Clustering Vehicle Trajectories with Hidden Markov Models Application to Automated Traffic Safety Analysis

Nicolas Saunier and Tarek Sayed

Department of Civil Engineering,
University of British Columbia

WCCI'06

saunier@civil.ubc.ca

Outline

1. Introduction to automated traffic safety analysis based on video sensors
2. Traffic conflict detection through semi-supervised learning
3. Experimental results
4. Future work

1. Motivation

- Traditional road safety is a reactive approach, based on historical collision data.
- Pro-active approach: "Don't wait for accidents to happen".
- Need for surrogate safety measures that
  - bring complementary information,
  - that can be easily collected,
  - are based on more frequent events,
  - are still related to safety (accidents).
- Traffic conflicts (near-misses). Video

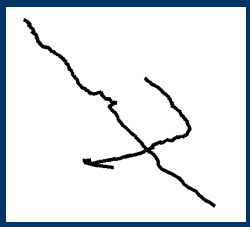
1. Video Sensors

- Main bottleneck of traffic conflict techniques
  - collection cost,
  - reliability and subjectivity of human observers.
- Advantages of video sensors
  - they are easy to install,
  - they can provide rich traffic description (vehicle tracking),
  - they can cover large areas,
  - they are cheap sensors.
- Computer vision is required to interpret video data.

1. Modular System

Vehicle Detection and Tracking

Conflicting Trajectories Detection



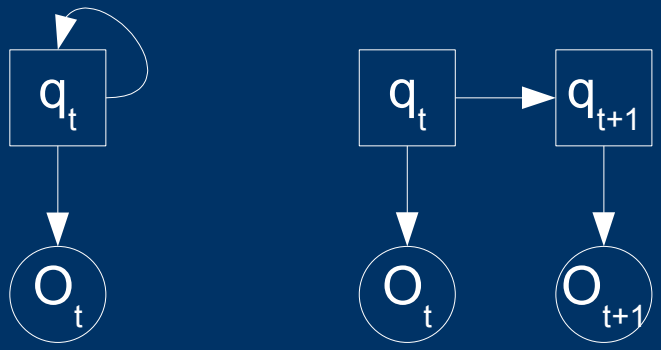
Implement a complete system

2. Detecting Traffic Conflicts

- Input
  - vehicle trajectories (x_1, y_1, ..., x_n, y_n), and velocities (vx_1, vy_1, ..., vx_n, vy_n).
- Output
  - actual traffic conflicts,
  - selected short sequences containing the traffic conflicts for further human review.
- Traffic conflicts are rare events. Data is limited for training and test.

2. Traffic Conflict Detection

- Direct extrapolation method is difficult because of imperfect tracking data.
- Learning is more generic
  - interaction classification,
  - learning and prediction of vehicle movements.
- Probabilistic models for sequential data: HMMs, DBNs.

