Apprentissage automatique des trajectoires des usagers de la route pour la sécurité routière

Conférence du CIRRELT 2018

Nicolas Saunier
nicolas.saunier@polymtl.ca

10 mai 2018

POLYTECHNIQUE
MONTRÉAL

LE GÉNIE
EN PREMIÈRE CLASSE

Plan de la présentation

Introduction

Méthodes d’apprentissage des trajectoires

Résultats expérimentaux

Conclusion

Plan de la présentation

Introduction

Méthodes d’apprentissage des trajectoires

Résultats expérimentaux

Conclusion

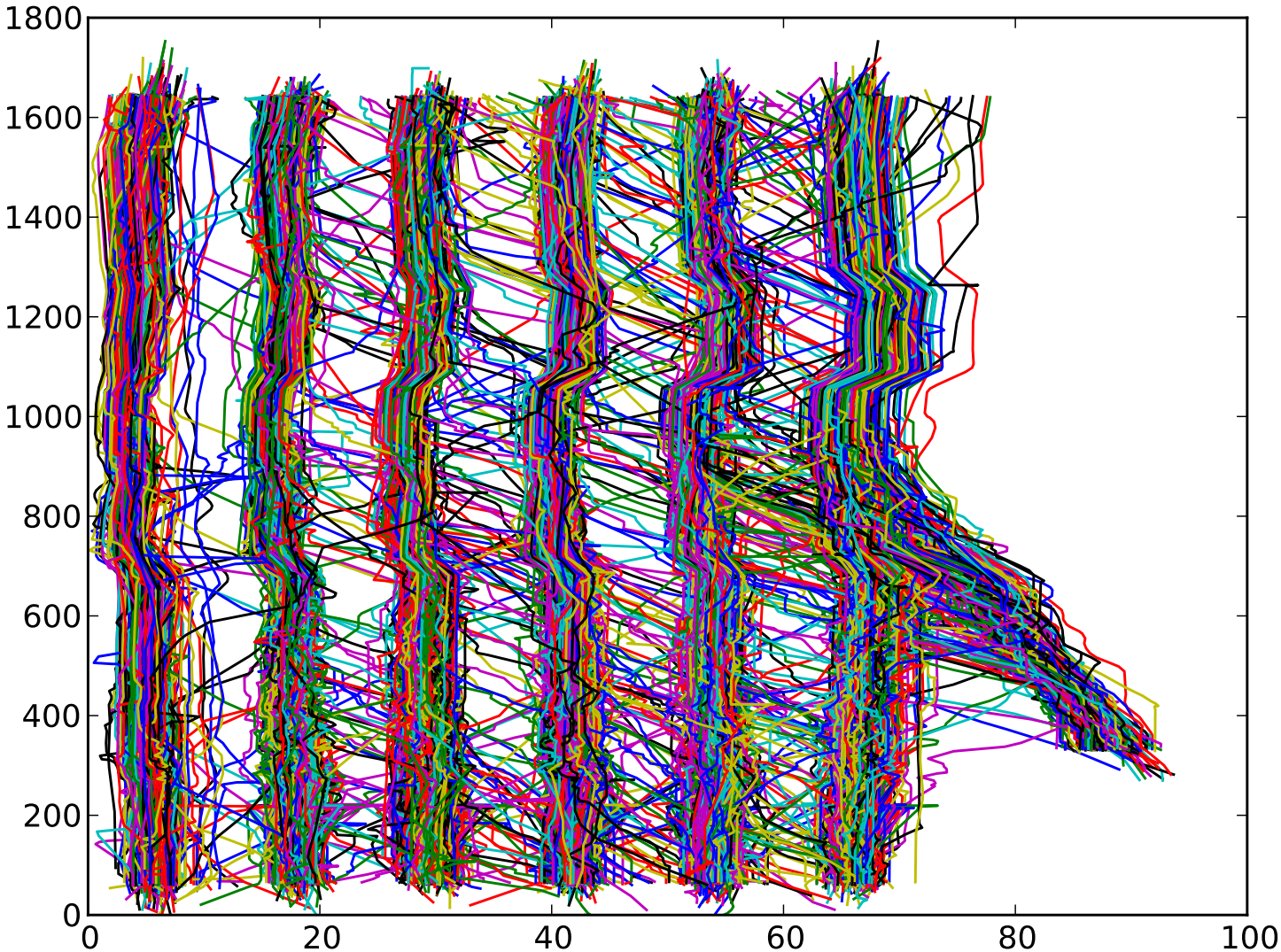
Quel est le type de données fondamental en transport?

Quel est le type de données fondamental en transport?

$$(t_i, x(t_i), y(t_i))$$

Exemples

Exemples



Données de trajectoires

- Des trajectoires d'usagers des transports et des marchandises sont de plus en plus facilement disponibles
  - capteurs GNSS (GPS), vidéo et autres méthodes d'identification unique (plaque d'immatriculation, Bluetooth, RFID, etc.)

Données de trajectoires

- Des trajectoires d'usagers des transports et des marchandises sont de plus en plus facilement disponibles
  - capteurs GNSS (GPS), vidéo et autres méthodes d'identification unique (plaque d'immatriculation, Bluetooth, RFID, etc.)
- Objectifs: compréhension des comportements, en particulier
  - groupes de trajectoires ou patrons de mouvements
  - détection d'anomalies
  - modèles de prédiction des mouvements des usagers

Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

1. méthode traditionnelle à base de données d'accidents

Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

1. méthode traditionnelle à base de données d'accidents
    - reconstitution des événements

Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

1. méthode traditionnelle à base de données d'accidents
    - reconstitution des événements
    - approche réactive

Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

1. méthode traditionnelle à base de données d'accidents
    - reconstitution des événements
    - approche réactive
2. méthode alternative à base d'observations de la circulation

Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

1. méthode traditionnelle à base de données d'accidents
    - reconstitution des événements
    - approche réactive
2. méthode alternative à base d'observations de la circulation

    - observation directe des événements, plus rapide

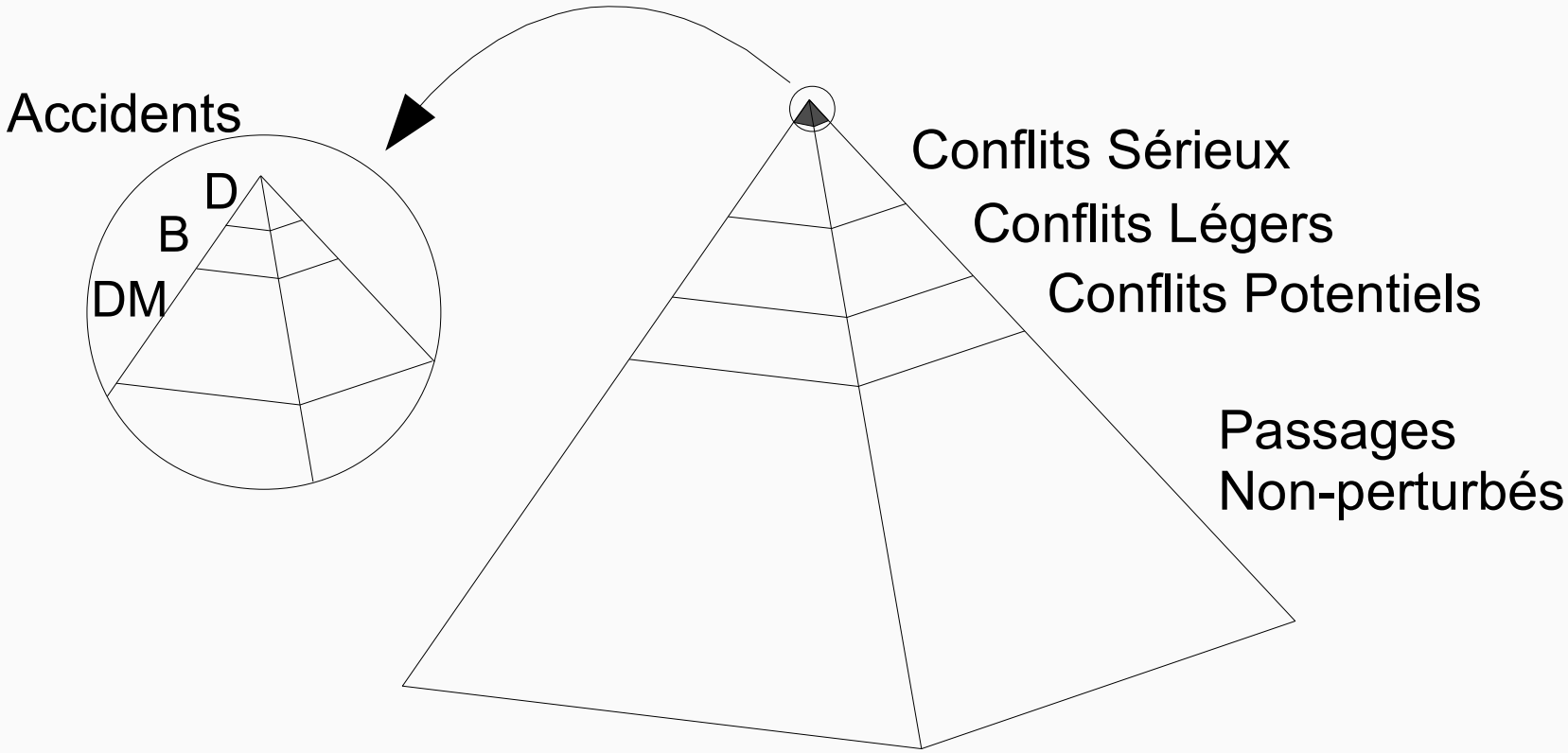
Méthodes de diagnostic de la sécurité routière

Il existe deux grandes catégories de méthodes de diagnostic de la sécurité routière

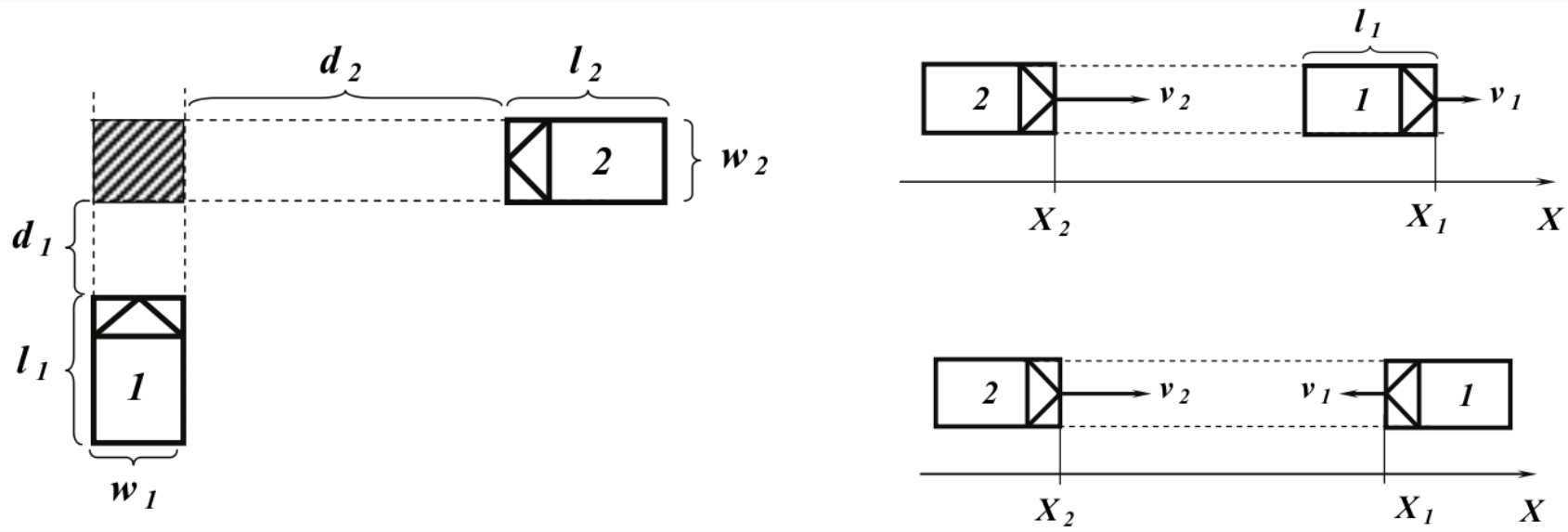
1. méthode traditionnelle à base de données d'accidents
   - reconstitution des événements
   - approche réactive
2. méthode alternative à base d'observations de la circulation

