



A Spatio-Temporal and a Probabilistic Approach for Video Retrieval

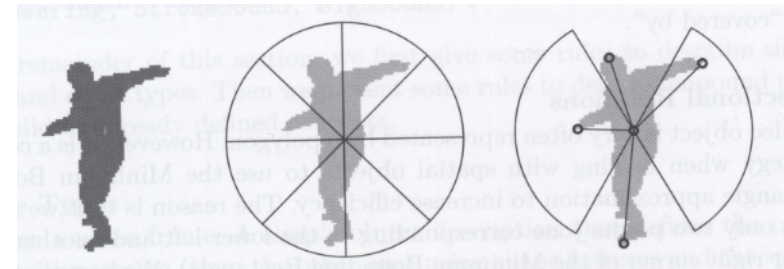
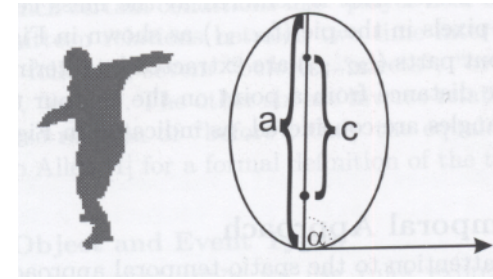
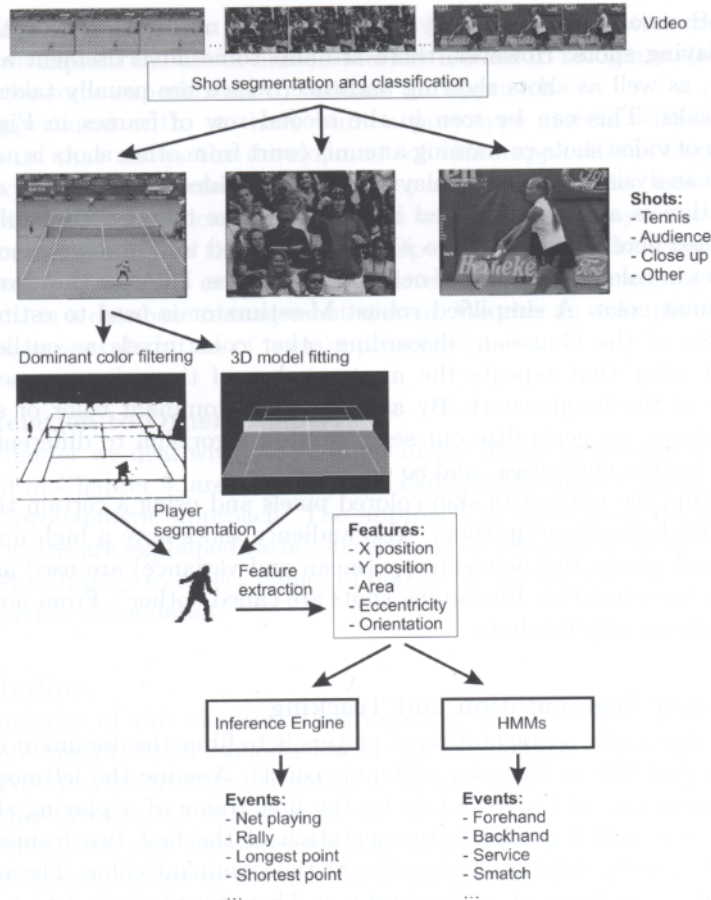
Zhirong Yang

5.4.2008

Introduction

- To extract high-level concepts from video footage
 - spatio-temporal approach
 - hidden Markov model
- Concentrate on the tennis case
- Empirical results with a prototype system are provided

Tennis Video Analysis





Spatio-temporal Approach

- Spatial Relations
 - topological: e.g. meet, equal, overlap
 - directional: e.g. north, east, north-east
- Temporal Relations
 - for example: before, during, starts, finishes

Rules for Object and Event Types

- **Feature types:**

{f1, f2, f3, f4, f5, f6, f7, f8, f9}

- **Visual object and audio event types:**

{SpatialObject, Ball, Net}

{Cheering, StrokeSound, SighSound}

- **Simple rules:**

PlayerCloserToCamera::=

{r1:SpatialObject, r2:rect(0,144,384,288)},

{700<f7(r1)<1200}, {contain(r2,r1)}

PlayerNearTheNet::=

{o1:PlayerCloserToCamera}, {o1.name="V.Williams"},

{y_distance(o1,Net)<50}, {duration>60}

cont.

