A Spatio-temporal Database Model for Transportation Surveillance Videos

Sept. 11, 2006

Xin Chen, Chengcui Zhang
University of Alabama at Birmingham
U.S.A

Introduction and Motivation

- There is a proliferation of transportation surveillance videos.
- With object tracking techniques, trajectories of vehicles can be extracted. Analysis focuses on the spatio-temporal relations of vehicles.
- Spatio-temporal multimedia database models are general-purpose. There is a need for domain specific model that provides efficient indexing and query schema for transportation surveillance videos.
- The proposed model bases its structure on MATNs [1] and adopts the concept of CAI [2].

Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

Video Preprocessing

- An unsupervised segmentation method called Simultaneous Partition and Class Parameter Estimation (SPCPE) algorithm coupled with a background learning and subtraction method, is used to identify the vehicle objects in a traffic video sequence [3].
- The rectangular area is the Minimal Bounding rectangle (MBR) of the vehicle that is represented by (x_{low}, y_{low}) and (x_{high}, y_{high}) -- the coordinates of the bottom right point and the upper left point of the MBR. ($x_{centroid}$, $y_{centroid}$) are the coordinates of that vehicle segment's centroid. It is used for tracking the positions of vehicles the across video frames.

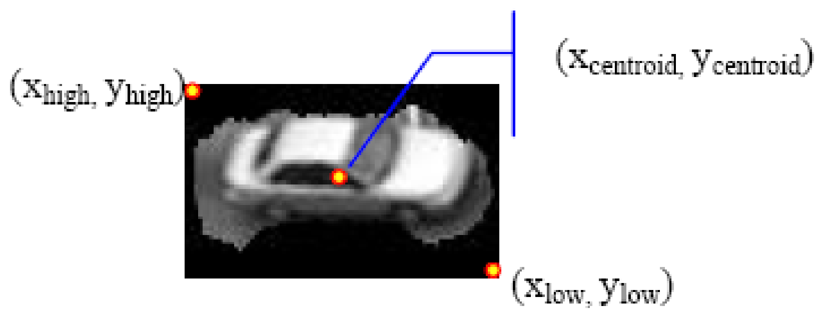


Figure 1 An example vehicle segment

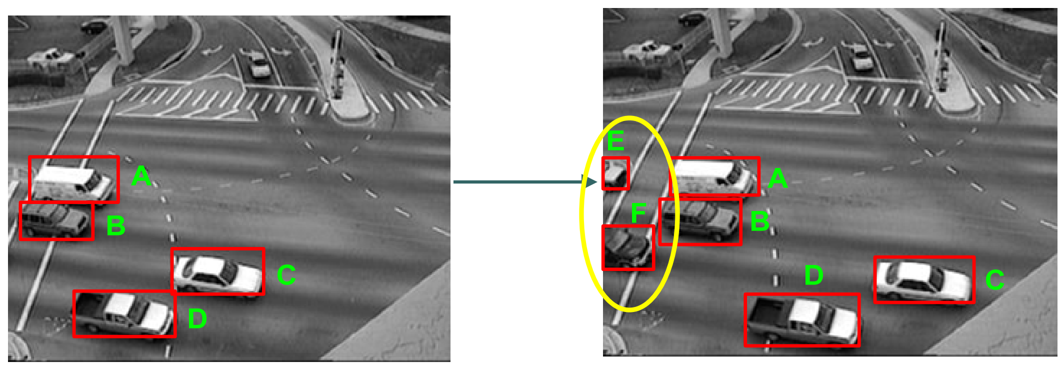
Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

