Fine Grain Annotation of Cricket Videos
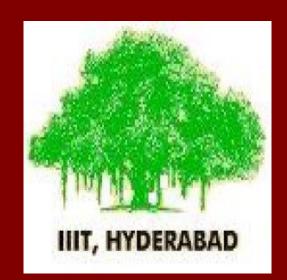
Rahul Anand Sharma, Pramod Sankar K¥ and C. V. Jawahar

CVIT, IIIT Hyderabad, India
¥Xerox Research Center India
http://cvit.iiit.ac.in/projects/cricketannotation

Problem: Fine Grain Annotation of Cricket Videos

Can we Detect and Generate detailed commentary from Cricket Videos?

Batsman: Gambhir
Description : "He pulls it off from outside off stump and just manages to clear the deep square leg rope"

Challenges

Semantic description of actions/activities is difficult using current visual-recognition techniques alone.
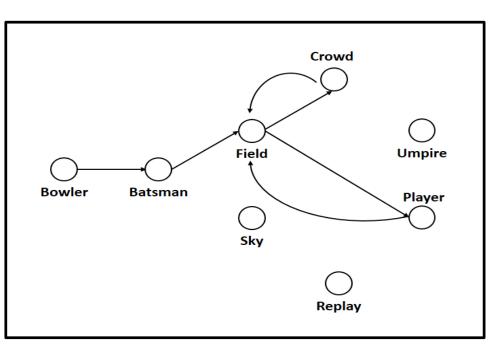Typical actions of interest are *tiny* in space-time (90x150 pixels across ~35 frames(1.2 seconds)
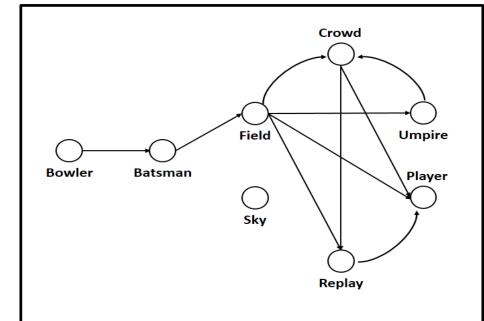
Goals

Annotate Cricket Videos with Semantic descriptions at a fine grain spatio-temporal scale.
Leverage *parallel* information in the form of online text commentaries, to augment visual recognition
Label actions at shot-level with semantic descriptions of actions and activities.
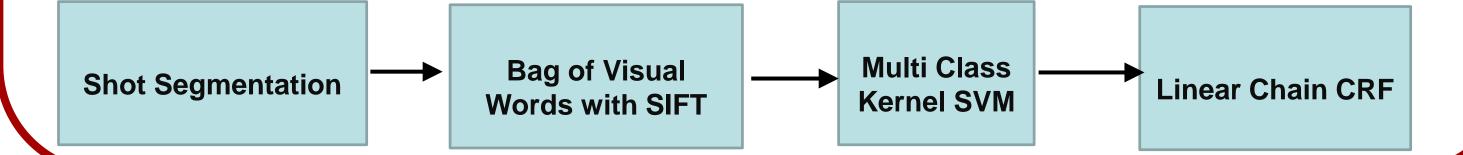
Scene Segmentation

