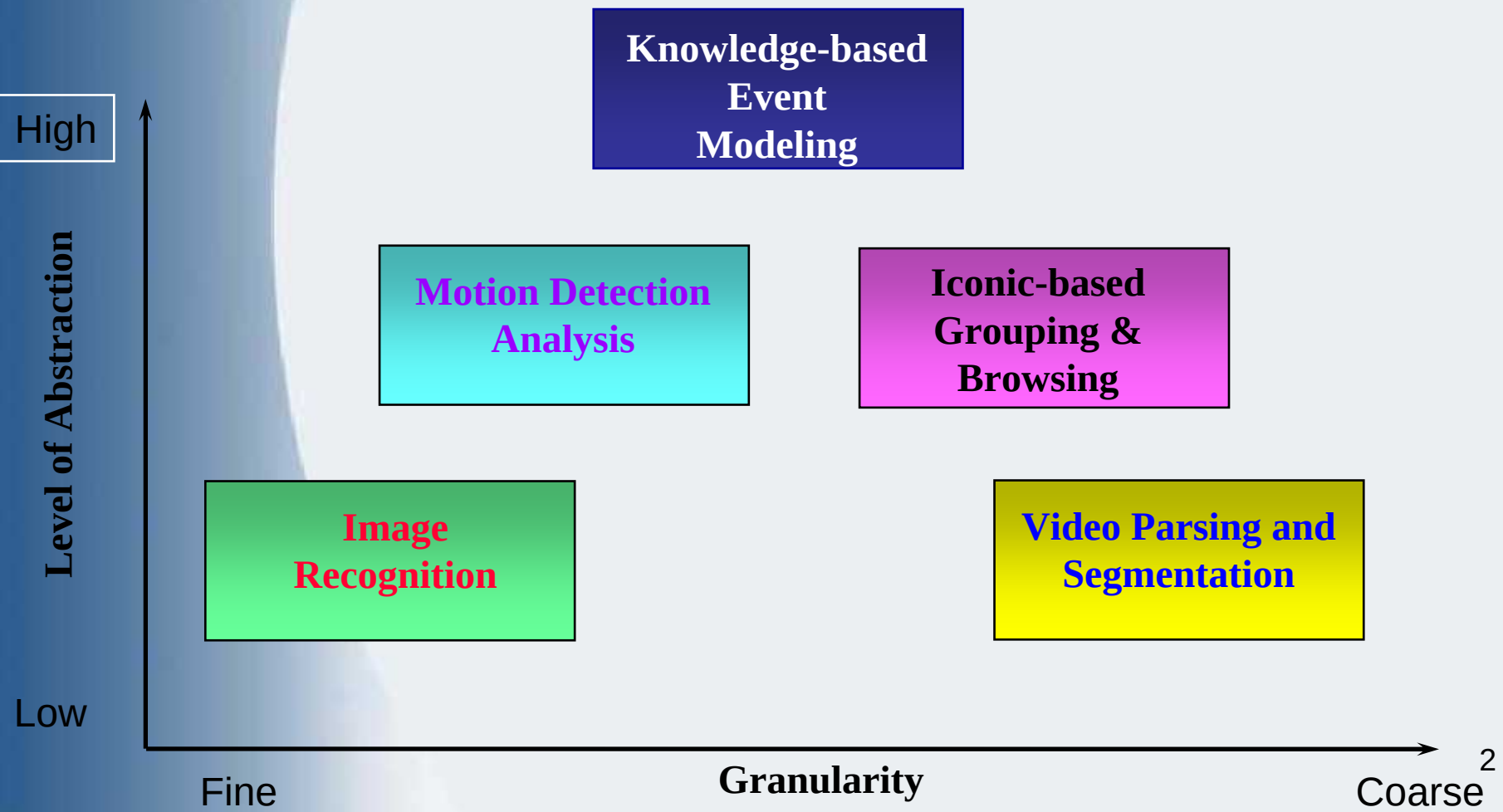
ECE624

Week 4-b

Video Semantic Modeling Paradigm



Video DBMS

Design and development of a video database management system that allows intelligent querying and searching