   - observation directe des événements, plus rapide
   - approche proactive

La hiérarchie de la sécurité des interactions



Indicateurs de sécurité des usagers



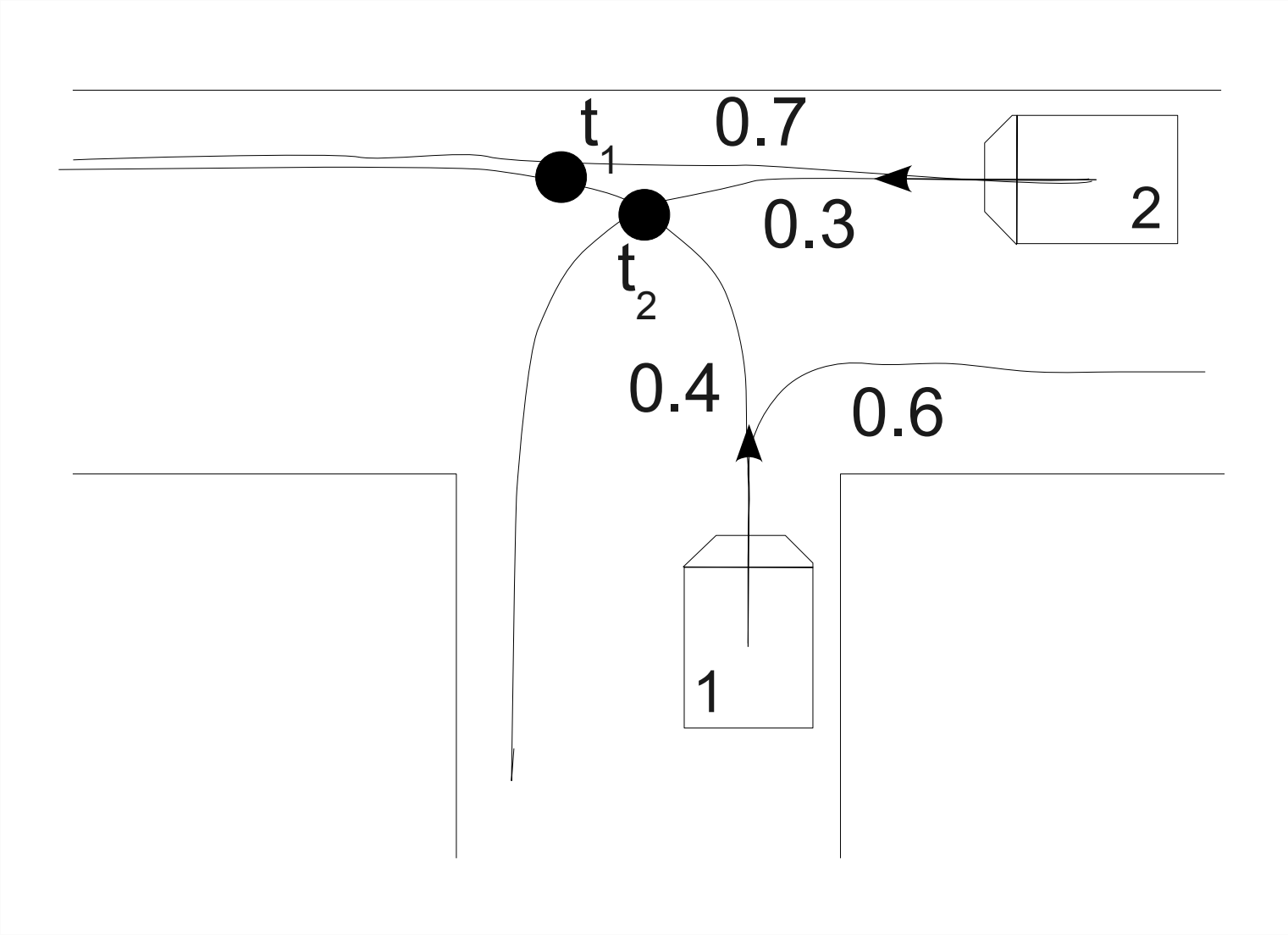
$$TTC = \frac{d_2}{v_2} \; if \; \frac{d_1}{v_1} < \frac{d_2}{v_2} < \frac{d_1 + l_1 + w_2}{v_1}$$

$$TTC = \frac{d_1}{v_1} \; if \; \frac{d_2}{v_2} < \frac{d_1}{v_1} < \frac{d_2 + l_2 + w_1}{v_2} \; (lateral)$$

$$TTC = \frac{X_1 - X_2 - l_1}{v_1 - v_2} \; if \; v_2 > v_1 \; (arriere)$$

$$TTC = \frac{X_1 - X_2}{v_1 + v_2} \; (de \; face)$$

Méthodes de prédiction des mouvements des usagers



Plan de la présentation

Introduction

Méthodes d’apprentissage des trajectoires

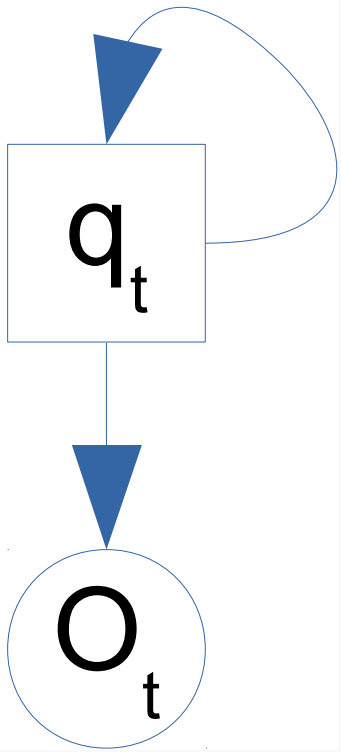
Résultats expérimentaux

Conclusion

Caractéristiques des données

- Erreurs sur les positions (GNSS, vidéo)
- Le taux d'échantillonnage peut varier dans une même trajectoire et entre trajectoires
- Séquences de longueur variable

Méthodes de regroupement de données séquentielles



- Extraction d'un nombre fixe de caractéristiques pour chaque séquence, par ex. coefficients de Fourier, pour appliquer des méthodes classiques
- Mesures de similarité pour des séquences
- Méthodes statistiques de regroupement de séquences: des séquences sont similaires si elles sont générées par le même modèle, mesuré par la vraisemblance $P(\mathit{observation}|\mathit{modele})$

Mesures de similarité

- Distances entre vecteurs de même taille: distance Euclidienne, etc.
  - nécessite un pré-traitement des données, sensible au bruit
- Distances “élastiques”
  - distance d’édition, “dynamic time warping” (DTW), “longest common subsequence” (LCSS)

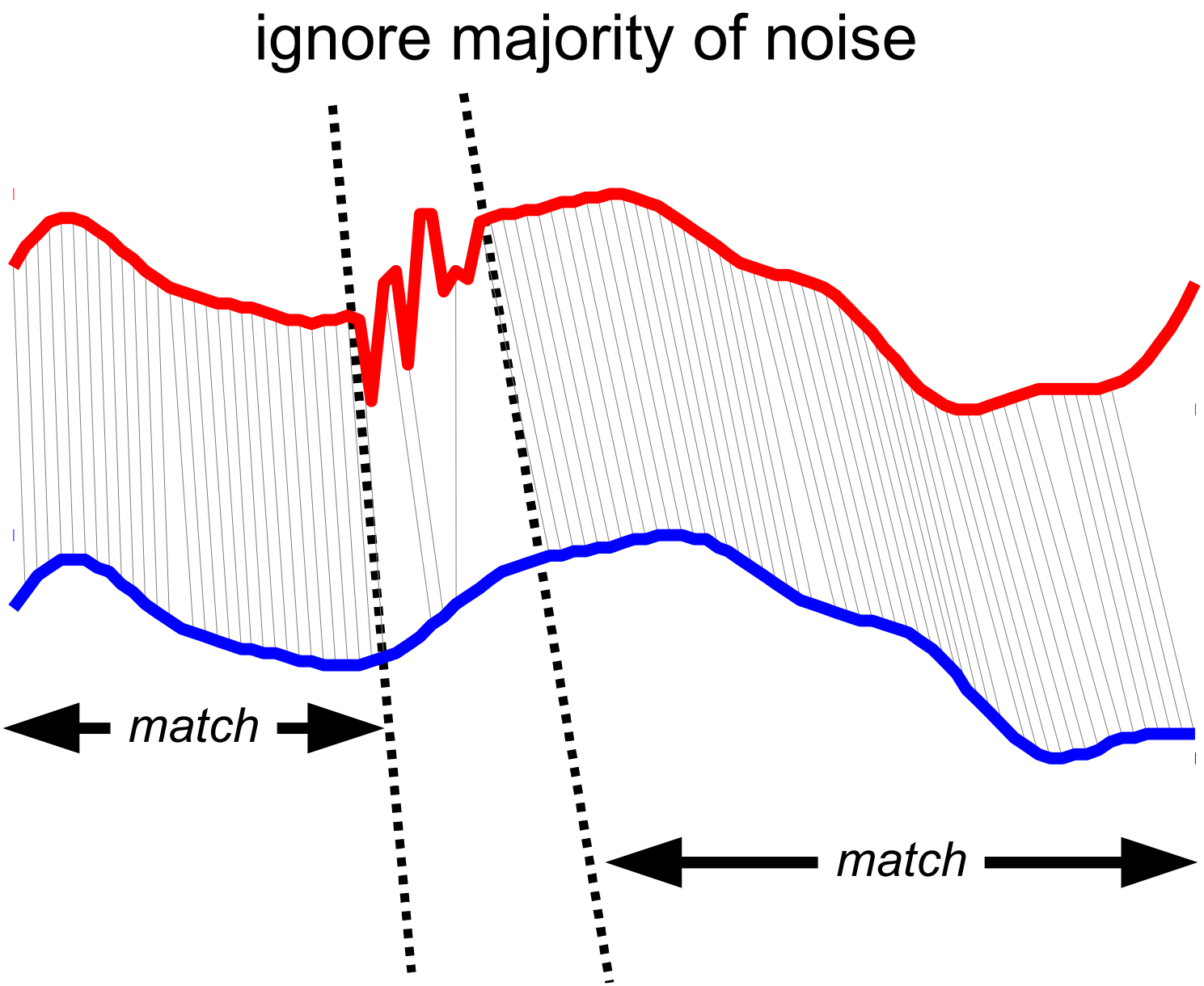
LCSS

- La LCSS est une distance d'édition modifiée, utilisée par ex. pour la correction d'orthographe, la reconnaissance d'écriture manuscrite, l'appariement de séquences d'ADN
- La LCSS est robuste au bruit
  - les séquences sont appariées avec des distortions, sans changer l'ordre des éléments, mais en permettant à certains de ne pas être appariés
- La LCSS est très flexible
  - la similarité est subjective et dépend de l'application