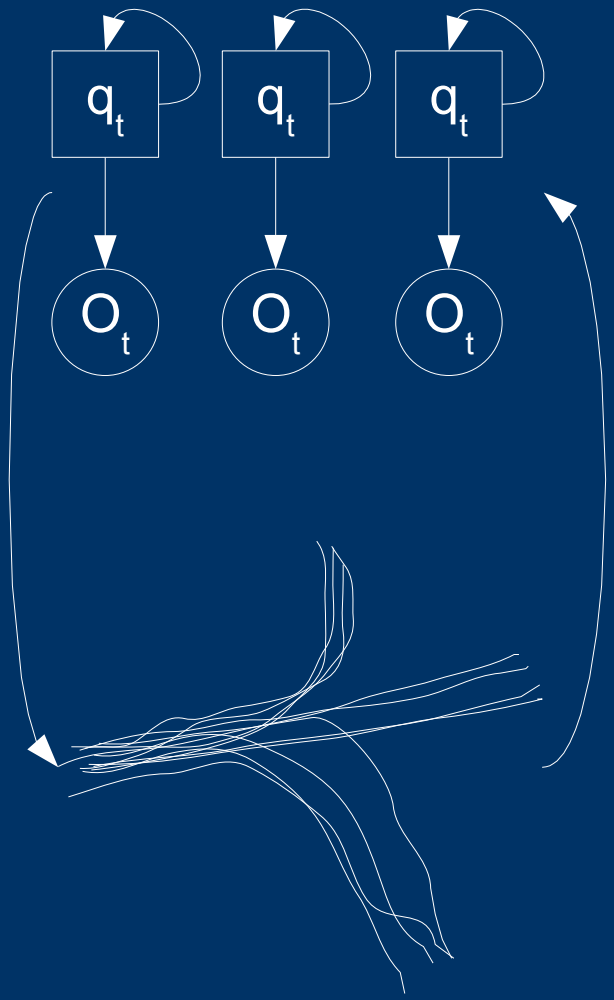


prior probabilities,
transition probabilities,
output distributions.

2. Sequential Data Clustering

- Sequence similarity: distance over sequences.
  - ex: edit distance, DTW, LCSS.
- Extract a set of features for each sequences, for use with traditional fixed length vector-based clustering methods.
  - ex: leading Fourier coefficients.
- Statistical sequence clustering: sequences are similar if they have a common similarity to a model, computed by the likelihood *P(Observation|Model)*.