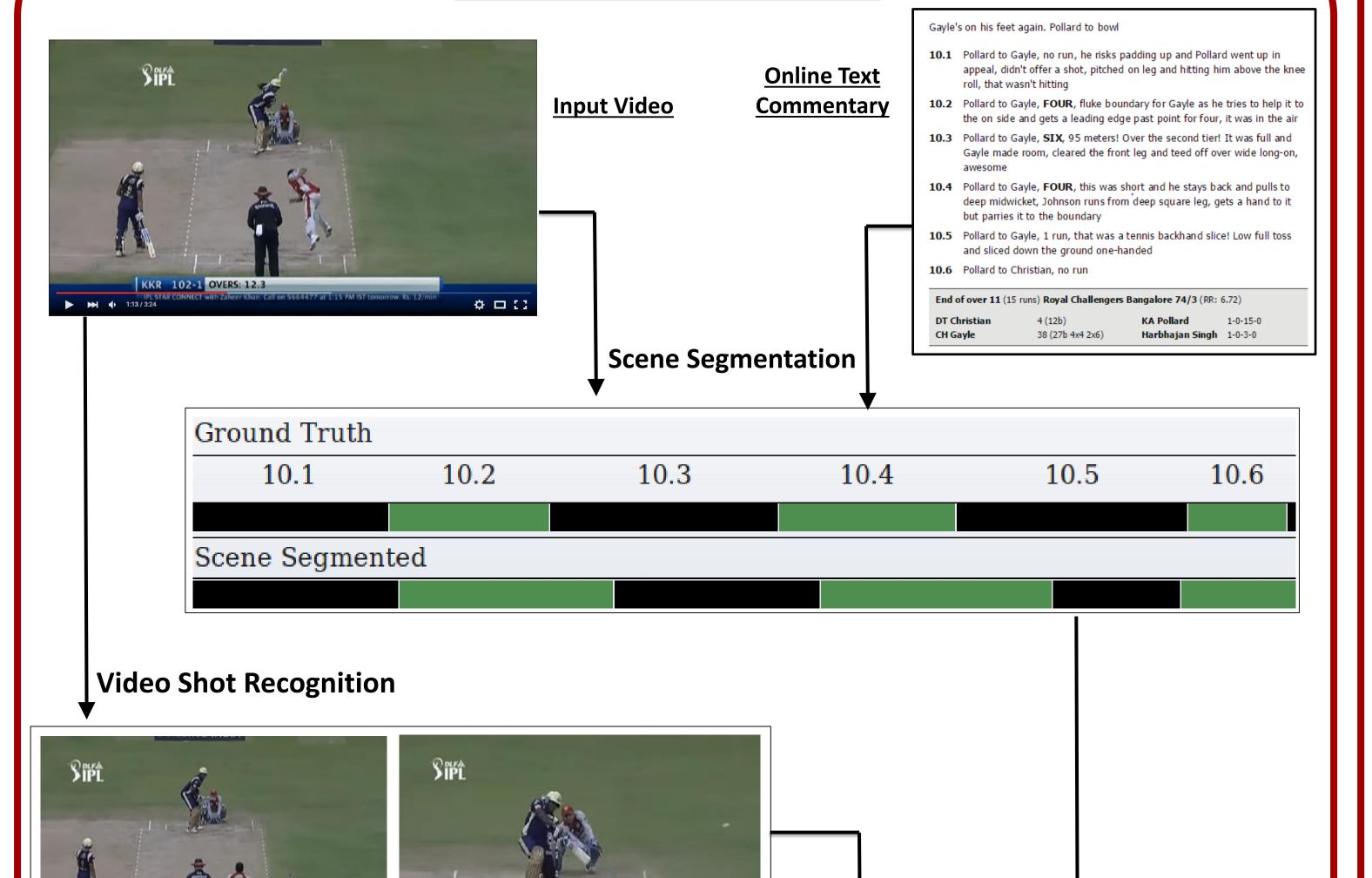**one-run model** includes only a few states
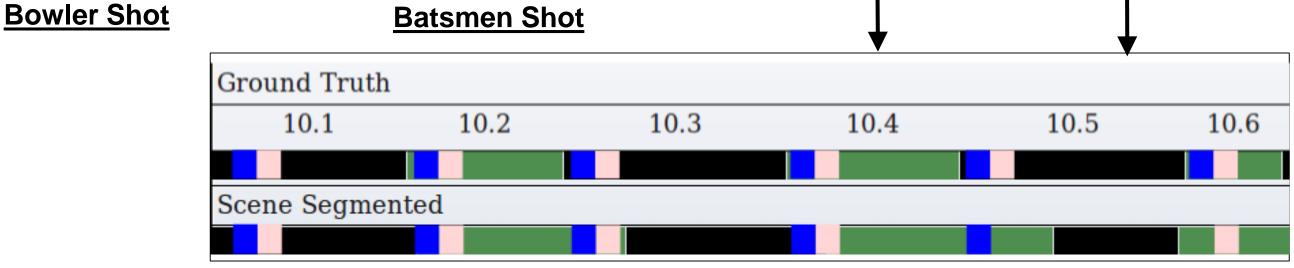
four-run model involves a variety of visuals

State transition diagrams for two scene categories: (left) One Run and (right) Four-Runs

Video Shot Recognition

Shot Segmentation → Bag of Visual Words with SIFT → Multi Class Kernel SVM → Linear Chain CRF

Solution Overview

Input Video   **Online Text Commentary**

Gayle's on his feet again. Pollard to bowl

10.1 Pollard to Gayle, no run, he risks padding up and Pollard went up in appeal, didn't offer a shot, pitched on leg and hitting him above the knee roll, that wasn't hitting
10.2 Pollard to Gayle, FOUR, fluke boundary for Gayle as he tries to help it to the on side and gets a leading edge past point for four, it was in the air
10.3 Pollard to Gayle, SIX, 95 meters! Over the second tier! It was full and Gayle made room, cleared the front leg and teed off over wide long-on, awesome
10.4 Pollard to Gayle, FOUR, this was short and he stays back and pulls to deep midwicket, Johnson runs from deep square leg, gets a hand to it but parries it to the boundary
10.5 Pollard to Gayle, 1 run, that was a tennis backhand slice! Low full toss and sliced down the ground one-handed
10.6 Pollard to Christian, no run