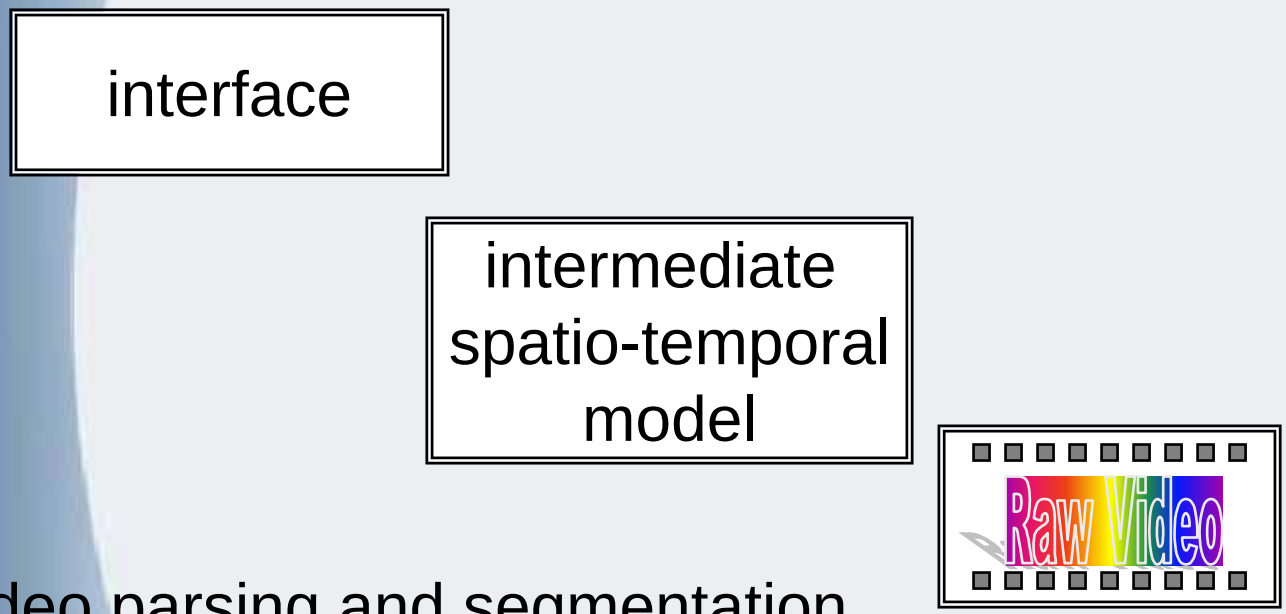
- Content-based searching of video data based on object motion
- A visual query interface to describe complex events in an intuitive manner

Example Video Queries

- “Find video clips where four T-33A’s fly in an arrow formation”
- “Find video clips where a Blue Angels aircraft draws a spiral trajectory”
- “Search for football clips that contains a touch-down”

Querying such events requires motion-based content description along with the use of conventional image features

Approach



- Video parsing and segmentation
- Low-level image-based searching by representative frames
- Motion-based modeling and indexing of video objects
- Spatial semantics based indexing of video objects

Video Parsing

- Scene change detection (uncompressed)
  - Abrupt vs. Gradual
    - Color histogram- and/or pixel-based difference.
  - Abrupt:
    - Nagasaka & Tanaka, Otsuji, Akutsu et al. Hsu et al.
  - Gradual:
    - Tonomura et al., Zhang et al., Shahraray, Zabih et al.

Scene change detection (compressed)

- DC Image Sequence Based (Princeton)
- DC Coefficients Based (Siemens)

Object Motion Tracking

- Color segmentation
  - Divide frame into predefined number of regions
  - Identify the individual color areas
- Motion analysis
  - Estimate motion between two consecutive frames

Object Motion Tracking

- Motion segmentation
  - Segment motion vector field
  - Identify the aggregate displacement of regions
- Motion compensation
- Rule-based technique to refine boundaries and distinguish between stationary and moving objects.

