Generic Architecture for Event Detection in Broadcast Sports Video

Chris Poppe, Sarah De Bruyne,
Rik Van de Walle
Multimedia Lab – Ghent University - IBBT

Context

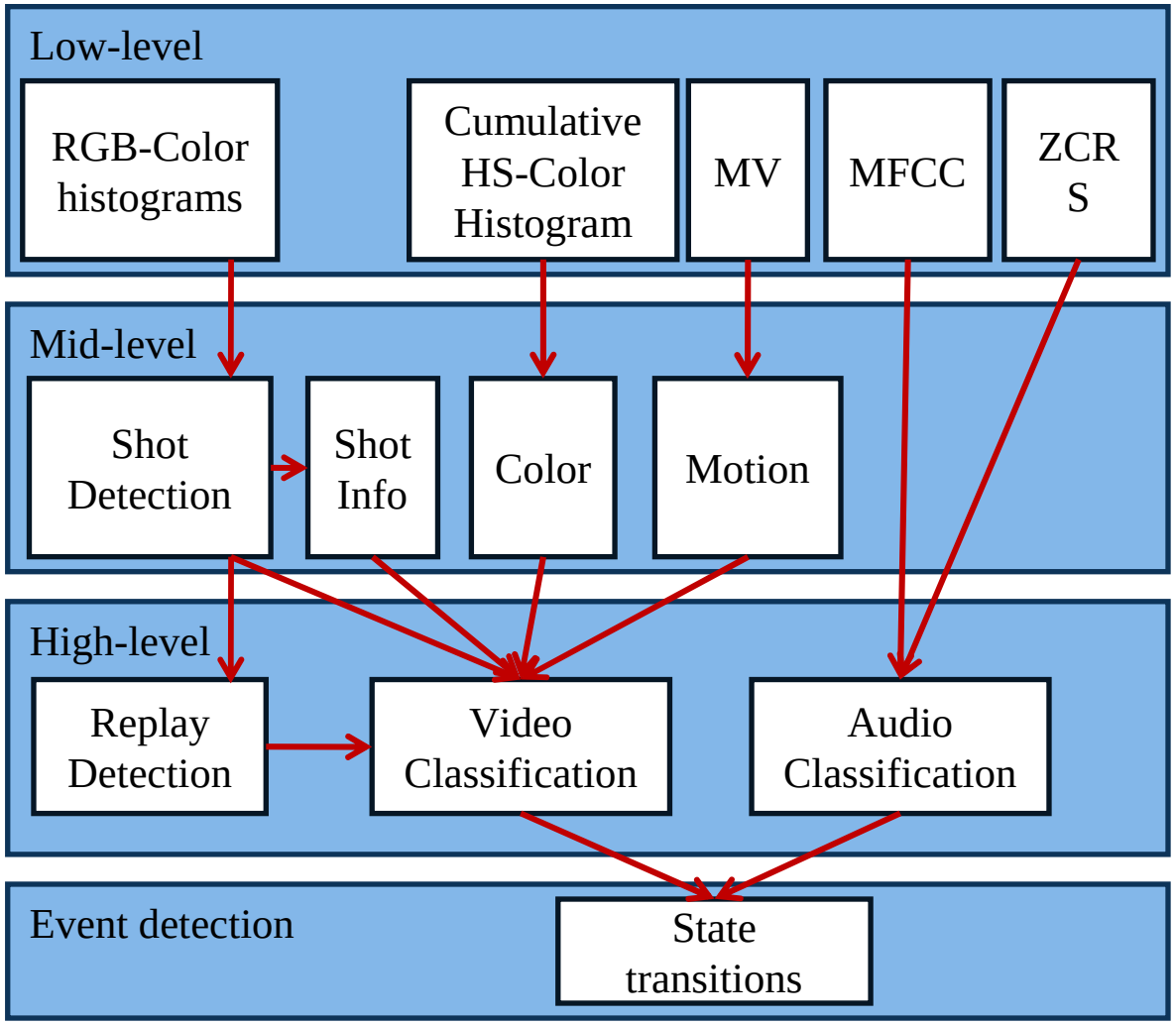
- Broadcast sports video

- Automatic annotation
  - summarization
  - retrieval
  - adaptation (e.g., for mobile devices)
  - statistical analysis (coaching)