- **Compound rules:**

ForehandTouch::=

{o1:PlayerCloserToCamera, o2:ball}, {s:StrokeSound},
{IsRighthanded(o1)}, {overlap(o1,o2), east(o2,o1)}

PlayerInRightCorner::=

{o1:PlayerCloserToCamera}, {f6.x(o1)>=190, f6.y(o1)>=170}

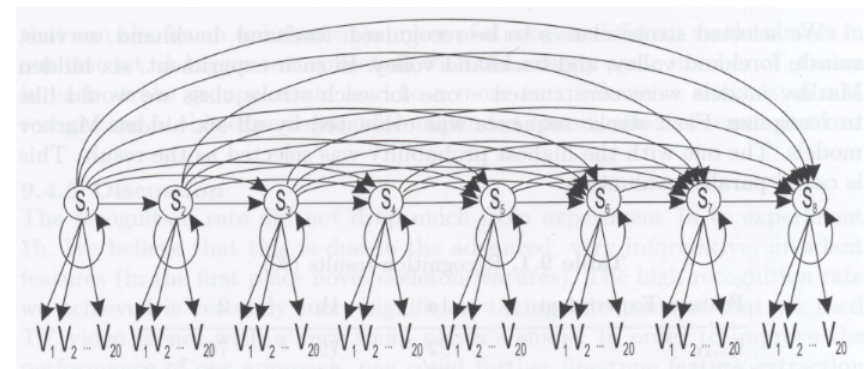
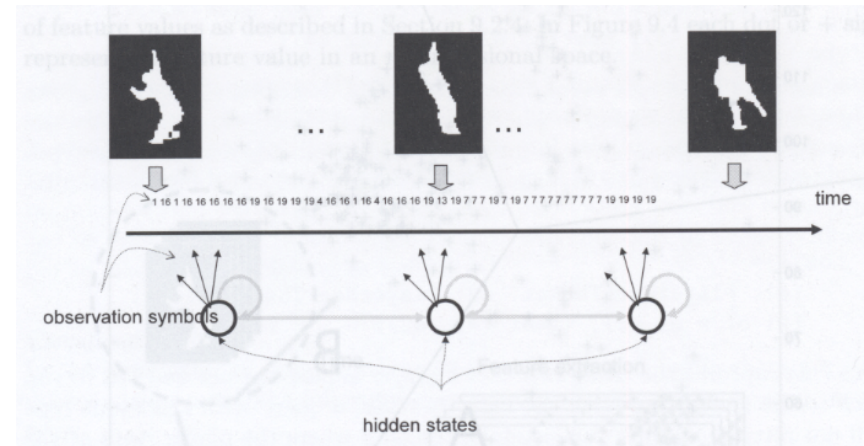
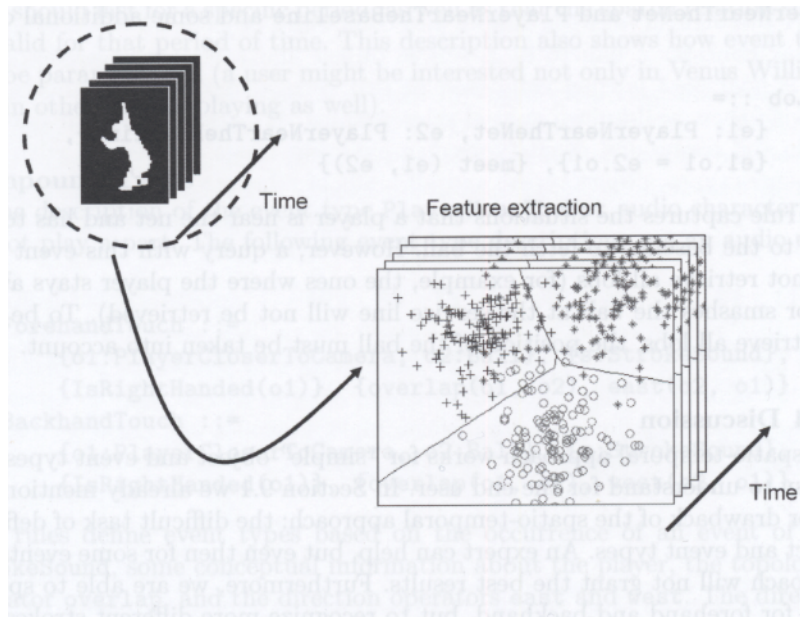
LtRRally::=