Motion-based Modeling and Indexing

Current Approaches

- STL & algebraic models: Italy, Purdue, MIT
- Trail model (VSDG): Purdue
- Trajectory model: ASU (Gulshani)

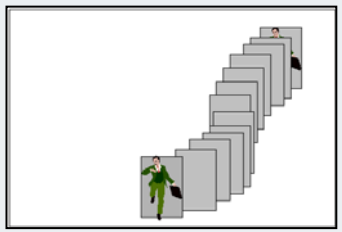
These formalisms have been extended by several researchers

Spatio-Temporal Logic

- Formal semantic modeling approach
- Efficient query processing
- Searching by exact comparisons
- Restricted to pre-defined formalisms
- Insufficient visual query support

Trail-based Modeling

- Bounding box salient objects
- Spatio-temporal information about objects
  - Mosaic representation of object motion
  - Attached time duration: added flexibility
- Object relationships can be represented in a graphical data structure (VSDG)

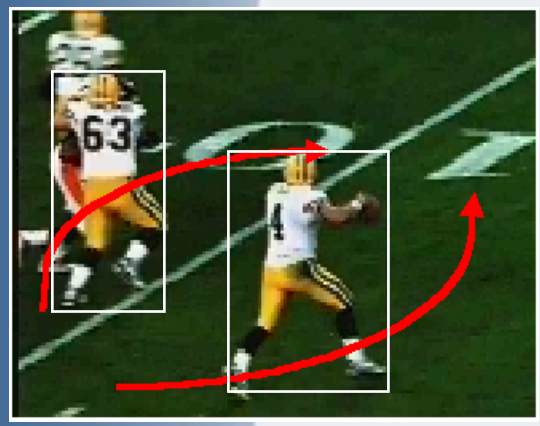


Trail Model: VSDG

Video **S**emantic **D**irected **G**raph

VSDG is a directed bipartite G=<V1, V2, E>
V1= Set of nodes representing objects in a video clip
D: V1 ---> I, mapping from nodes to set of duration in terms of number of frames
W: V1----> Z
where Z = Set of motion vectors associated with the bounding boxes of objects
V2 = Set of nodes marking the appearances of new object(s) of interest.

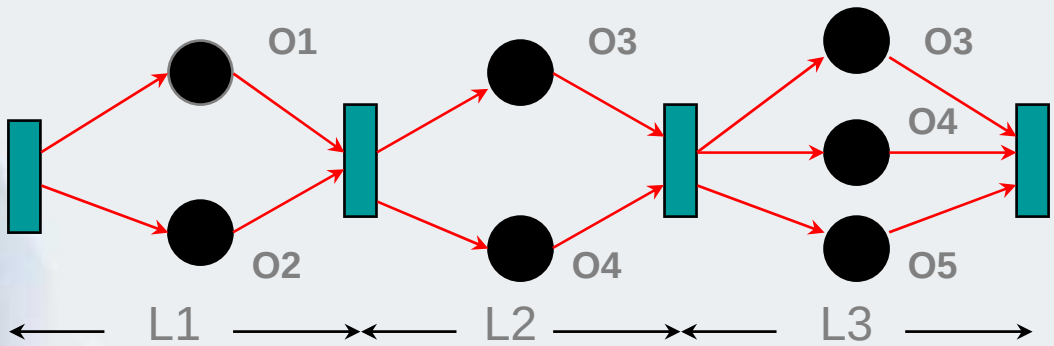
VSDG Example



L1

L2

L3



Trail Model

- Parallel to emerging video standards such as MPEG4 and MPEG7
- Practical and natural for visual querying
- Low precision
- Computationally expensive in query evaluation

Trajectory Model

- Special case of VSDG, where the centroid of each bounding box is tracked.
  - Good low-level support: MPEG
  - No searching technique exists.

Trail-based Search

Issues:

- Fuzziness
- Translational invariance
- Speed invariance
- Performance

Trail Search Types

Trail image comparison:

- Spatial absolute - exact screen location
- Spatial invariant - translation
  - Fourier-based convolution
- Scale invariant - object and trail sizes
  - Mellin-based pattern matching algorithm