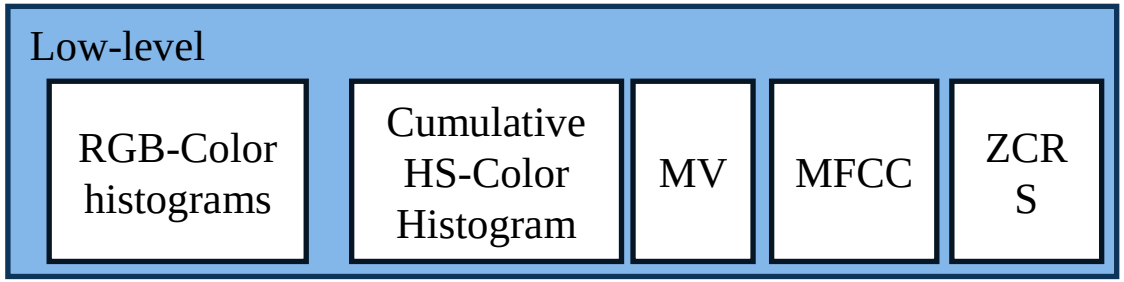
Context

- Generic!
  - soccer
  - tennis
  - football
  - cricket
  - hockey
  - …
- common aspects: broadcast video, edited, game rules
- Generic = easily configurable ☺

System


Low-level
RGB-Color histograms
Cumulative HS-Color Histogram
MV
MFCC
ZCR S
Mid-level
Shot Detection
Shot Info
Color
Motion
High-level
Replay Detection
Video Classification
Audio Classification
Event detection
State transitions

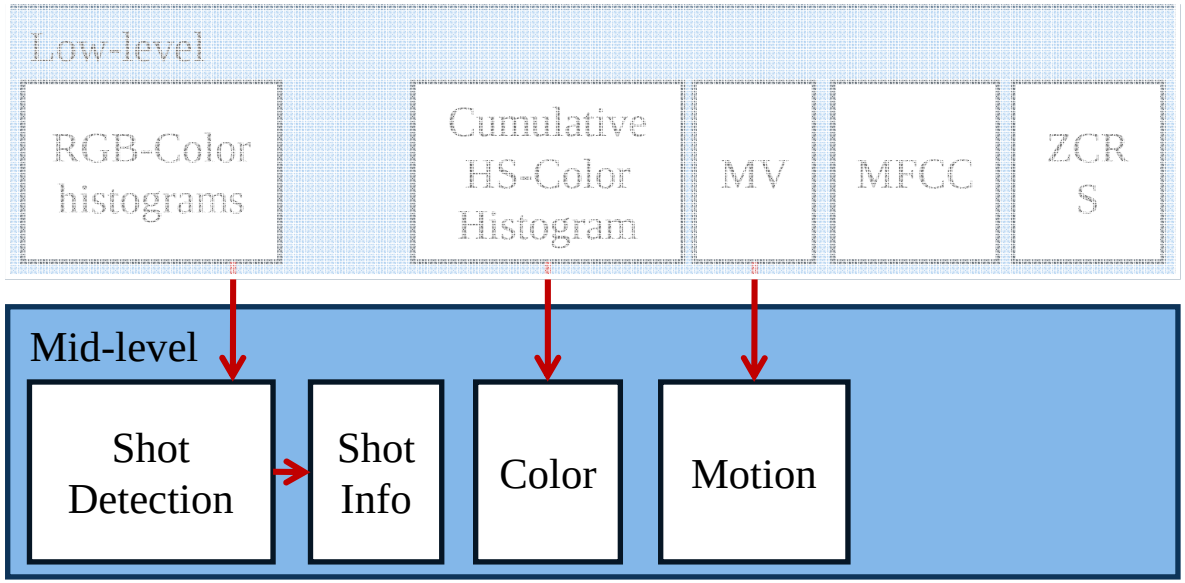
Low-level



- Video
  - Motion Vectors
  - RGB histogram
  - Cummulative HS(V)-histogram
    - determines dominant color in video
- Audio
  - MARSYAS (Music Analysis, Retrieval and Synthesis for *Audio* Signals) an open source software framework for *audio* processing
    - Mel-Frequency Cepstral Coefficients (MFCC)
    - Zero Crossings (ZCRS)