LCSS

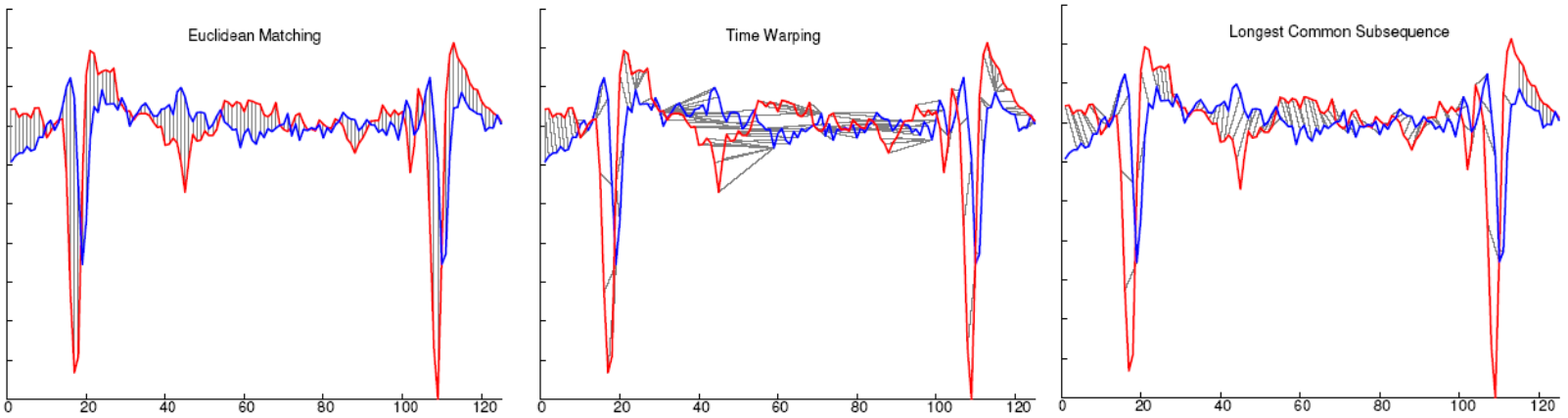
- Trajectoire à des pas de temps réguliers: $P_i = \{p_{i,1}, ...p_{i,n}\}$ où $p_{i,k} = (x_{i,k}, y_{i,k})$
- $Head(P_i) = \{p_{i,1}, ...p_{i,n-1}\}$
- Soit un seuil $\epsilon > 0$, P_i et P_j deux trajectoires de longueurs m et n, $LCSS_\epsilon(P_i, P_j)$ est défini par
  - 0 si $m = 0$ ou $n = 0$
  - $1 + LCSS_\epsilon(Head(P_i), Head(P_j))$ si les points $p_{i,n}$ et $p_{j,m}$ sont similaires
  - $max(LCSS_\epsilon(Head(P_i), P_j), LCSS_\epsilon(P_i, Head(P_j)))$ sinon
- Exemple de similarité entre points: p_{i,k_1} et p_{j,k_2} sont similaires si $|x_{i,k_1} - x_{j,k_2}| < \epsilon$ et $|y_{i,k_1} - y_{j,k_2}| < \epsilon$
- Distance: $DLCSS_\epsilon(P_i, P_j) = 1 - \left(\frac{LCSS_\epsilon(P_i,P_j}{min(n,m)}\right)$

LCSS



LCSS

Method	Complexity	Elastic Matching	One-to-one Matching	Noise Robustness
Euclidean	*O(n)*	×	✓	×
DTW	*O(n*δ)*	✓	×	×
LCSS	*O(n*δ)*	✓	✓	✓



Éléments d'un algorithme de regroupement ("clustering")

Algorithme de type k-moyennes, où on spécifie une mesure seuil de similarité plutôt que le nombre de groupes

1. Mesure de similarité (distance)
2. Représentation des groupes
3. Méthode de mise à jour du groupe

Éléments d'un algorithme de regroupement ("clustering")

Algorithme de type k-moyennes, où on spécifie une mesure seuil de similarité plutôt que le nombre de groupes

1. Mesure de similarité (distance): mesure LCSS
2. Représentation des groupes
3. Méthode de mise à jour du groupe

Éléments d'un algorithme de regroupement ("clustering")

Algorithme de type k-moyennes, où on spécifie une mesure seuil de similarité plutôt que le nombre de groupes

1. Mesure de similarité (distance): mesure LCSS
2. Représentation des groupes: trajectoires prototypes
3. Méthode de mise à jour du groupe

Éléments d'un algorithme de regroupement ("clustering")

Algorithme de type k-moyennes, où on spécifie une mesure seuil de similarité plutôt que le nombre de groupes

1. Mesure de similarité (distance): mesure LCSS
2. Représentation des groupes: trajectoires prototypes
3. Méthode de mise à jour du groupe: garder la trajectoire la plus longue

Plan de la présentation

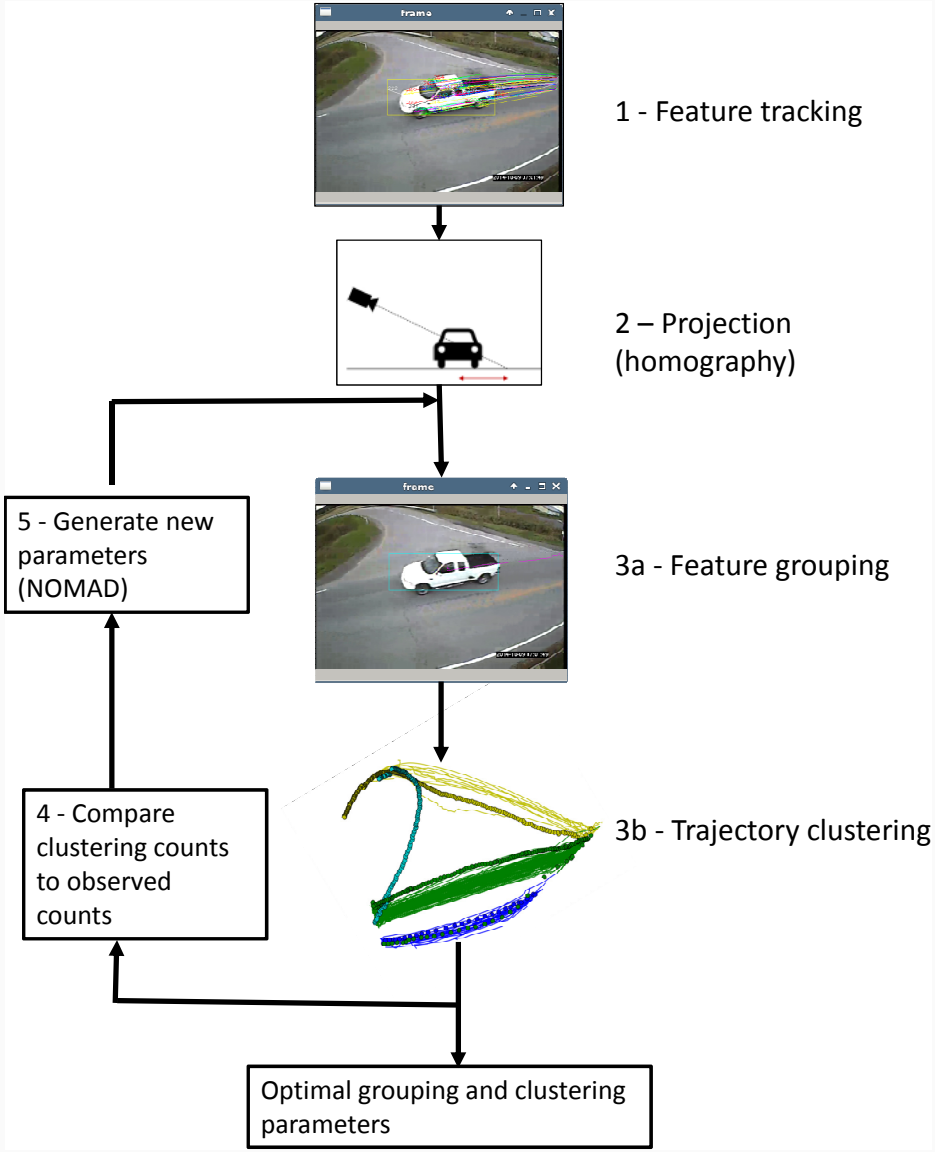
Introduction

Méthodes d’apprentissage des trajectoires

Résultats expérimentaux

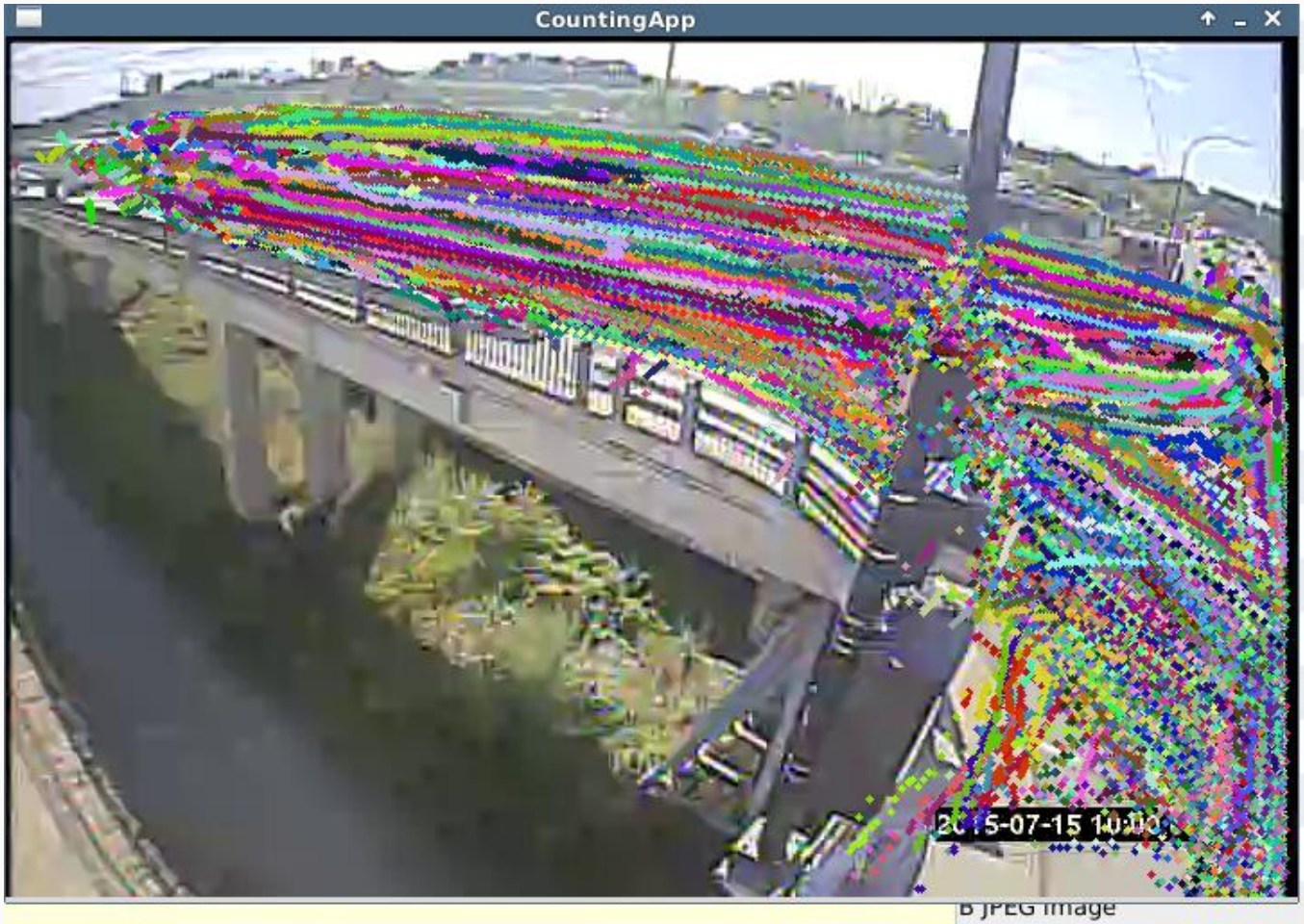
Conclusion

Application aux comptages de mouvements tournants

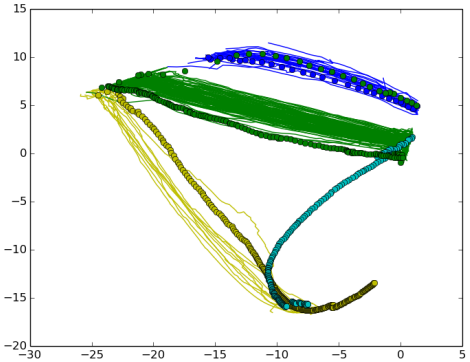


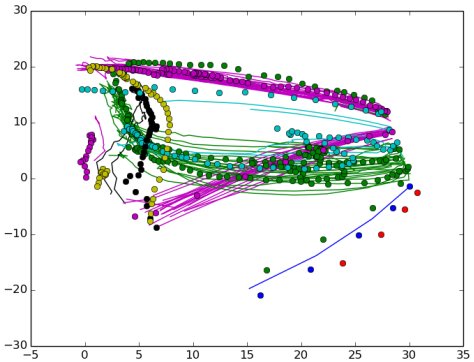
avec F. Bélisle (WSP), G.-A. Bilodeau et S. le Digabel