Mellin Transform - Features

$$[M(x)](u,v) = \Sigma\ \Sigma\ x(k,l)\ k^{-(ju+1)}\ l^{-(jv+1)}$$

- Scale invariance:

$$x_{ab}(k,l)=x(ak,bl)$$

$$[M(x_{ab})](u,v) = a^{-ju}\ b^{-jv}[M(x)](u,v)$$

$$|[M(x_{ab})](u,v)|= |[M(x)](u,v)|$$

- Computational efficiency:

$$[M(x_{ab})](u,v) = [FFT(x_{ab})](\log u, \log v)$$

$$O(N \log N) \text{ complexity}$$

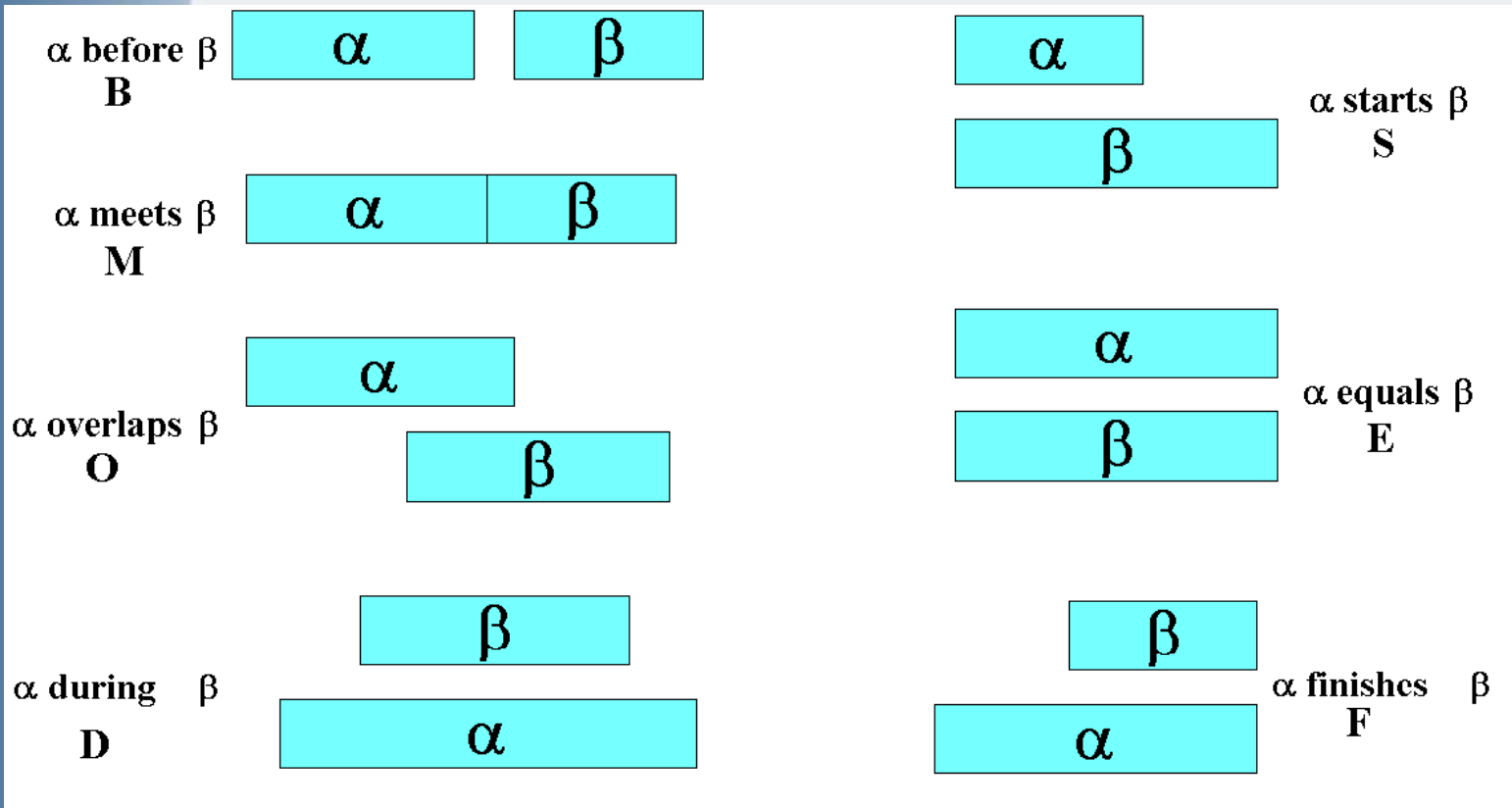
Problems - Issues

- Camera motion
  - Solution: Mosaic images
- Temporal information not preserved
  - Solution: Aggregate duration specification
    - Limited control on speed information