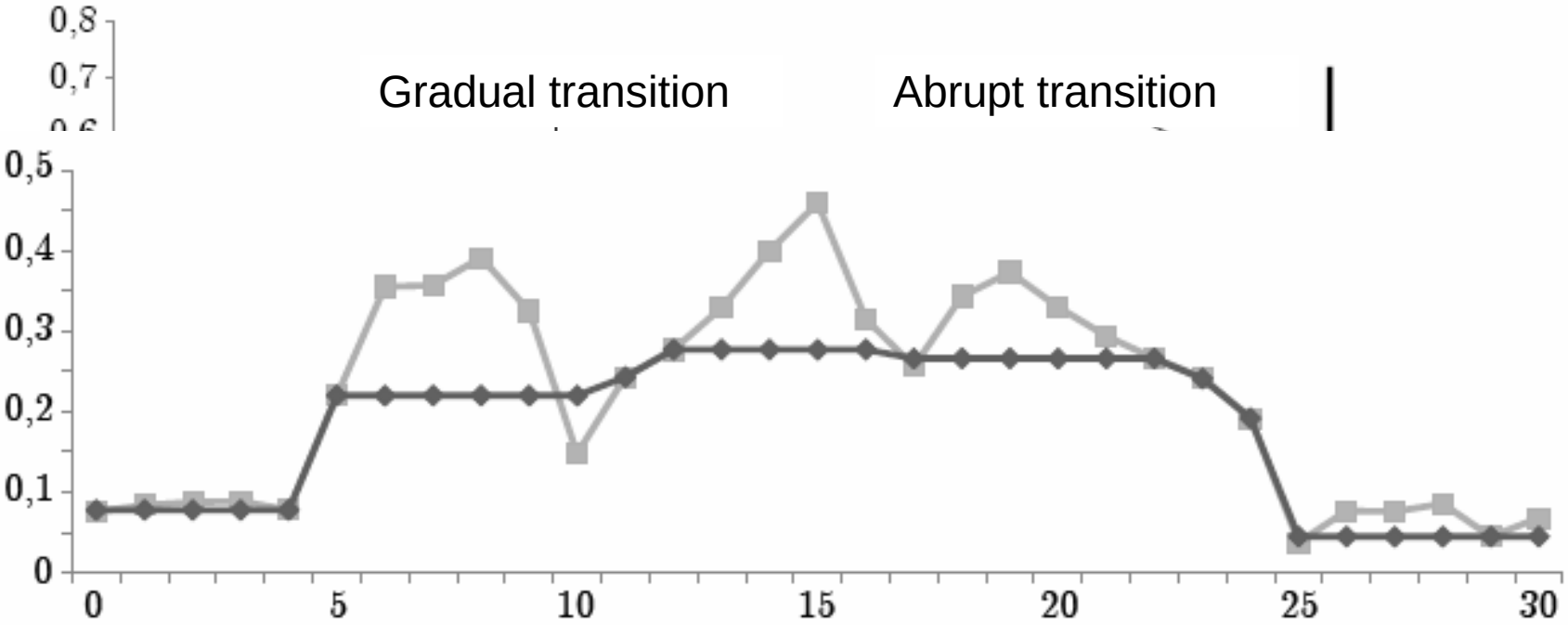
Mid-level

- Rely on low-level or mid-level modules
- Create features
  - information compression



Shot Detection

- Evaluate differences of histograms
- Local search window to find maxima
- Morphological processing for gradual transitions