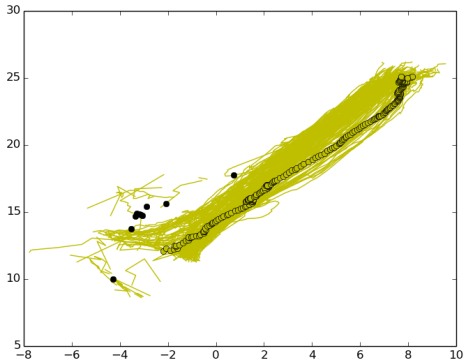
Application aux comptages de mouvements tournants



Application aux comptages de mouvements tournants







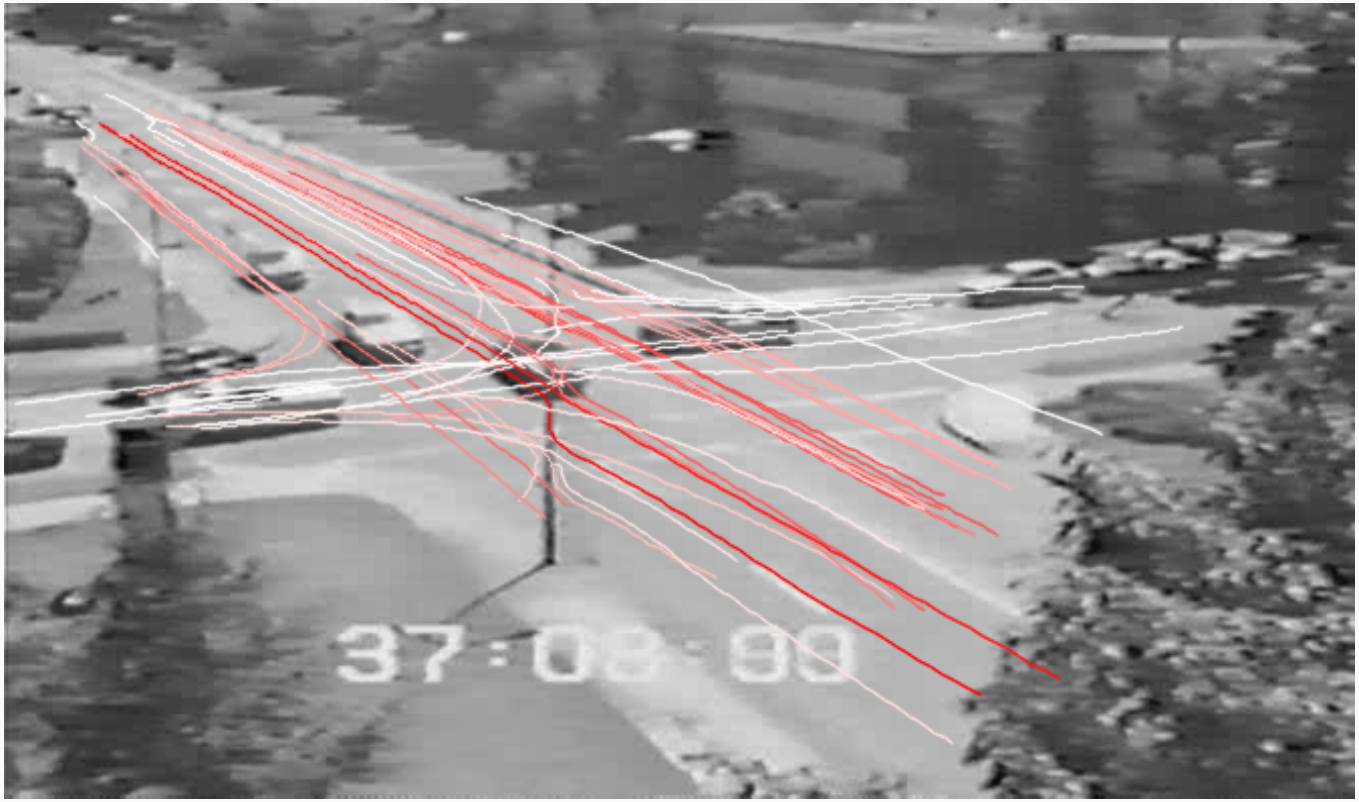
Application aux comptages de mouvements tournants

ID	Town	# Mvts	Training Err.	Auto. Gen. Err.
030	Montréal	1	0.7 %	10.8 %
004	Ste-Cat. de Hatley	4	0 %	9.5 %
006	Westbury	6	12 %	19.2 %

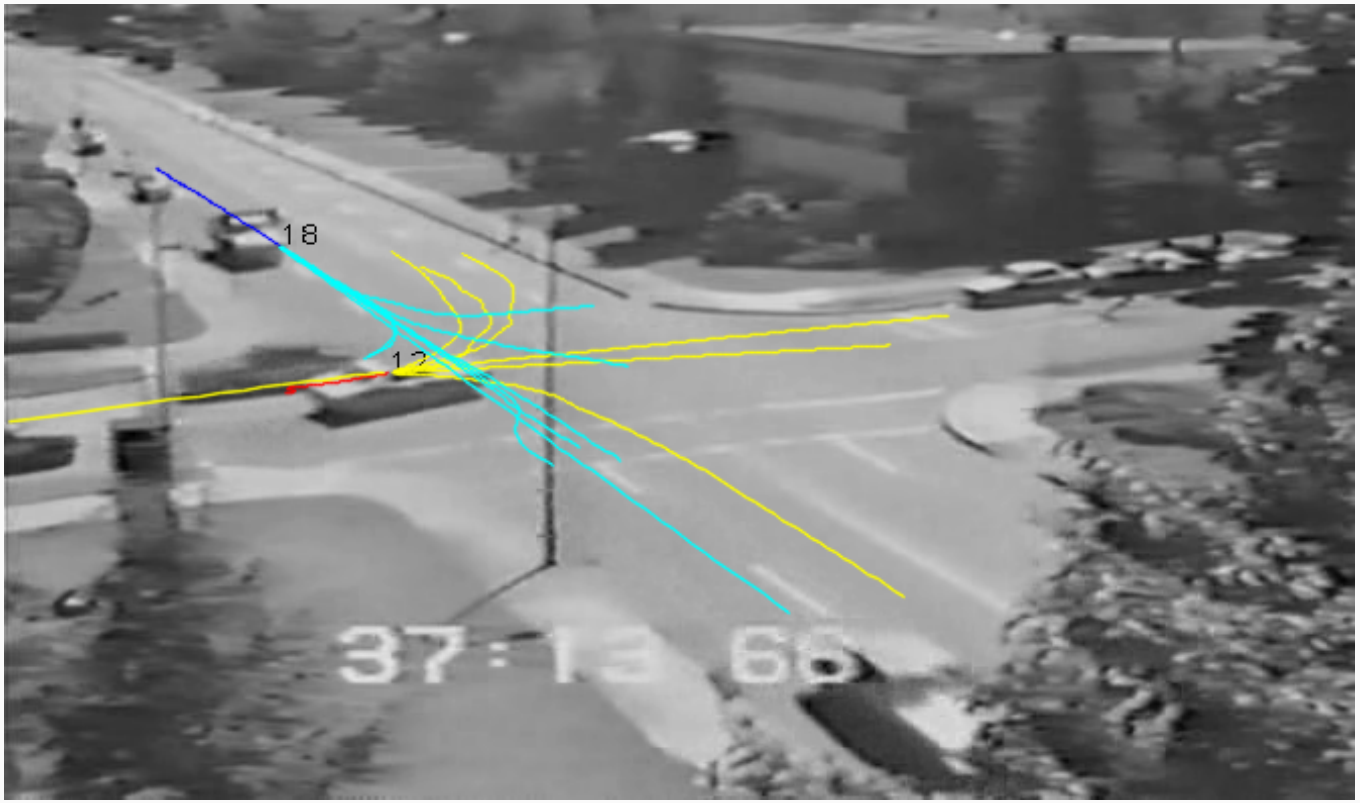
$$WAPE = \frac{\sum_{i=1}^{n} |t_i - r_i|}{\sum_{i=1}^{n} t_i}$$

- t_i vrai comptage pour le mouvement i
- r_i comptage automatique pour le mouvement i
- n nombre de mouvements