HMM-based clustering of vehicle trajectories



train each HMM on its assigned trajectories

assign trajectories to HMMs

- K-means approach
- discard small clusters

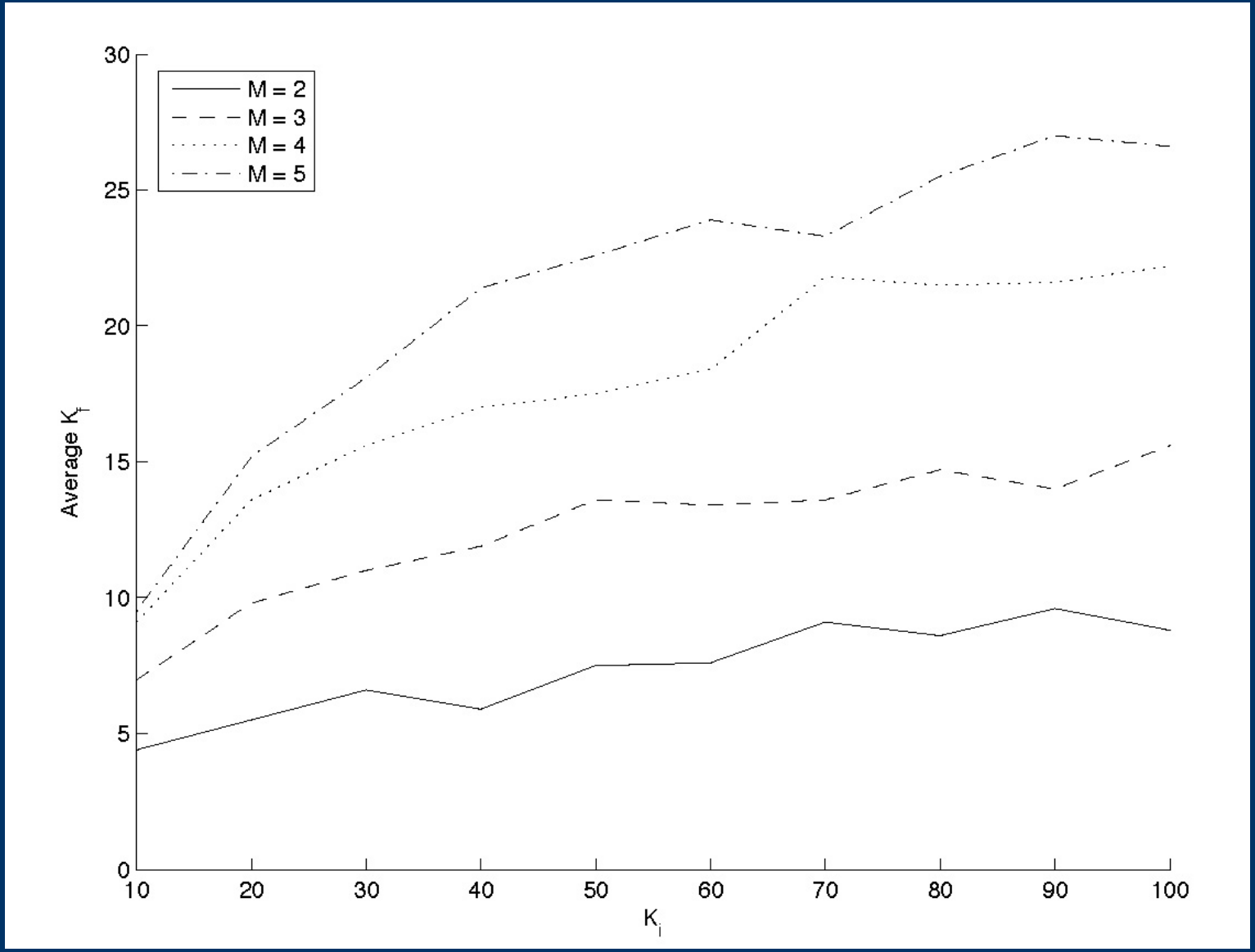
2. Semi-Supervised Learning

- Use of available traffic conflict instances
  - adaptation of HMMs to trajectories involved (means and covariances of the Gaussian output distributions)
  - memorization of "conflicting" models.
- Detection
  - interacting vehicles are detected
  - the trajectories are assigned to models
  - if the models were memorized as conflicting, a traffic conflict is detected