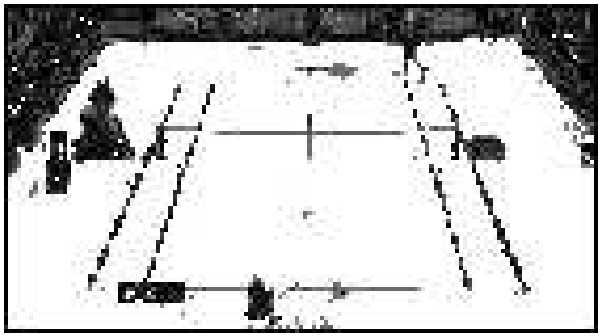
Shot detection results

Precision and recall values for the detection of shot transitions

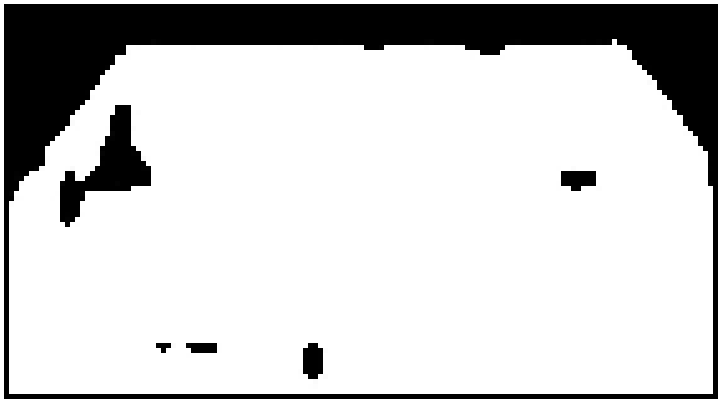
	Precision (%)	Recall (%)
Roland Garros	99.44	98.90
Manchester United	93.86	100

