEEE Transactions on Broadcasting*, vol. 61, num. 2, pp. 195–209, June 2015.

[7] Somnath Senguptak Maheshkumar H. Kolekar, “Caption content analysis based automated cricket highlight generation”, 2007.

[8] M. Merler, K. C. Mac, D. Joshi, Q. Nguyen, S. Hammer, J. Kent, J. Xiong, M. N. Do, J. R. Smith, and R. Feris, “Automatic curation of sports highlights using multimodal excitement features”, *IEEE Transactions on Multimedia*, vol. , pp. 1–1, 2018.

[9] P. Shukla, H. Sadana, A. Bansal, D. Verma, C. Elmadjian, B. Raman, and M. Turk, “Automatic cricket highlight generation using event-driven and excitement-based features”, In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1881–18818, June 2018.

[10] Dian Tjondronegoro, Yi-Ping Phoebe Chen, and Binh Pham, “Sports video summarization using highlights and play-breaks”, In *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval*, MIR ’03, pp. 201–208, New York, NY, USA, 2003, ACM.