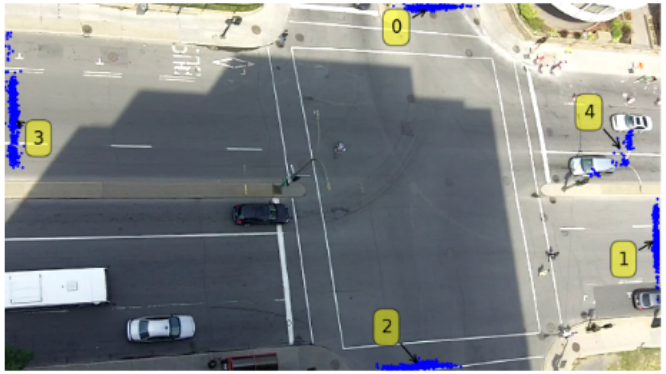
Application à la sécurité

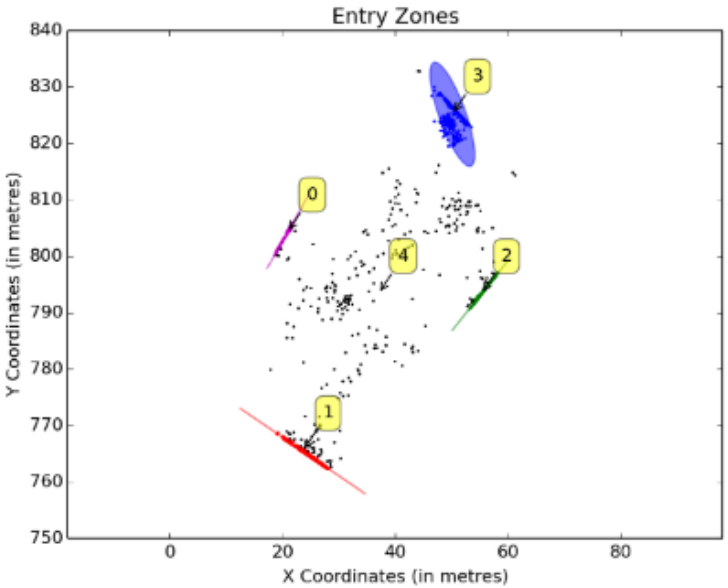


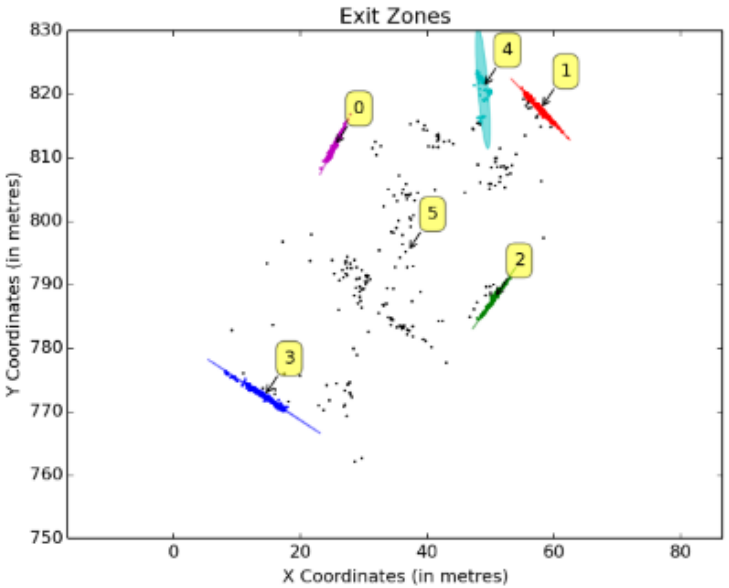
Application à la sécurité



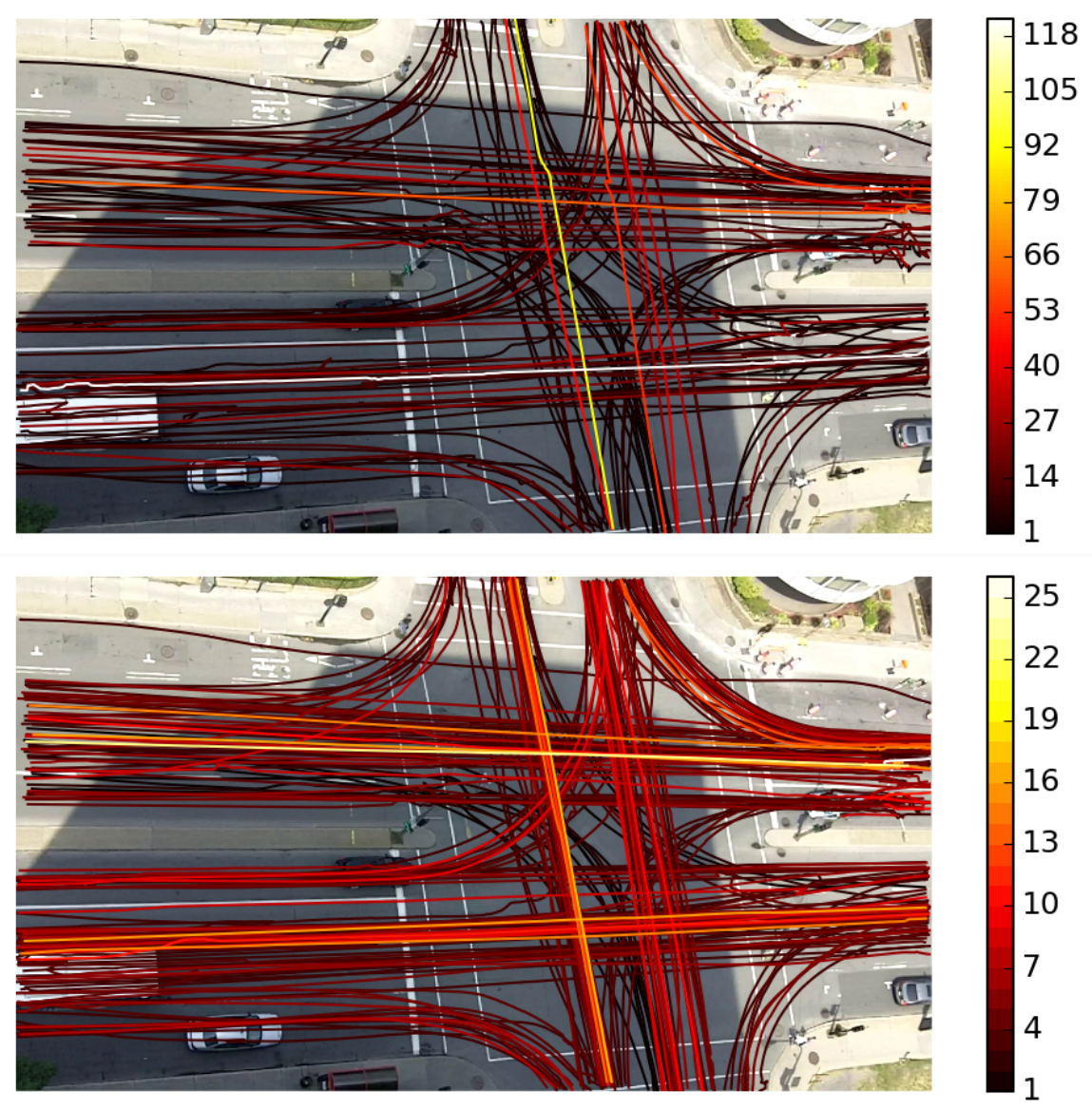
Apprentissage en deux étapes (M. Gomaa Mohamed)



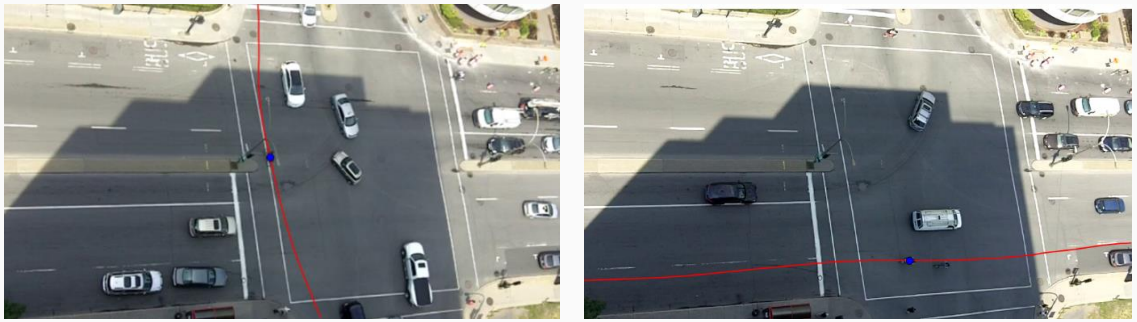




Apprentissage en deux étapes (M. Gomaa Mohamed)

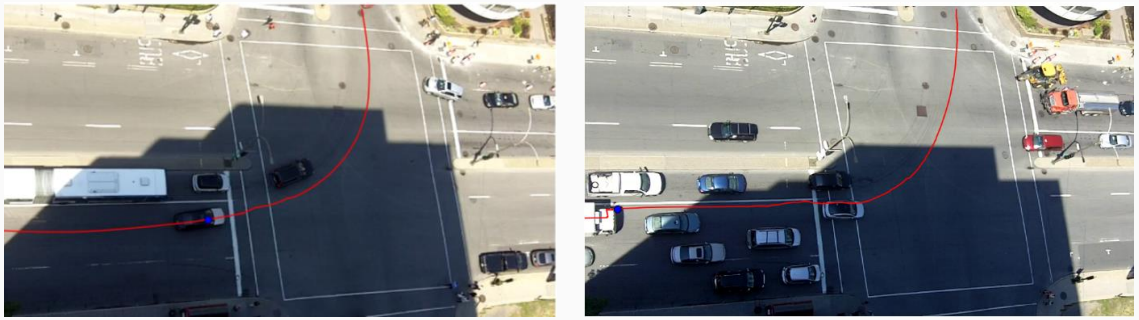


Application à la détection d'anomalies (M. Gomaa Mohamed)



a) Misclassification as vehicle (cyclist)

b) Misclassification as vehicle (pedestrian jogging)



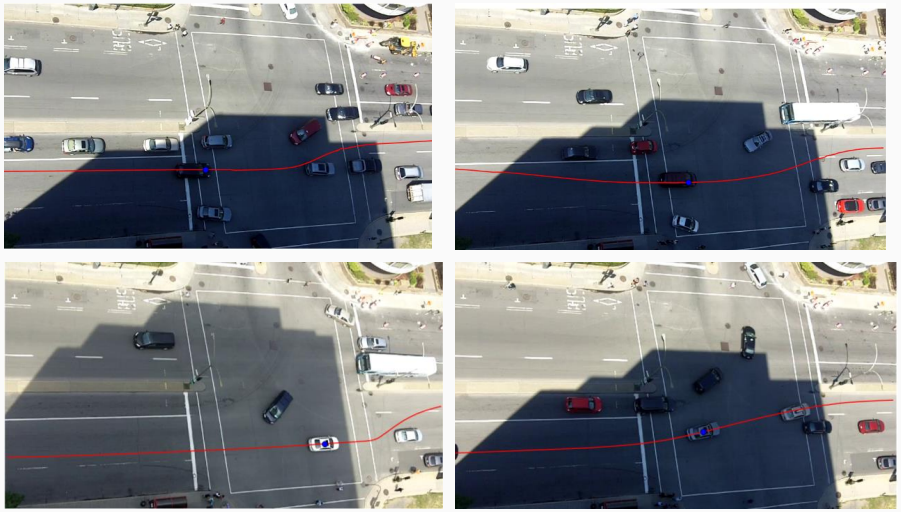
b) Abnormal Left Turn

Figure 4.16: Samples of detected spatial anomalies at the Guy intersection

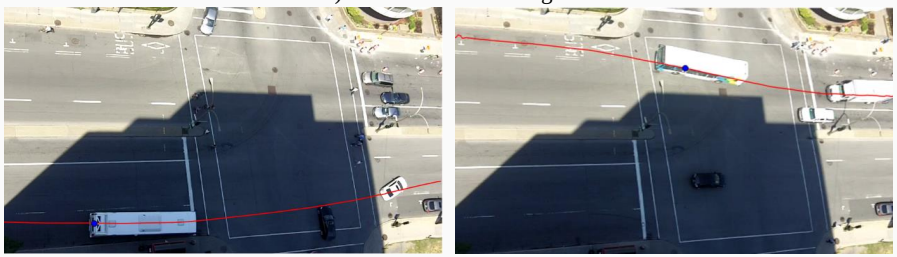
Application à la détection d'anomalies (M. Gomaa Mohamed)

83

c) Abnormal Right Turn



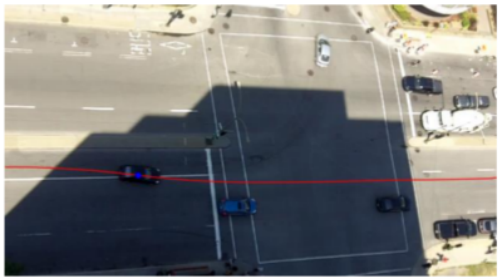
d) Abnormal lane Changes



e) Normal movements that rarely happens (bus lane change to its stop station or to its exclusive lane)

Figure 4.16: Samples of detected spatial anomalies at the Guy intersection (cont'd)

Application à la détection d'anomalies (M. Gomaa Mohamed)



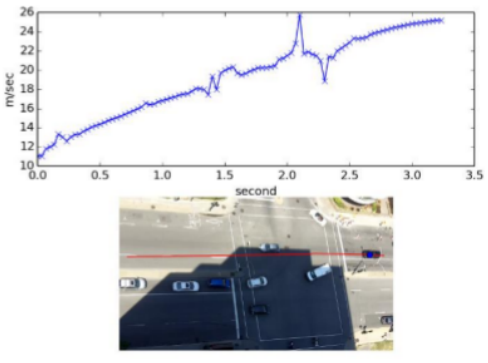
f) Lane change over solid lane markings with no surrounding vehicle

g) Side-swipe interactions (one vehicle changes lane over a solid lane marking while another vehicle adapts its path at the same time)

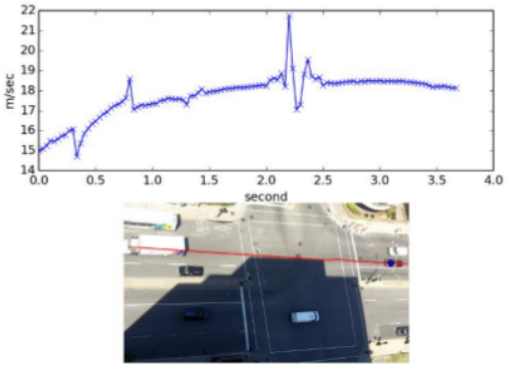
h) Abnormal Overtaking

i) Connection error

Figure 4.16: Samples of detected spatial anomalies at the Guy intersection (cont'd)



a) Excessive speed (maximum 85km/hr)



b) Excessive speed (maximum 65km/hr)

Figure 4.17: Samples of detected excessive speed profile anomalies at the Guy intersection

