Data-driven characterization of pedestrian traffic

Marija Nikolić, Michel Bierlaire

June 14, 2016
Outline

1. Introduction

2. Related research

3. Methodology
   - Discretization framework
   - Definitions of the indicators

4. Empirical analysis

5. Conclusion and future work
Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
4. Empirical analysis
5. Conclusion and future work
Motivation
Background

Importance

- Understanding, reproducing and forecasting phenomena that characterize pedestrian traffic is necessary in order to provide services related to pedestrian safety and convenience.

Vehicular traffic

- Well-established theory
- Regulated and separated by directions
Background

Pedestrian traffic

• Multidirectional, without strict rules for pedestrian to follow
• Pedestrians can occupy any part of the walkable area

Indicators

• Density $k \ (\text{ped/m}^2)$, speed $v \ (m/s)$ and flow $q \ (\text{ped/ms})$
• Used to observe and to model the flows of pedestrians
• Consistent and unified approach to the definitions of the indicators is missing
Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
4. Empirical analysis
5. Conclusion and future work
Grid-based (GB) method

\[ k(A) = \frac{N}{|A|} \]
\[ v(A) = \frac{\sum v_i}{N} \]
\[ q(A) = k(A)v(A) \]

\[ v_i = \sqrt{\left(\frac{\Delta x_i}{\Delta t}\right)^2 + \left(\frac{\Delta y_i}{\Delta t}\right)^2} \]
\[ \Delta x_i = x_i(t + \Delta t) - x_i(t), \quad \Delta y_i = y_i(t + \Delta t) - y_i(t) \]

[Duives et al., 2015]
Range-based (RB) method

\[
k(A_r) = \frac{N}{|A_r|} \\
v(A_r) = \sum_{i}^{N} v_i \\
q(A_r) = k(A_r)v(A_r)
\]

\[
v_i = \sqrt{\left(\frac{\Delta x_i}{\Delta t}\right)^2 + \left(\frac{\Delta y_i}{\Delta t}\right)^2}
\]

\[
\Delta x_i = x_i(t + \Delta t) - x_i(t), \quad \Delta y_i = y_i(t + \Delta t) - y_i(t)
\]

[Duives et al., 2015]
Exponentially Weighted (EW) method

\[ k(x, y, t) = \sum_{i=1}^{f} f \left( \left( \begin{array}{c} x_i(t) \\ y_i(t) \end{array} \right) - \left( \begin{array}{c} x \\ y \end{array} \right) \right) \]

\[ \bar{v}(x, y, t) = \frac{\sum_{i=1}^{\bar{v}_i(x,y,t)} f \left( \left( \begin{array}{c} x_i(t) \\ y_i(t) \end{array} \right) - \left( \begin{array}{c} x \\ y \end{array} \right) \right)}{\sum_{i=1}^{f} f \left( \left( \begin{array}{c} x_i(t) \\ y_i(t) \end{array} \right) - \left( \begin{array}{c} x \\ y \end{array} \right) \right)} \]

\[ \bar{q}(x, y, t) = k(x, y, t) \bar{v}(x, y, t) \]

\[ f \left( \left( \begin{array}{c} x_i(t) \\ y_i(t) \end{array} \right) - \left( \begin{array}{c} x \\ y \end{array} \right) \right) = \frac{1}{\pi R^2} \exp \left( - \frac{\left\| \left( \begin{array}{c} x_i(t) \\ y_i(t) \end{array} \right) - \left( \begin{array}{c} x \\ y \end{array} \right) \right\|^2}{R^2} \right) \]

[Helbing et al., 2007], [Duives et al., 2015]
XY-T method

\[ k(V) = \frac{\sum_{i} t_i}{XYT} \]

\[ \bar{q}(V) = \left( \begin{array}{c} q_x(V) \\ q_y(V) \end{array} \right) = \left( \begin{array}{c} \frac{\sum_{i=1}^{N} x_i}{XYT} \\ \frac{\sum_{i=1}^{N} y_i}{XYT} \end{array} \right) \]

\[ \bar{v}(V) = \left( \begin{array}{c} v_x(V) \\ v_y(V) \end{array} \right) = \left( \begin{array}{c} \frac{q_x(V)}{k(V)} \\ \frac{q_y(V)}{k(V)} \end{array} \right) \]