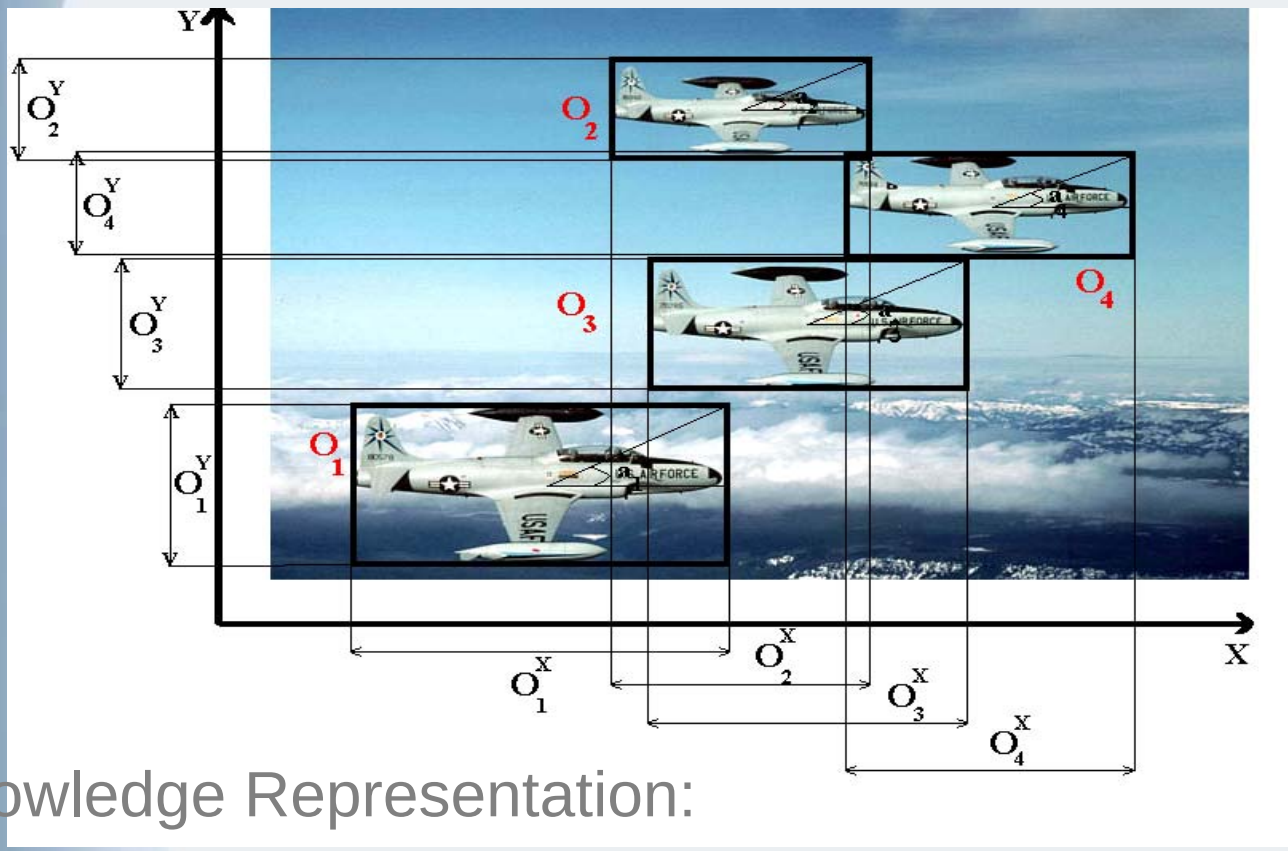
Spatial Semantic-based Search

Spatial semantics: Arrow formation

Binary Spatio/Temporal Relations



Example of Spatial Meta-Knowledge



Meta-Knowledge Representation:

$\mathbf{O}\ (O_1^x, O_2^x, O_3^x, O_4^x, \tau_{12}, \tau_{23}, \tau_{13}, \tau_{34})\ \wedge\ \mathbf{B}\ (O_1^y, O_3^y, O_4^y, O_2^y, \tau_1, \tau_2, \tau_3)$