Color Module

- Cummulative histogram for detection of playing field
- Back projection

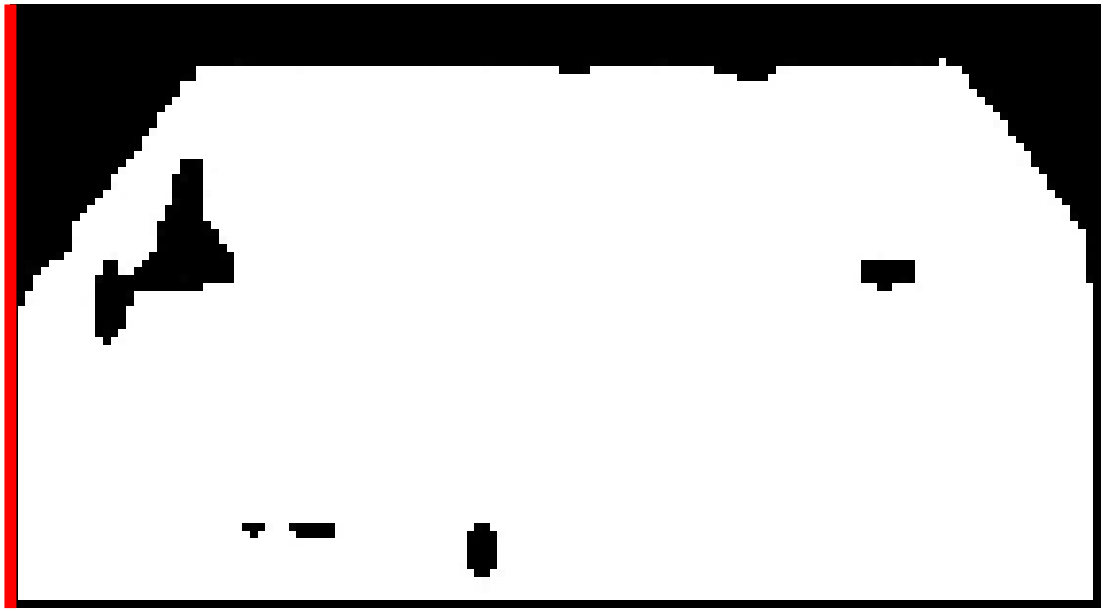


- Gaussian smoothing



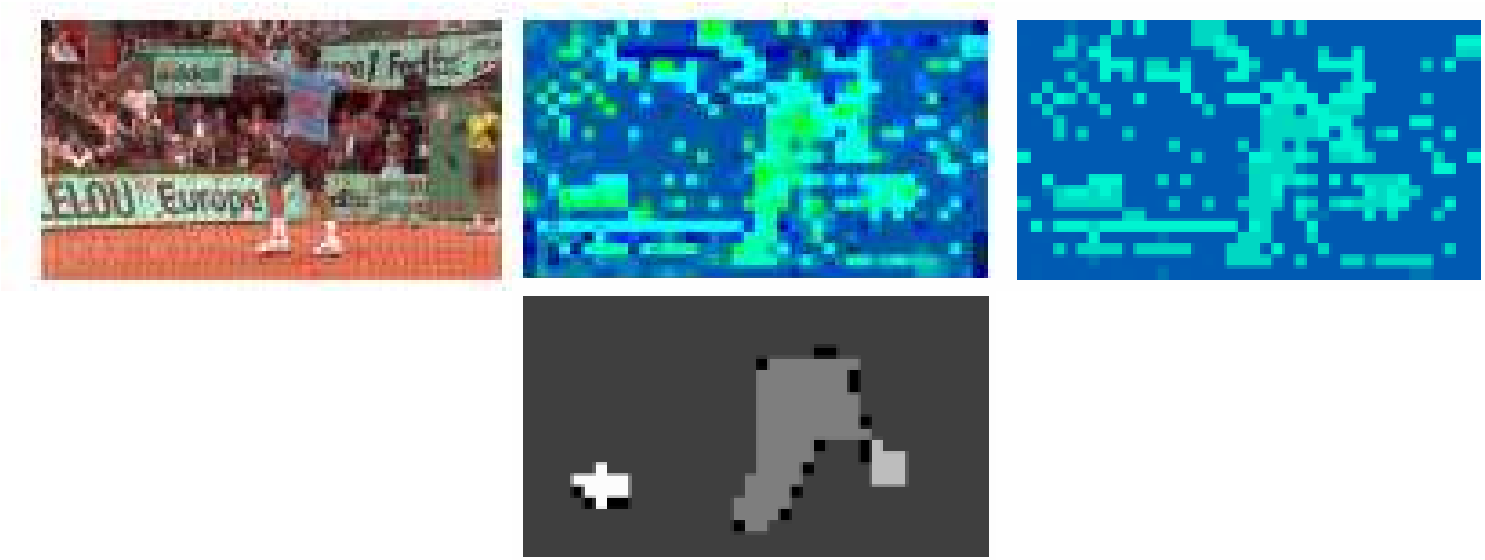
Color Module

- Calculate ratio of field pixels for each vertical line
  - *ColorMax*
  - *ColorStDev*