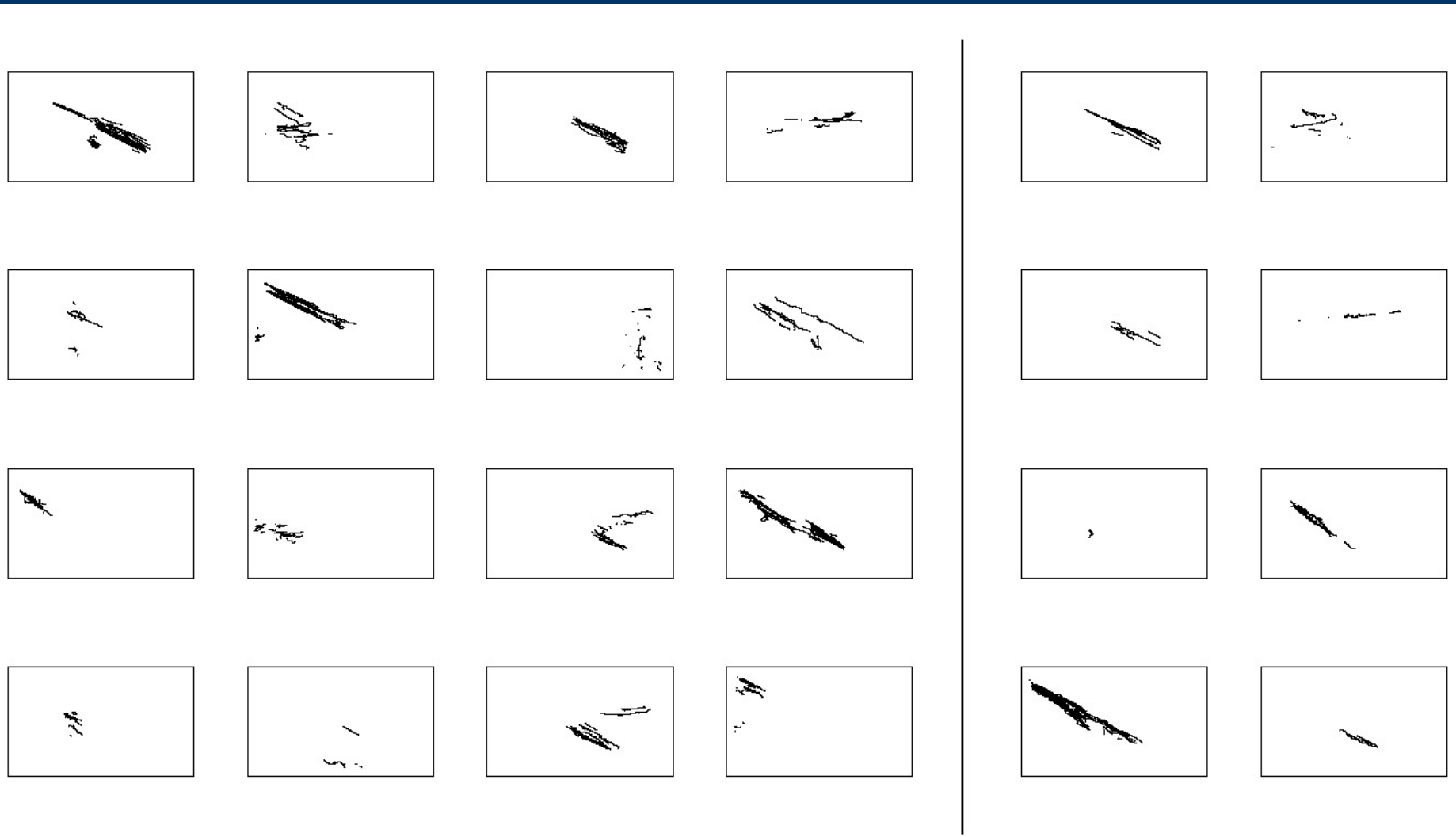
3. Experimental Results

- 10 video sequences used for the training of traffic conflict observers (1980s),
- 560 trajectories in 8 sequences used for learning,
- only 5 traffic conflicts.

3. Number of Models



3. Example of Trajectories Clustering



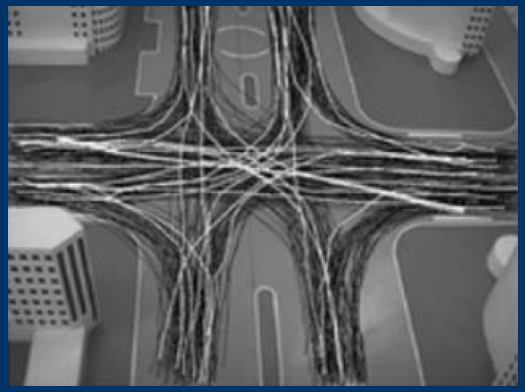
3. Detection Results

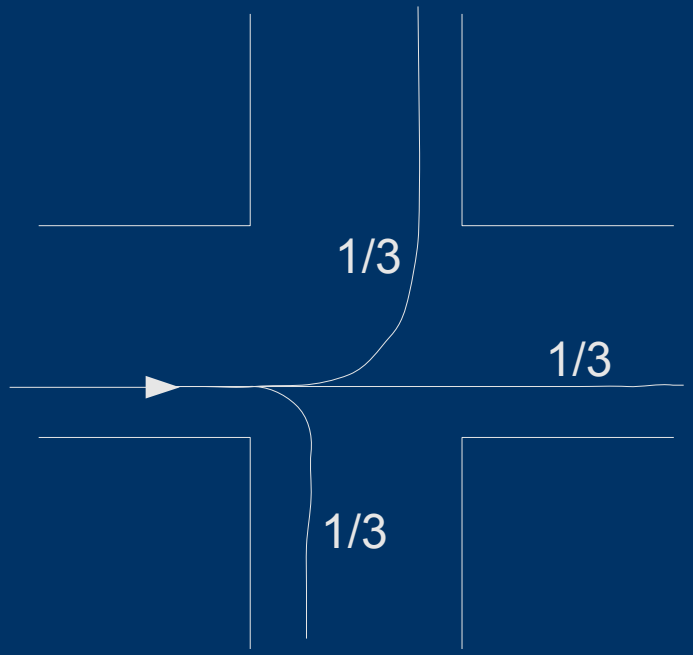
α	CD	Uncertain TC	FA
”0”	10	17	38
0.05	10	13	6
0.10	10	13	10
0.15	10	12	6
0.20	10	3	3
0.25	10	5	2
0.30	10	5	2
0.35	10	4	1
0.40	10	4	0
0.45	10	4	0
0.50	10	3	0

- HMM-based clustering is very sensitive to initialization.

4. Conclusion and Future Work

- Traffic conflict detection is feasible.
- Collecting more data
  - other sources,
  - artificial data,
  - interactive labeling, active learning.
- Collision probability computation.





Thank you !