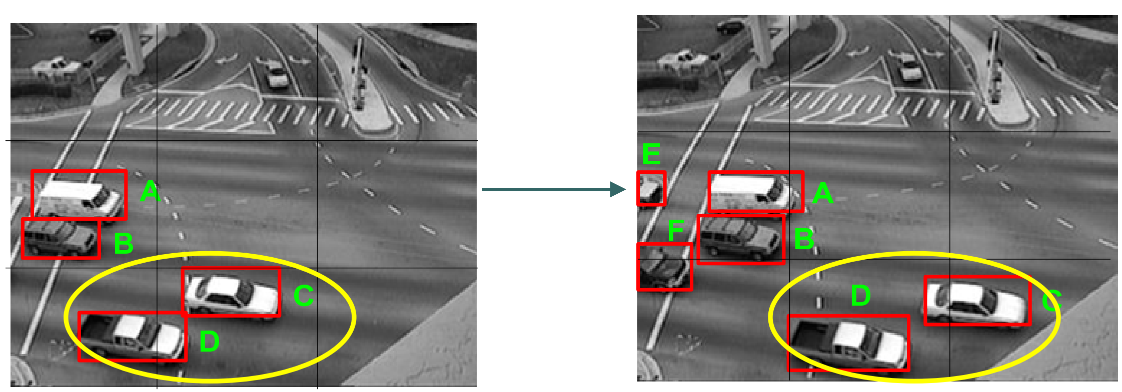
CAI for Modeling Media Streams

- It is unnecessary to record all frames as it will introduce redundancy. Videos shall be segmented and only key frames are recorded.
- This is not easy since transportation surveillance videos are continuous and do not contain obvious boundaries in between.
- Common Appearance Interval (CAI)[2] is an interval in which vehicle objects appear all together. A new *CAI* starts when there is a new vehicle appears in the video or an old one disappears or both.
- CAI has the flavor of “shot” in movies.

CAI Example



In the proposed model, CAI's are further divided into sub-intervals where relative positions of all vehicles remain unchanged.

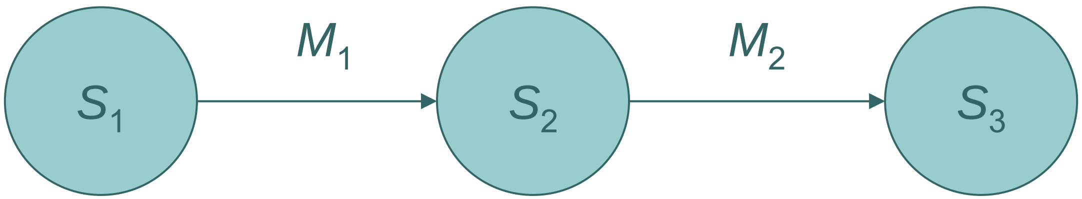


Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

MATN Based Structure

- Multimedia Augmented Transition Network
- MATN model is good at modeling the replay of multimedia presentations.
- It also provides an efficient mechanism in modeling the spatial relations of semantic objects in the video.



M: Media Streams

S: Transition States

Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

Media Streams

- MATN structure has main network and subnetwork.
- Media streams in main network is a CAI and media streams in sub-network is a sub-CAI
- Media streams in main network (CAI): (Vehicle ID)&… The symbol “&” means concurrent. *e.g. A&B&C*
- Media streams in sub network (sub-CAI): (VehicleID)(Relative Position)(Driving Direciton)&… e.g. A9NE&B1N&C20W

Media Streams -- Relative Positions

- Spatial relations of moving objects are recorded based on 27 three dimensional relative positions.
- Only 9 relative positions are used in our model for 2-D video sequence.
- Relative positions of vehicles in a video frame are used to record vehicle positions at coarse granularity.

Number	Relative Coordinates	Number	Relative Coordinates
1	$x_s \approx x_t, y_s \approx y_t, z_s \approx z_t$	15	$x_s < x_t, y_s < y_t, z_s > z_t$
2	$x_s \approx x_t, y_s \approx y_t, z_s < z_t$	16	$x_s < x_t, y_s > y_t, z_s \approx z_t$
3	$x_s \approx x_t, y_s \approx y_t, z_s > z_t$	17	$x_s < x_t, y_s > y_t, z_s < z_t$
4	$x_s \approx x_t, y_s < y_t, z_s \approx z_t$	18	$x_s < x_t, y_s > y_t, z_s > z_t$
5	$x_s \approx x_t, y_s < y_t, z_s < z_t$	19	$x_s > x_t, y_s \approx y_t, z_s \approx z_t$
6	$x_s \approx x_t, y_s < y_t, z_s > z_t$	20	$x_s > x_t, y_s \approx y_t, z_s < z_t$
7	$x_s \approx x_t, y_s > y_t, z_s \approx z_t$	21	$x_s > x_t, y_s \approx y_t, z_s > z_t$
8	$x_s \approx x_t, y_s > y_t, z_s < z_t$	22	$x_s > x_t, y_s < y_t, z_s \approx z_t$
9	$x_s \approx x_t, y_s > y_t, z_s > z_t$	23	$x_s > x_t, y_s < y_t, z_s < z_t$
10	$x_s < x_t, y_s \approx y_t, z_s \approx z_t$	24	$x_s > x_t, y_s < y_t, z_s > z_t$
11	$x_s < x_t, y_s \approx y_t, z_s < z_t$	25	$x_s > x_t, y_s > y_t, z_s \approx z_t$
2	$x_s < x_t, y_s \approx y_t, z_s > z_t$	26	$x_s > x_t, y_s > y_t, z_s < z_t$
13	$x_s < x_t, y_s < y_t, z_s \approx z_t$	27	$x_s > x_t, y_s > y_t, z_s > z_t$
14	$x_s < x_t, y_s < y_t, z_s < z_t$		