{e1,e2:PlayerInLeftCorner, e3,e4:PlayerInRightCorner},
{e1.o1=e2.o1=e3.o1=e4.o1},
{meet(e1,e3), meet(e3,e2), meet(e2,e4)}

Lob::=

{e1:PlayerNearTheNet, e2:PlayerNearTheBaseline}
{e1.o1=e2.o1}, {meet(e1,e2)}

Stroke Recognition Using Hidden Markov Models

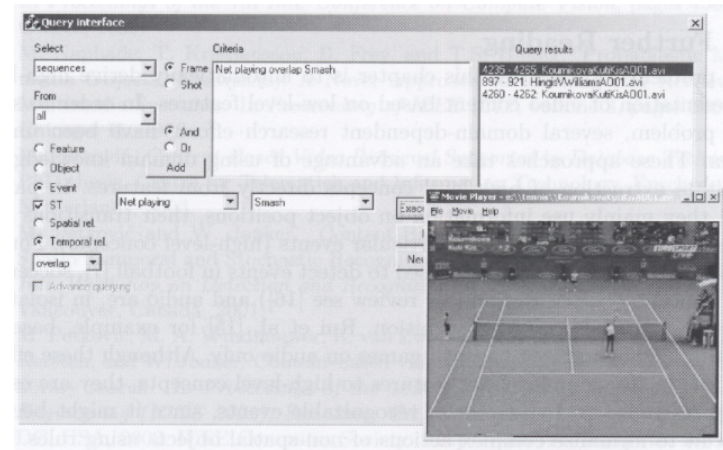
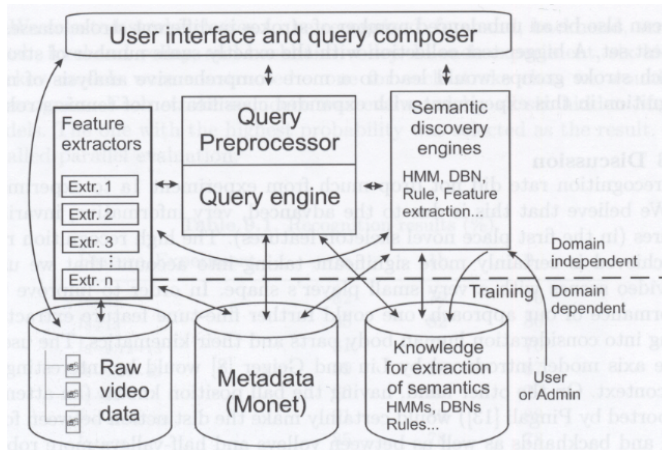


Recognition Results (%)

Feature/Experiment	1a	1b	2
f_{8-11}	82	79	76
f_{8-13}	85	82	80
$f_{8-9,12-13}$	81	78	76
$f_{8-9,22-23}$	89	88	87
f_{8-23}	86	82	79
$f_{9-11,22-23}$	91	89	88
f_{14-21}	85	78	78
f_{14-23}	93	87	86

- Recognizing strokes of 6 classes
 - 1a: same player in training and evaluation
 - 1b: different player groups for training and evaluation
- Recognizing strokes of 11 classes
 - 2: same setting as 1b, with new stroke classes added

Prototype



```
SELECT vi.frame_seq
FROM vidso vi
WHERE s_contains (vi.frame_seq,
    SmashOnNet=({e1: PlayerNearTheNet, e2:Smash},
    {overlap(e2,e1)}, e1.o1=e2.o1))=1 AND
vi.name='KournikovaKutikis'
```

Summary

- A framework for automatic extraction of high-level concepts from raw video data
 - rule-based component
 - stochastic component
- Integrated retrieval system for particular domain of tennis game videos
- The mapping from features to high-level concepts is still restricted, and more advanced techniques are required.