Motion Module

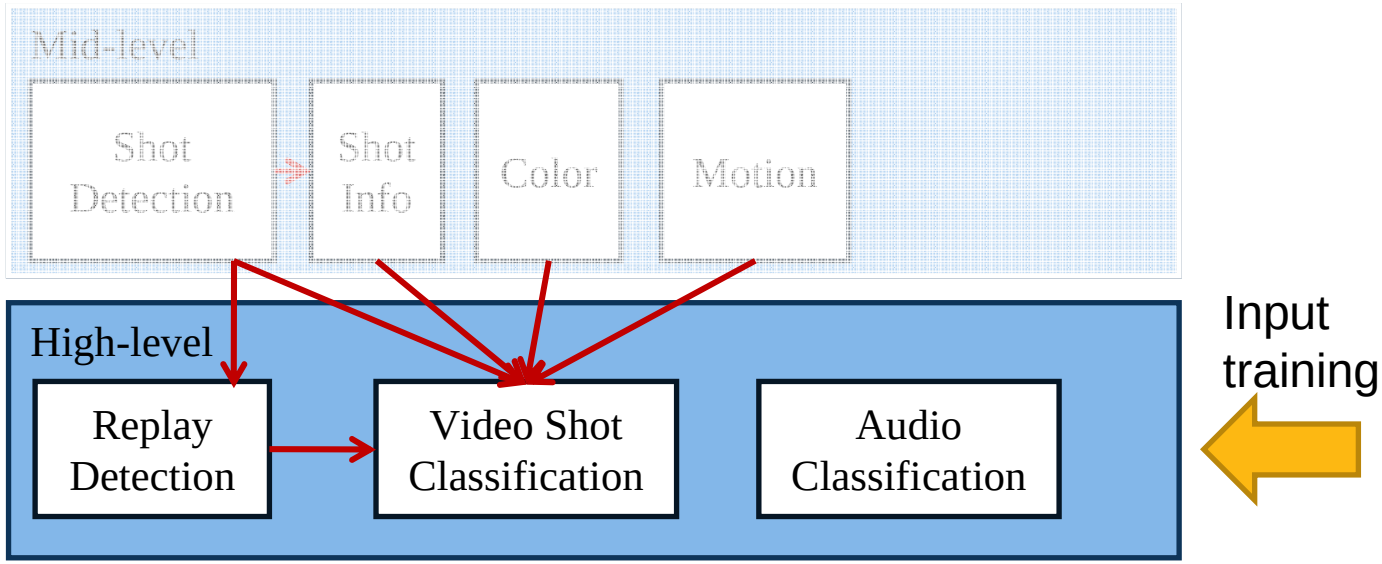
- MV Space
  - represent MVs as HSV values according to size and angle
  - Cluster (image processing)



  - *MotionEntropy, HorizontalMotion, VerticalMotion, DiagonalMotion*

High-level

- Add semantic meaning based on Mid-level features
  - famous semantic gap
- Replay detection
  - assume gradual transitions correspond to replay
- Classification



Video Classification

- Labeled training set

Video labels for soccer sequences	
Audience	Reaction of audience
Field	Overview of actions
Goal	Field view with goal
Close-up of player	No action
Tracking of player	Medium view of action
Medium image of player	Before or after an action

SVM Classification

1. Train SVM
2. Input feature vector
3. Output label

???

ColorMax, MotionEntropy, …

"Field" "Tracking of Player"

ColorMax, MotionEntropy, …

SVM

"Field"

Classification results

Classification results for soccer (in %).

SVM	Precision (%)	Recall (%)
Audience	100	30.43
Field	75.32	99.58
Goal	89.47	32.08
Close-up of player	90.03	92.18
Tracking of player	89.15	78.23
Medium image of player	61.76	42.00
Replay	88.99	97.00