Application au regroupement des profils de TTC

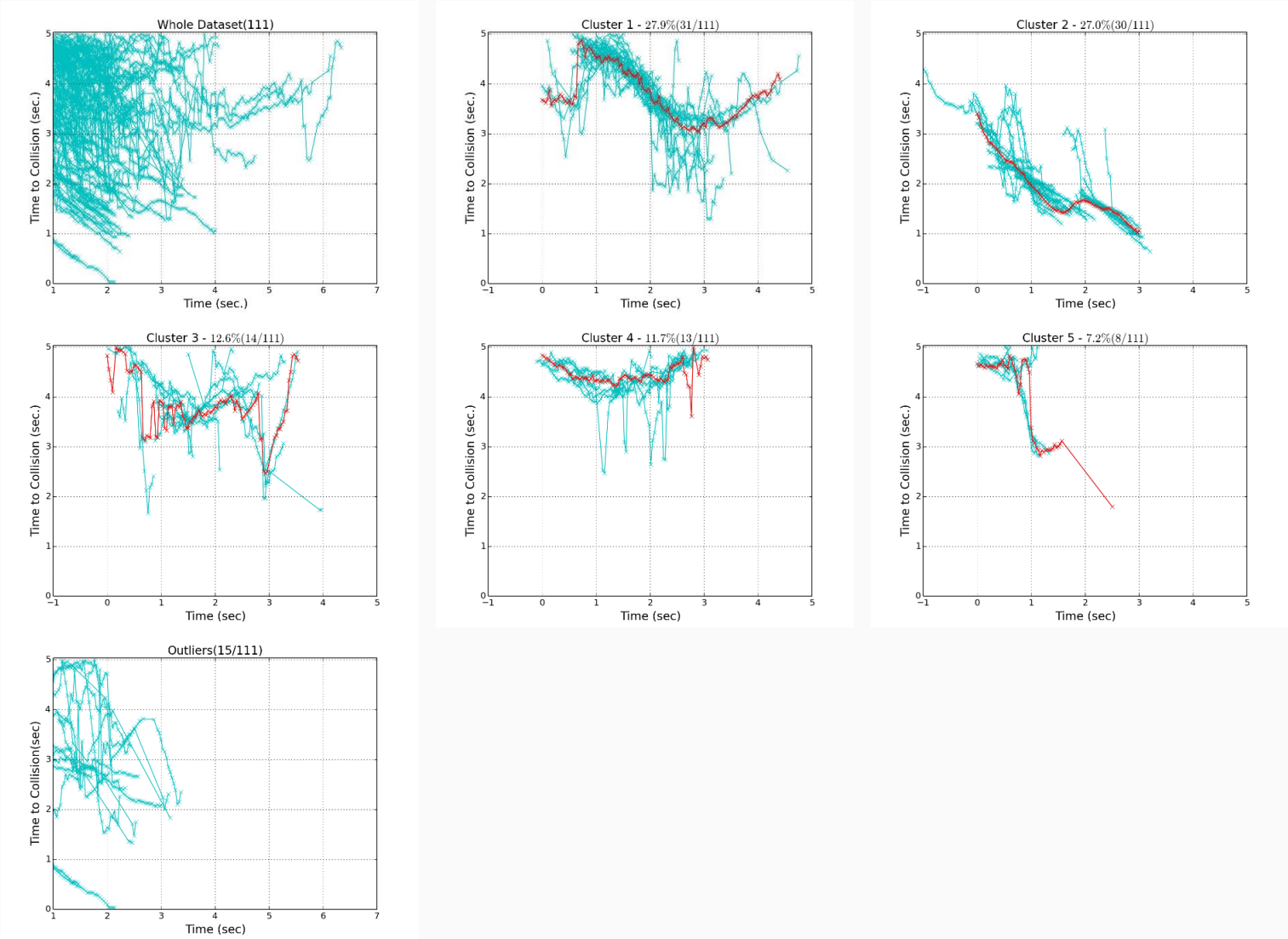
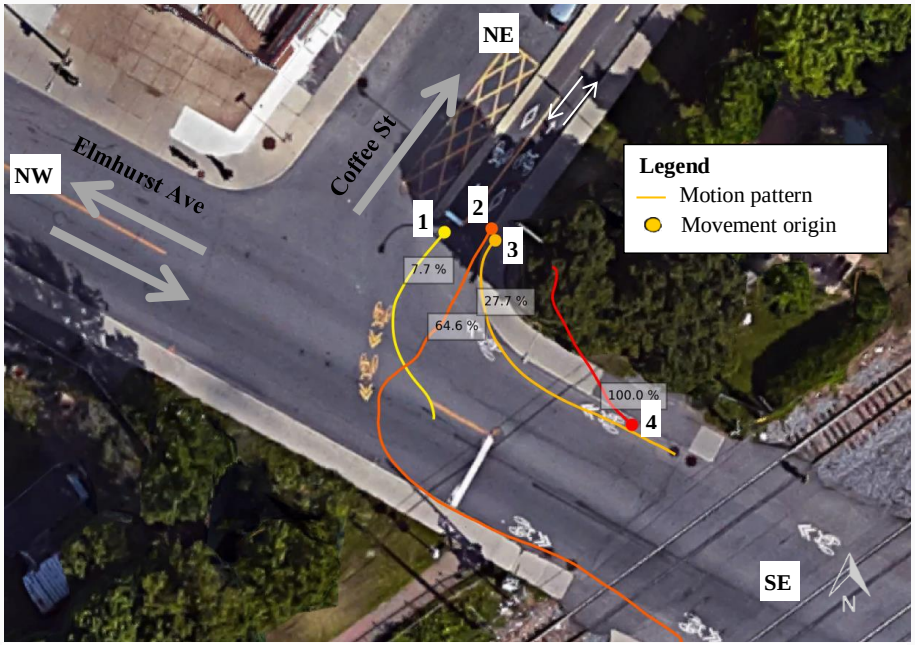
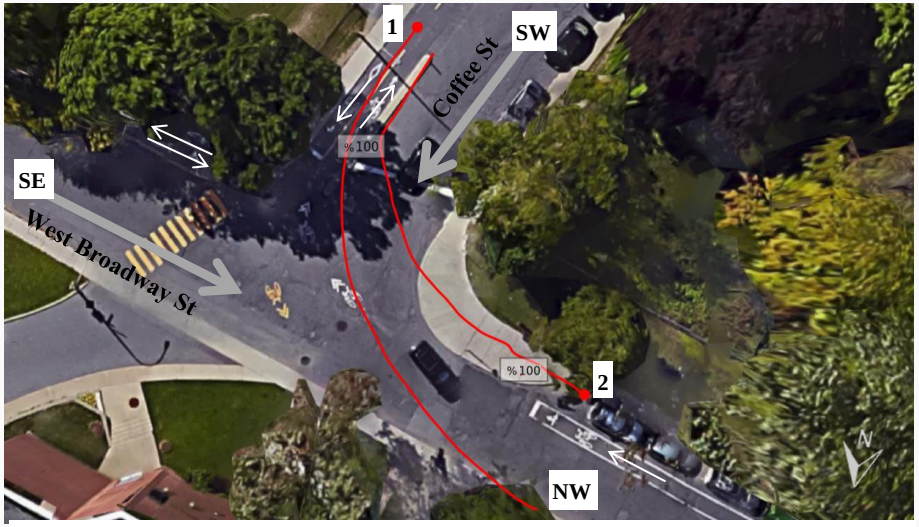


Figure 6.13: Clusters of the TTC profiles from the Guy Intersection (whole dataset, 5 clusters, and outliers)

Application à l'analyse de comportements (M. Nabavi Niaki)



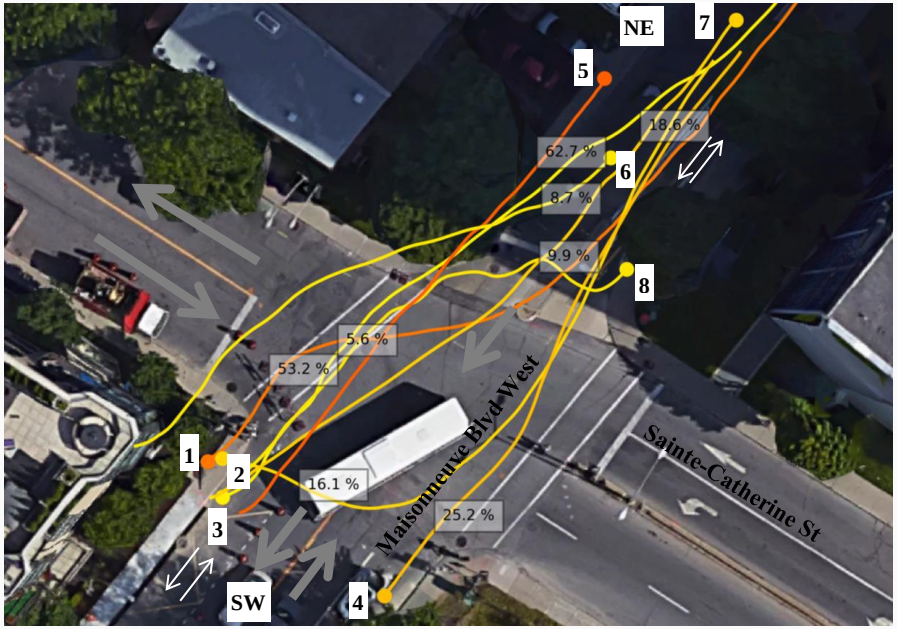
a. Coffee St and Elmhurst Ave (discontinuity)



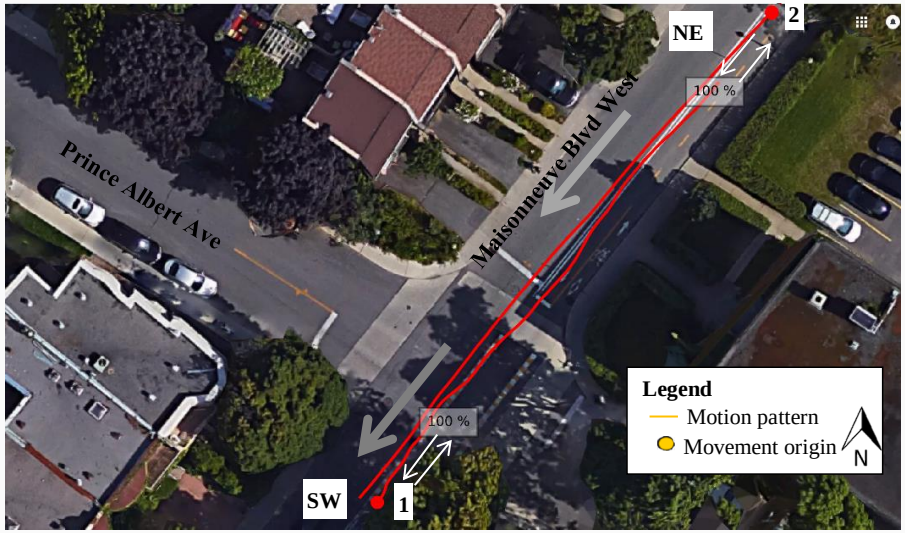
b. Coffee St and West Broadway St (control)

Figure 2 Cyclist motion patterns (represented by their prototype trajectories) for the change in cycling facility type discontinuity

Application à l'analyse de comportements (M. Nabavi Niaki)



a. Maisonneuve Blvd and Sainte-Catherine St (discontinuity)



b. Maisonneuve Blvd west and Prince Albert Ave (control)

Figure 3 Cyclist motion patterns (represented by their prototype trajectories) for the change in cycling facility side discontinuity

Plan de la présentation

Introduction

Méthodes d’apprentissage des trajectoires

Résultats expérimentaux

Conclusion

Conclusion

- Les données de trajectoires sont de plus en plus faciles à collecter
- Il existe des méthodes puissantes et flexibles pour les comparer (mesurer leur similarité)
- Les applications sont généralement non-supervisées
  - difficile de valider les méthodes en général
  - besoin d'interpréter les résultats