End of over 11 (15 runs) Royal Challengers Bangalore 74/3 (RR: 6.72)
DT Christian   4 (12b)   KA Pollard   1-0-15-0
CH Gayle   38 (27b 4x4 2x6)   Harbhajan Singh   1-0-3-0

Scene Segmentation

Ground Truth
| 10.1 | 10.2 | 10.3 | 10.4 | 10.5 | 10.6 |

Scene Segmented

Video Shot Recognition

Bowler Shot   **Batsmen Shot**

Ground Truth
| 10.1 | 10.2 | 10.3 | 10.4 | 10.5 | 10.6 |

Scene Segmented

Blue and Pink denotes Bowler shot and Batsmen shot respectively

Text Classification

10.4, Pollard to Gayle, FOUR, this was short and he stays back and pulls to deep midwicket, Johnson runs from deep square leg, gets a hand to it but parries it to the boundary

event number and players involved along with the outcome (Four run)
(red) bowler actions, *(blue)* batsman action and *(green)* other player actions

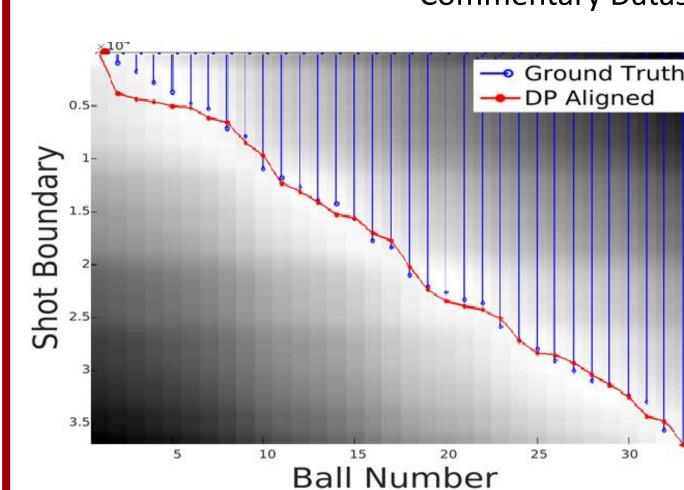**Center for Visual Information Technology, IIIT Hyderabad**

Experimental Results

Datasets

Name	Matches	No. of Phrases	Role
IPL Video Dataset	4 Matches (20Hrs)	960 Phrases	Video Shot Recognition
CricInfo Dataset	300 Matches	1500 Bowler Phrases 6000 Batsmen Phrases	Text Classification

Dataset is collected from the YouTube channel of the Indian Premier League (IPL) Tournament. Commentary Dataset is collected from CricInfo.com

Results of Scene Segmentation depicted over five-overs. The background of the image is the cost function of the scene segmentation.

Red Line - Optimal backtrack path
Red Circles - Inferred scene boundaries
Blue Lines - GT

Phrase Prediction

R	Bowler Shot	Batsman Shot
2	22.15	39.4
4	43.37	47.6
6	69.09	69.6
8	79.94	80.8
10	87.87	88.95

R is Neighborhood to search for correct Shot

Video Shot recognition

Kernel + CRF	Vocab: 300	Vocab: 1000
Linear	78.02	**82.25**
Polynomial	80.15	81.16
RBF	**81**	82.15
Sigmoid	77.88	80.53

Text Classifier

Method	SVM
Bag of Words (SVM)	89.09

Calculated using 2 Fold cross Validation

shortish and swinging away lots of bounce

on the pads once more Dravid looks to nudge it to square leg, off the pads for a leg bye

this is again slammed towards long-on

short outside off

Summary and Conclusion

- Our solution enables rich semantic annotation of Cricket Videos at a fine temporal scale.
- Circumvention of technical challenges in visual recognition by utilizing online text-commentaries.
- We obtain high annotation accuracy as evaluated over a large video collection.
- In future, our labelled dataset could be used to learn classifiers for fine-grain activity recognition.
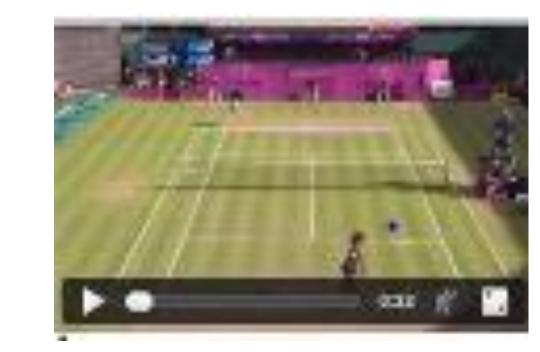
Other Applications

Highlight Generation
Automatic Highlight generation from a Match

Sports
Event Detection in Tennis Videos

Activity learning
Learning classifier for Cover drive and so on.