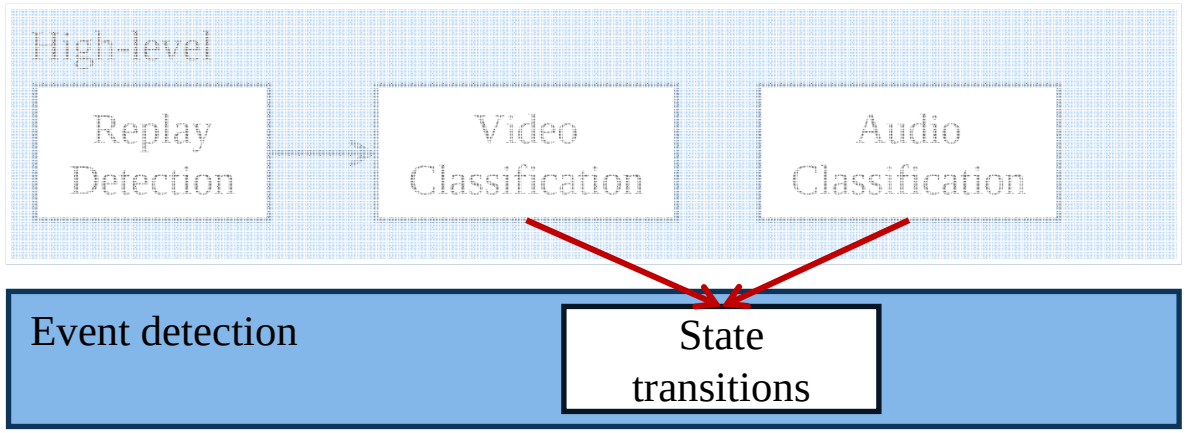
Audio Classification

- Labeled training set
- Feature vectors (*MFCC, ZCRS*)
- Audio samples

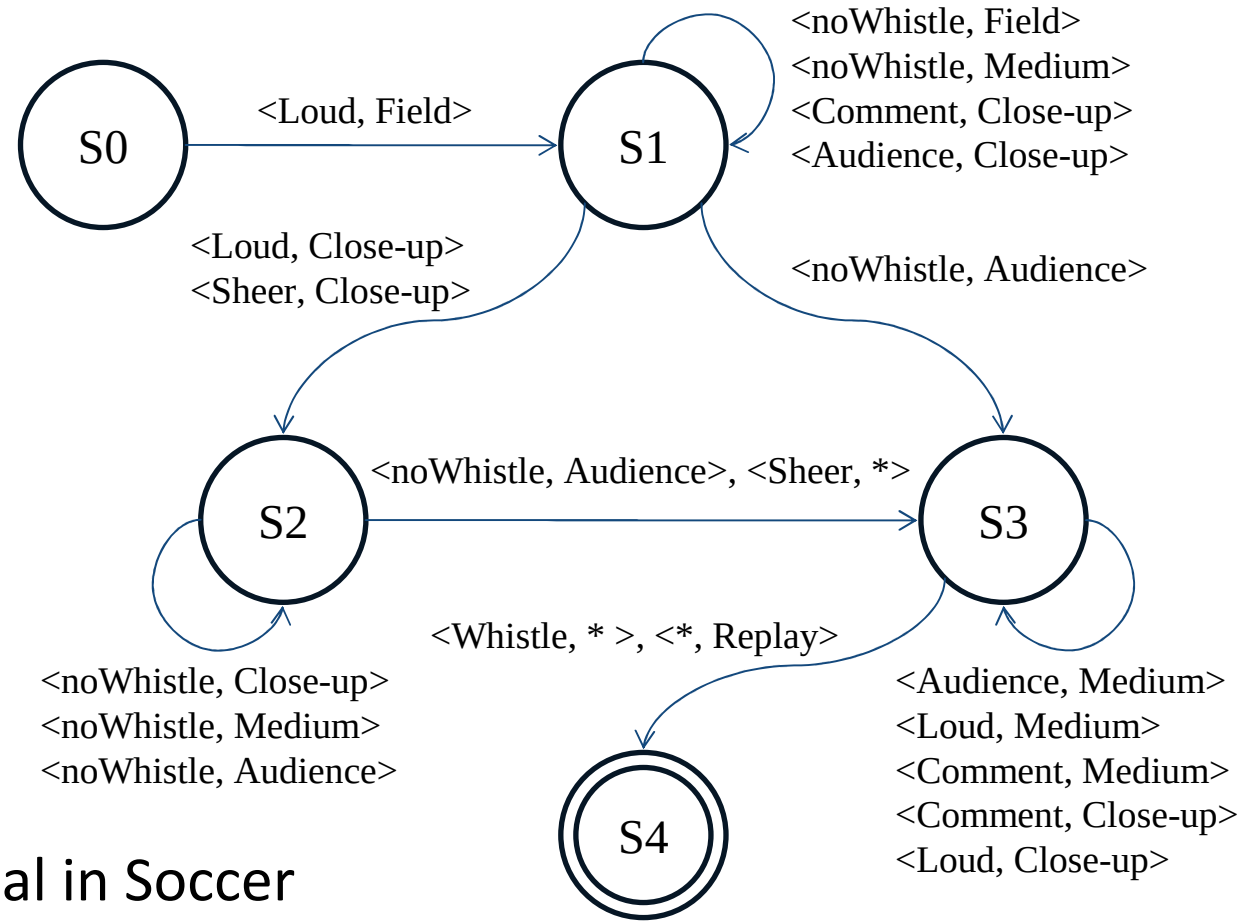
SVM	Precision (%)	Recall (%)
Whistle	95.00	97.44
Loud comments	96.00	72.73
Comments	77.78	87.50
Sheering	83.33	95.24
Audience	90.45	95.30