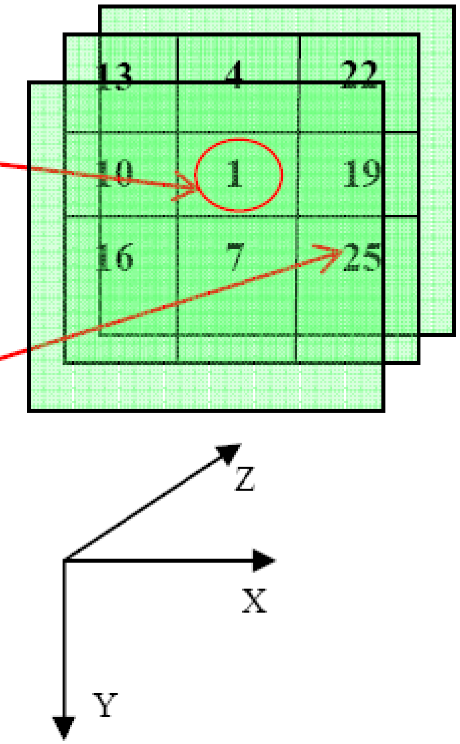


Figure 2 Three dimensional relative positions for vehicle objects.

(x_t, y_t, z_t) and (x_s, y_s, z_s) represent the X-, Y-, and Z-coordinates of the target and any semantic object, respectively. The '≈' symbol means the difference between two coordinates is within a threshold value.

Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

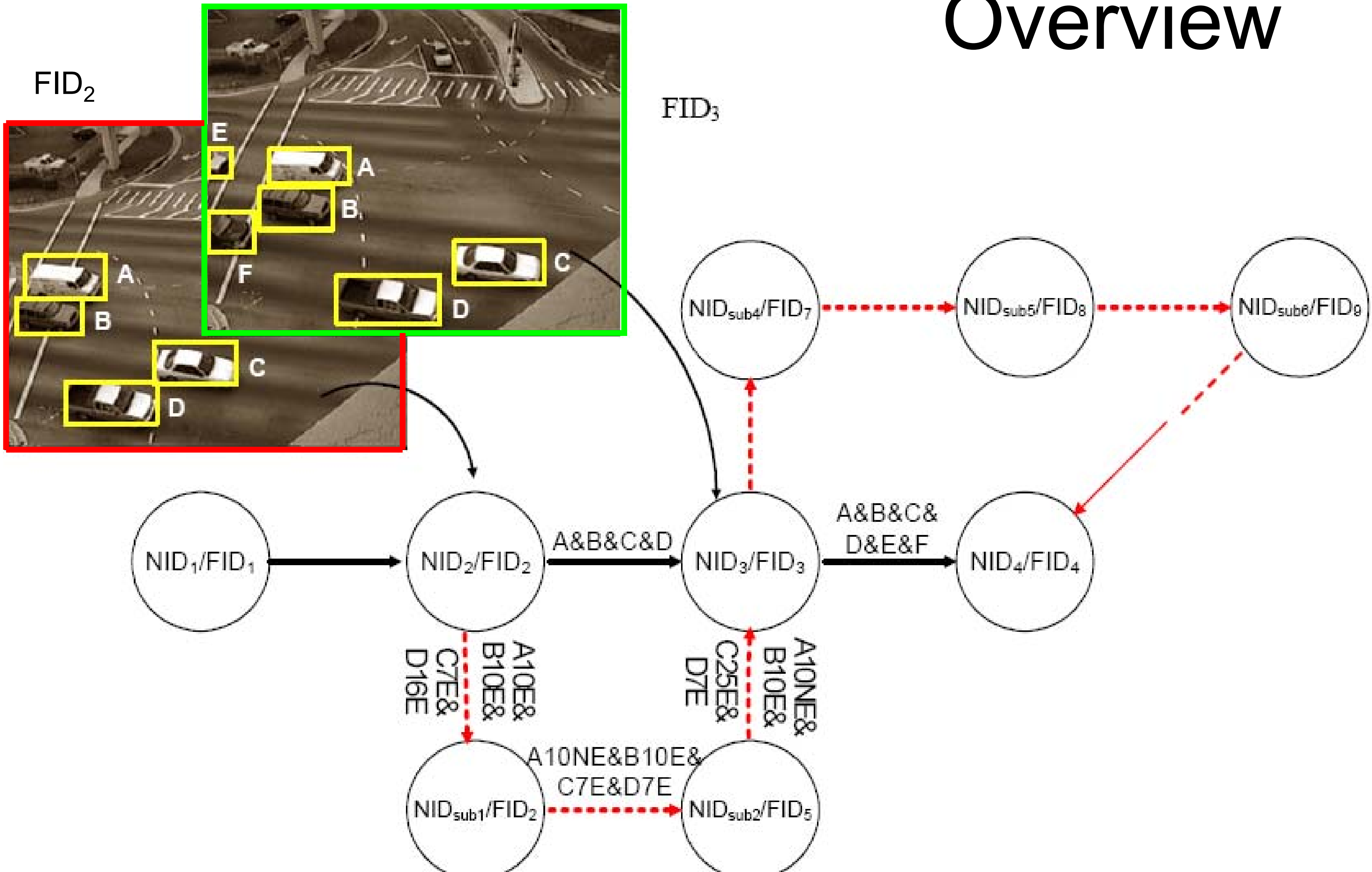
Transition States

- Nodes (states) in the main network -- 4-tuple *(NID, FID, OID_{in}, OID_{out}).*
  - *NID* is the node ID.
  - *FID* is the starting frame ID of the next CAI that is on the outgoing arc of this node (state).
  - OID_{in} is the list of IDs of the vehicle objects that newly appear in the next *CAI*.
  - OID_{out} is the list of IDs of the vehicle objects that disappear in the next *CAI*.
- Nodes (states) in the sub-network -- 2-tuple *(NID_{sub}, FID).*
  - NID_{sub} is the ID of a node in the subnetwork.
  - Each node is associated with a *FID* which is a frame ID. This frame is the starting frame of the next CAI_{sub} that is on the outgoing arc of this node (state).

Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

Overview

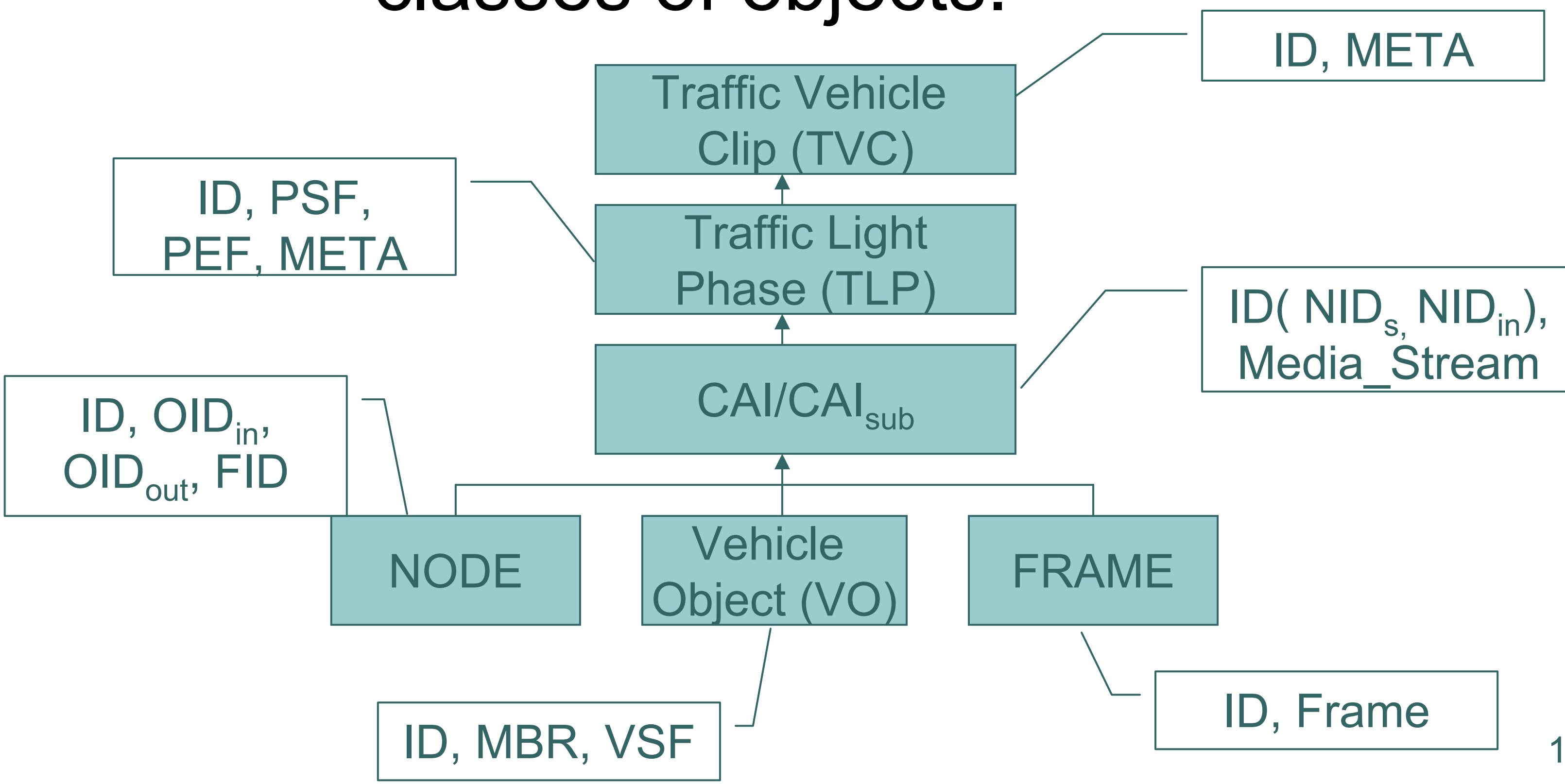

FID2
FID3
A
B
C
D
E
F
NID1/FID1
NID2/FID2
NID3/FID3
NID4/FID4
A&B&C&D
A&B&C&
D&E&F
NIDsub4/FID7
NIDsub5/FID8
NIDsub6/FID9
A10E&
B10E&
C7E&
D16E
A10NE&
B10E&
C25E&
D7E
NIDsub1/FID2
NIDsub2/FID5
A10NE&B10E&
C7E&D7E

Outline

- Video Preprocessing
- Proposed Model
  - CAI for Modeling Media Streams
  - MATN Based Structure
    - Media Streams
    - Transition States
  - Overview
- Database Queries
  - Queries
  - Examples

Object-oriented Database Management

- In the proposed model, there are 6 classes of objects.



Queries Strings

Table 2. Special Input Symbols

Special Symbol	Meaning
&&	Logical And
||	Logical Or
~	Logical Not
*	Wildcard
α	Arithmetic operators such as ‘+’, ‘-’…
β	Condition operators such as ‘<’, ‘>’, ‘= =’, ‘!=’…

Examples

- **Example 1: Find a vehicle that drives toward illegal direction in the traffic light phase when only north-bound and south-bound vehicles are allowed.**

  ~(CAIsub.(A*S)||CAIsub.(A*N))

- **Example 2: Find a vehicle that stops.**

  CAI.(A)&&

  (dist(A.mbr (CAI.NIDs.FID).centroid, A.mbr (CAI.NIDe.FID).centroid) /

  ((CAI.NIDs.FID –CAI.NIDe.FID) /TVC.META.FR) == 0)

Conclusion

- The proposed model is domain-specific.
  - Focus on transportation surveillance video database.
  - Target at the specific characteristics of transportation video.
  - Extract, index, and store the key information in the video.
  - Transportation video data can be efficiently accessed and queried.

- This model combines the strength of two general-purpose spatio-temporal database models – MATN and CAI.
  - Follow MATN’s basic structure and its way of modeling spatial relations among objects.
  - Adopt the concept of CAI.
  - Can better meet the needs of a transportation surveillance video database.

- Only frequently queried information is stored.
  - The relative spatial-relation of vehicles are only recorded at a coarse granularity based on MATN model.
  - The direction information of a moving vehicle is also recorded since this is a big concern of the user’s queries.
  - CAIs are further divided into sub-intervals which enables us to model the video streams at a finer granularity.

References

1. Chen, S.-C., and Kashyap, R. L. A spatio-temporal semantic model for multimedia database systems and multimedia information systems. *IEEE Transactions on Knowledge and Data Engineering,* vol. 13, no. 4, pp. 607-622, July/August 2001.

2. Chen, L., and Özsu, M. T. Modeling of video objects in a video database. In *Proc. IEEE International Conference on Multimedia,* Lausanne, Switzerland, August 2002, pp.217-221.

3. Chen, S.-C., Shyu, M.-L. , Peeta, S., and Zhang, C. Learning-Based spatio-temporal vehicle tracking and indexing for transportation multimedia database systems", *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, no. 3, pp. 154-167, September 2003.