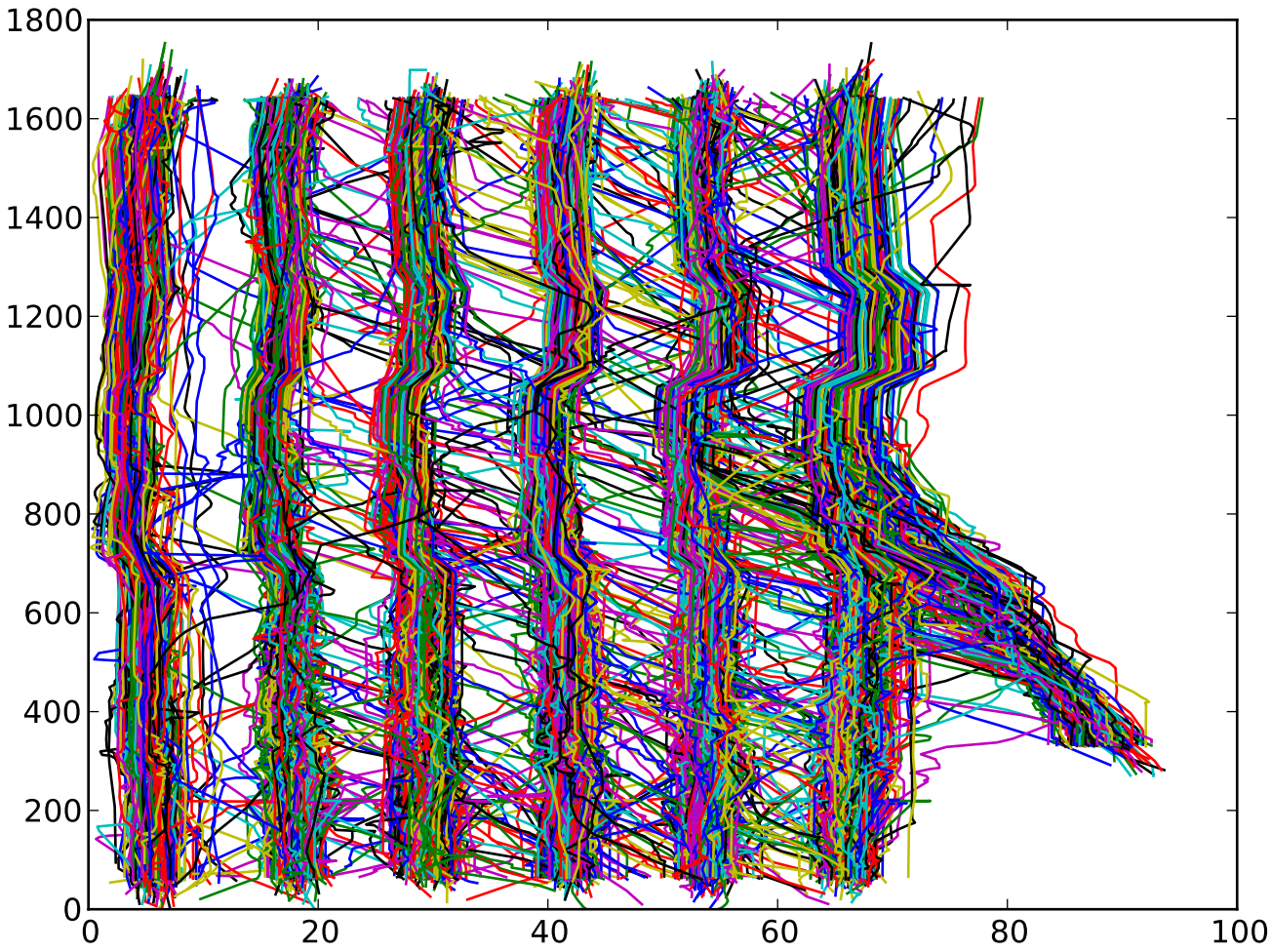
Perspectives

- Application au suivi dans des données vidéos: projet IVADO avec G.-A. Bilodeau, G. Pesant et D. Aloise
  - comparaison des méthodes
- Adaptation aux méga-données: travail de P. St-Aubin
- Validation élusive des mesures de sécurité

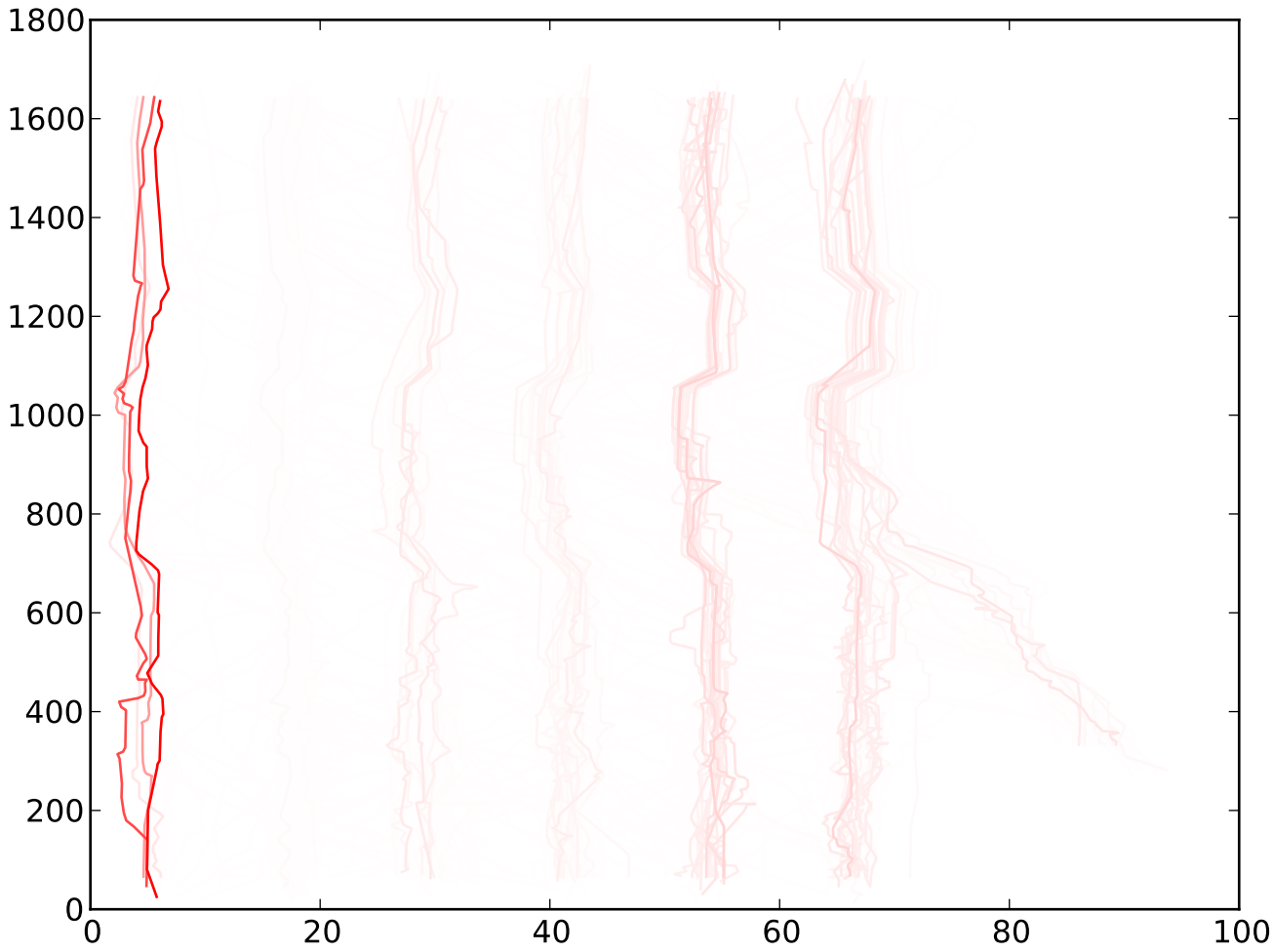
Questions?

http://nicolas.saunier.confins.net
nicolas.saunier@polymtl.ca

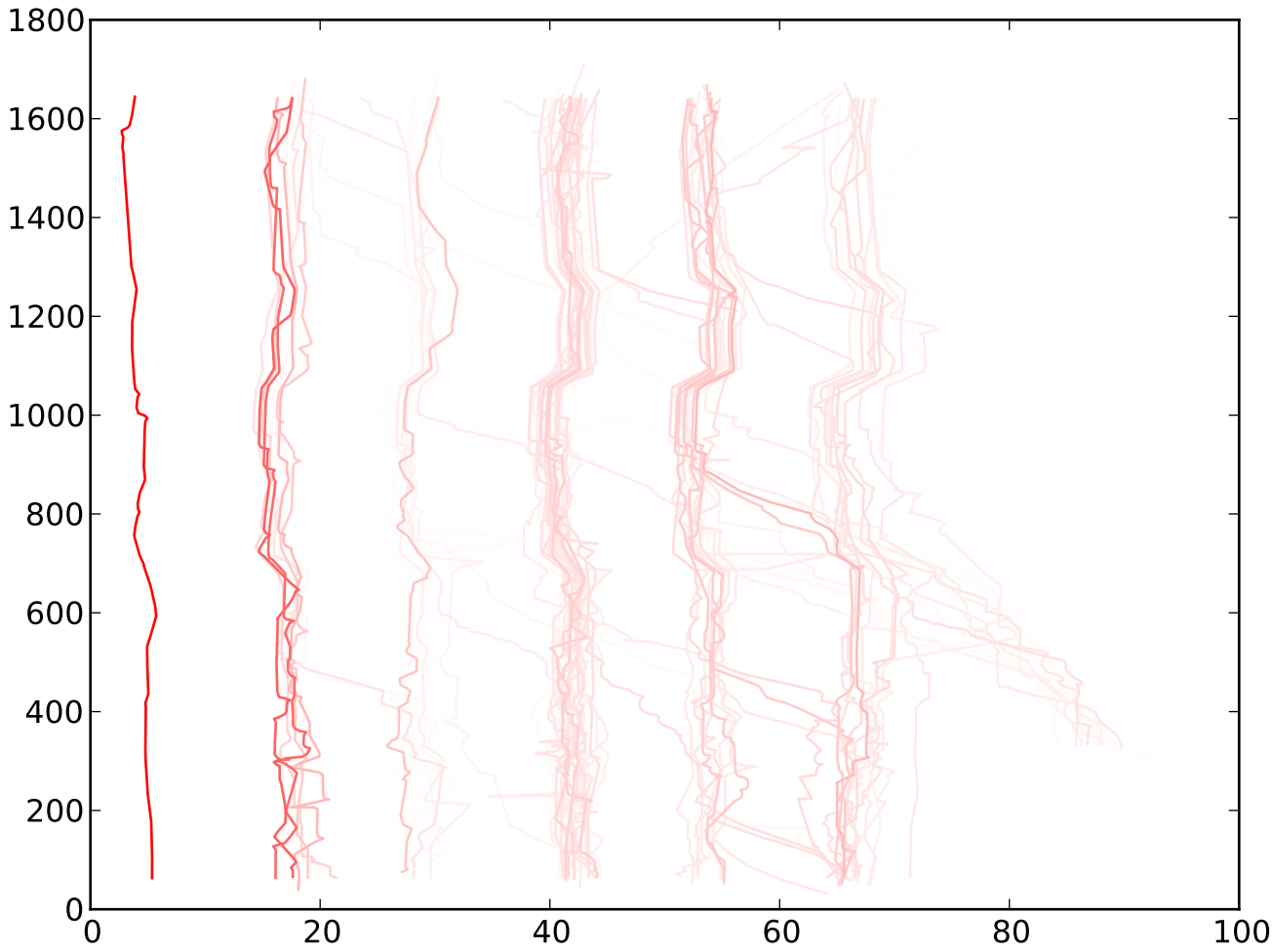
Clustering Examples: NGSIM Dataset (2052)