Event detection

- Detect events using cinematographic principles
  - Broadcast sports video is composed by director
  - Game-rules can also be used
- Search for patterns of video shots and combine with audio
  - Finite State Machine (FSM)
    - models an event (e.g., goal in soccer or ace in tennis)



Event description



- Goal in Soccer

Results

Detection results for tennis points and soccer goals (in %)

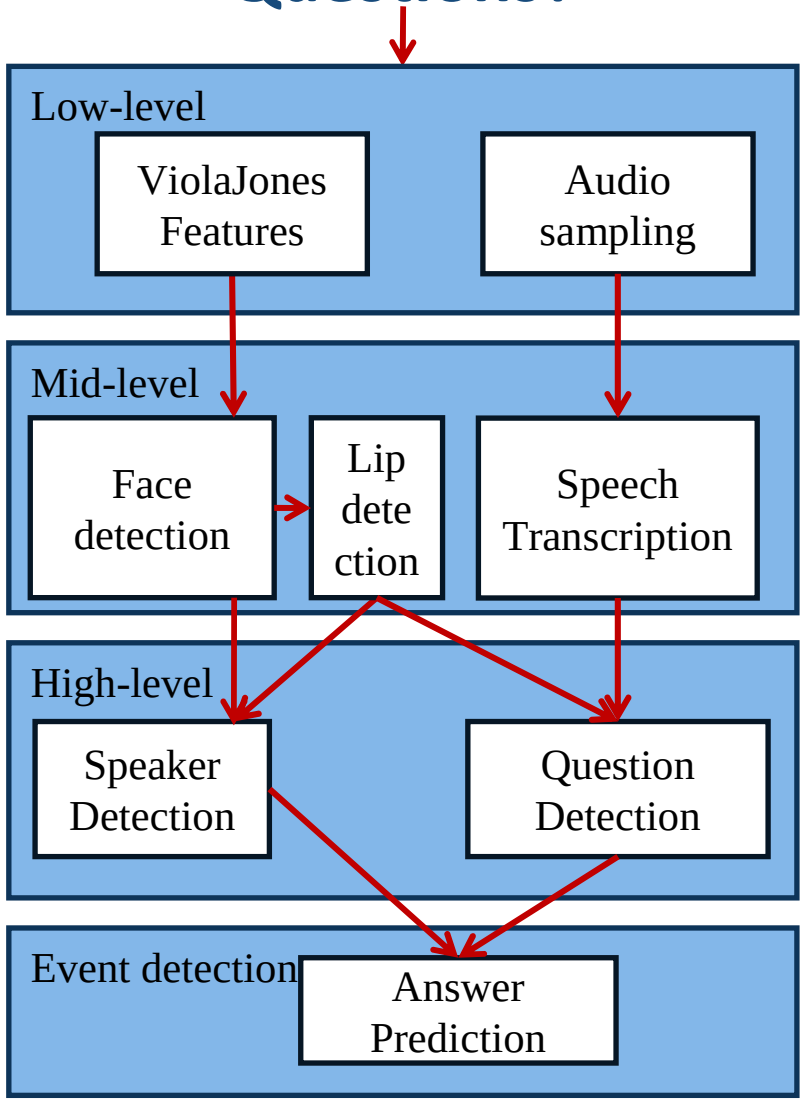
Tennis Point	Precision (%)	Recall (%)
Wimbledon	82.14	95.83
Roland Garros	84.44	95.00
US Open	97.67	85.71

Goal	Precision (%)	Recall (%)
Manchester United	100	100
Marseille	100	66.67

Conclusions

- System to detect events in broadcast sports video
- Generic (configurable ☺)
  - Training data (SVM)
  - Event description (FSM)
- Future work
  - semi-automatic selection of lower-level modules
  - new events/ new modules/ new sports

Questions?


Low-level
ViolaJones Features
Audio sampling
Mid-level
Face detection
Lip dete ction
Speech Transcription
High-level
Speaker Detection
Question Detection
Event detection
Answer Prediction