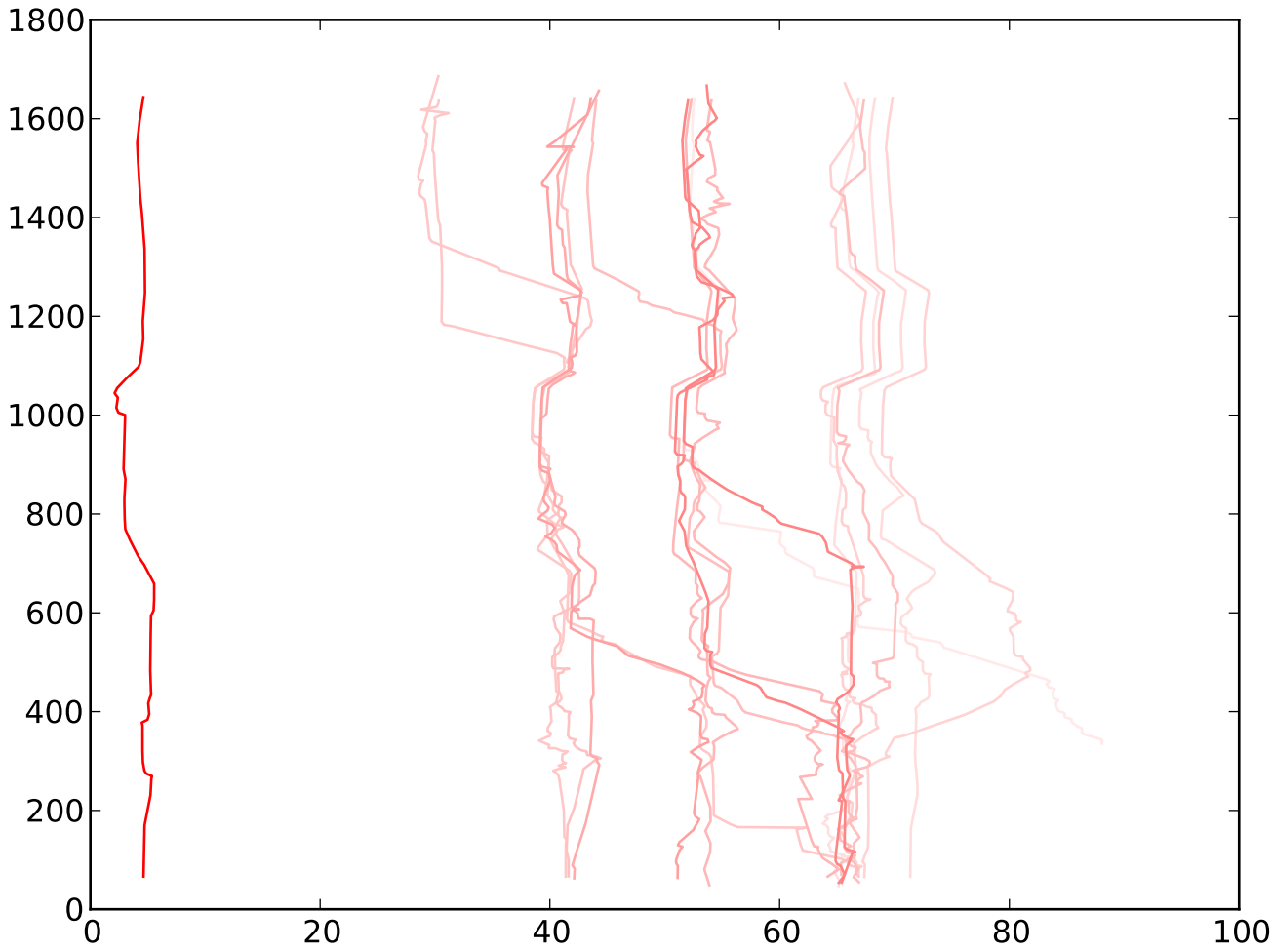
Clustering Examples: NGSIM Dataset (333)



Clustering Examples: NGSIM Dataset (96)

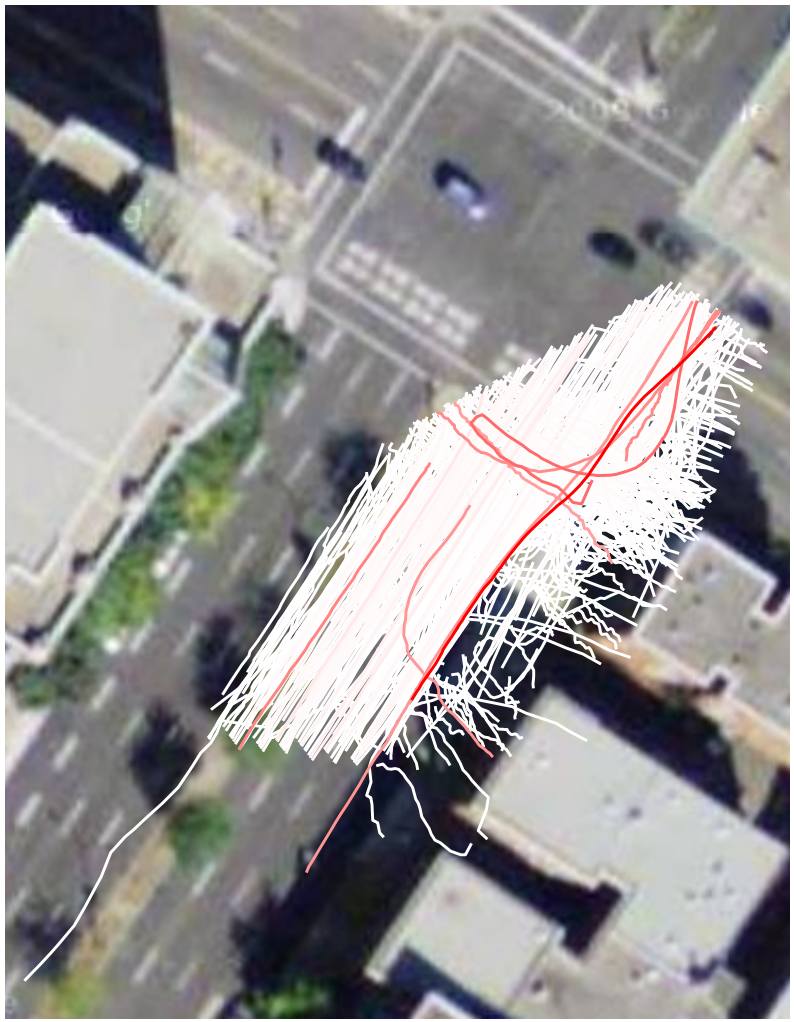


Clustering Examples: NGSIM Dataset (19)

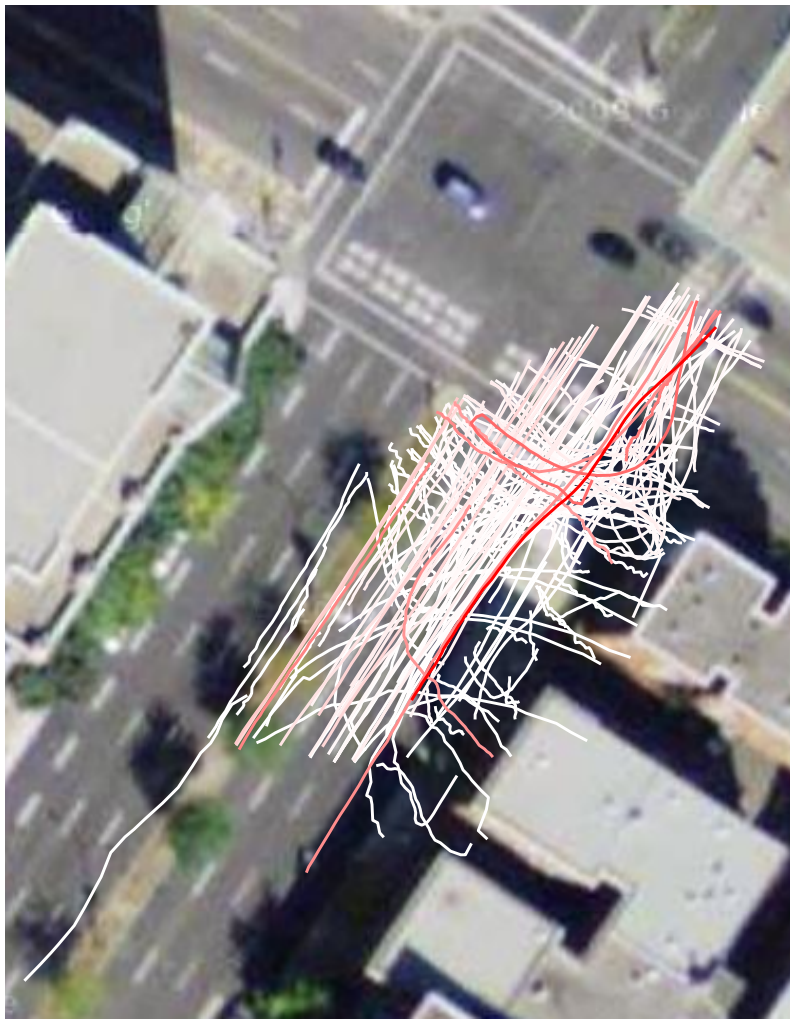


Clustering Examples: Montréal Intersection (6777)

Clustering Examples: Montréal Intersection (587)

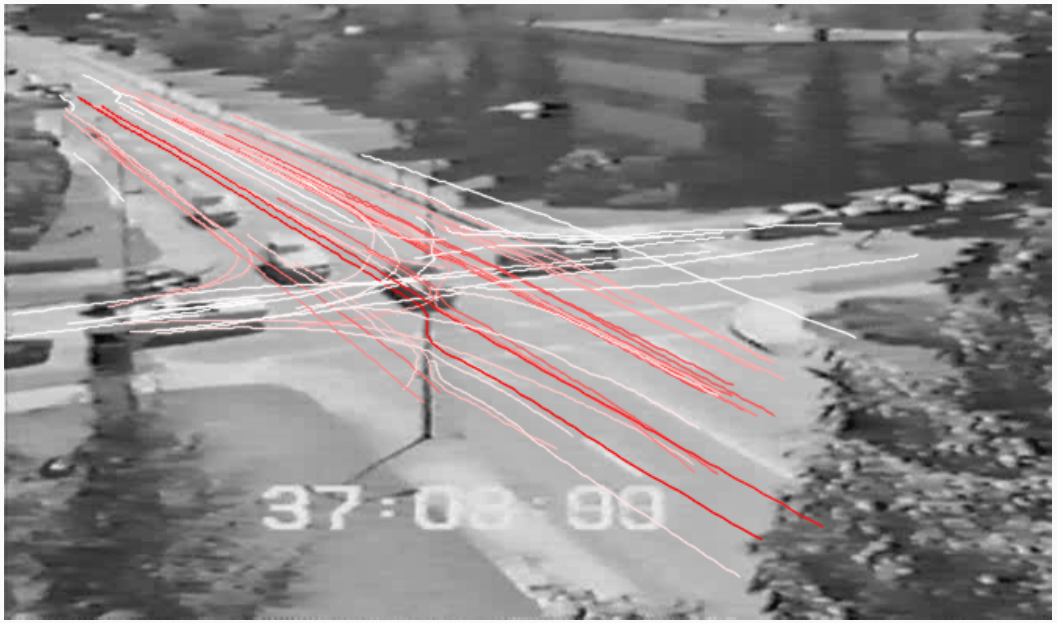


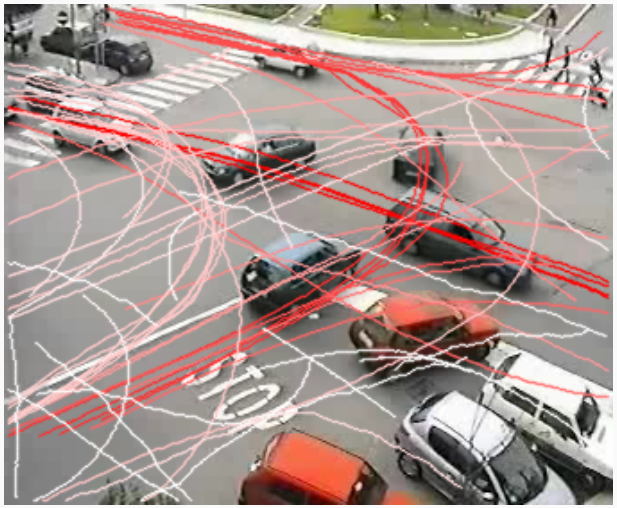
Clustering Examples: Montréal Intersection (168)



Clustering Examples: Montréal Intersection (9)

Motion Pattern Learning





Traffic Conflict Dataset, Vancouver	Reggio Calabria, Italy
58 prototype trajectories (2941 trajectories)	58 prototype trajectories (138009 trajectoires)