[van Wageningen-Kessels et al., 2014], [Saberi and Mahmassani, 2014]
Voronoi-based (VB) method

A personal region $A_i$ is assigned to each pedestrian $i$: each point $p$ in the personal region of pedestrian $i$ is closer to $i$ than to any other, with respect of $d_E$

$$A_i = \{ p | d_E(p, p_i) \leq d_E(p, p_j), \forall j \}$$

$$k(A_i) = \frac{1}{|A_i|}$$

$$v(A_i) = \sqrt{\left( \frac{\Delta x_i}{\Delta t} \right)^2 + \left( \frac{\Delta y_i}{\Delta t} \right)^2}$$

$$q(A_i) = k(A_i) v(A_i)$$

$$\Delta x_i = x_i(t + \Delta t) - x_i(t), \quad \Delta y_i = y_i(t + \Delta t) - y_i(t)$$

[Steffen and Seyfried, 2010], [Duives et al., 2015]
Arbitrary discretization

Sensitivity of results

The results might be very sensitive to minor changes

Unrealistic results

Velocity and flow vectors may cancel out when 2 equally sized streams of pedestrians walk with the same speed but in the opposite directions
How to define the discretization...

...independent of arbitrary chosen values?
One size and shape fits all?

It is all about adjustments...
Outline

1 Introduction

2 Related research

3 Methodology
   - Discretization framework
   - Definitions of the indicators

4 Empirical analysis

5 Conclusion and future work
Data-driven approach

Keep calm and let data speak!
Outline

1. Introduction

2. Related research

3. Methodology
   - Discretization framework
   - Definitions of the indicators

4. Empirical analysis

5. Conclusion and future work
Data-driven discretization framework

Pedestrian trajectories

$$\Gamma_i : \{ p_i(t) | p_i(t) = (x_i(t), y_i(t), t) \}$$

3D Voronoi diagrams associated with trajectories

Each trajectory $\Gamma_i$ is associated with a 3D Voronoi ‘tube’ $V_i$

$$V_i = \{ p | \min \{ d_*(p, p_i) | p_i \in \Gamma_i \} \leq \min \{ d_*(p, p_j) | p_j \in \Gamma_j \}, \forall j \}$$

$d_*(p, p_i)$ - spatio-temporal assignment rule
Data-driven discretization framework

Sample of points

\[ \Gamma_i : \{ p_{is} \mid p_{is} = (x_{is}, y_{is}, t_s) \} \], \quad t_s = [t_0, t_1, \ldots, t_f] \]

3D Voronoi diagrams associated with the points

Sequences of 3D Voronoi cells \( V_{is} \) are assigned to the sequence of points for each pedestrian

\[ V_i = \{ V_{is} \mid V_{is} = \{ p \mid d_*(p, p_{is}) \leq d_*(p, p_{js}) \}, \quad \forall j \} \]

\( d_*(p, p_i) \) - spatio-temporal assignment rule
Spatio-temporal assignment rules

Naive assignment rule (N-3D Voro)

\[ d_N(p, p_i) = \begin{cases} 
\sqrt{(p - p_i)^T(p - p_i)}, & \Delta t = 0 \\
\infty, & otherwise
\end{cases} \]

Time-Transform assignment rules (TT\{1,2,3\}-3D Voro)

\[ d_{TT_1}(p, p_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + \alpha^2(t - t_i)^2} \]
\[ d_{TT_2}(p, p_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + \alpha_i(t_i)|(t - t_i)|} \]
\[ d_{TT_3}(p, p_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + \alpha_i^2(t_i)(t - t_i)^2} \]

\(\alpha\) and \(\alpha_i\) - conversion constants expressed in meters per second
Spatio-temporal assignment rules

Predictive assignment rule (P-3D Voronoi)

\[ d_P(p, p_i) = \begin{cases} \sqrt{(x_i(t) - x)^2 + (y_i(t) - y)^2}, & t - t_i \geq 0 \\ \infty, & \text{otherwise} \end{cases} \]

The anticipated position of pedestrian \( i \) at time \( t \):
\[ x_i(t) = x_i(t_i) + (t - t_i)v_i^x(t_i), \ y_i(t) = y_i(t_i) + (t - t_i)v_i^y(t_i) \]

The speed of pedestrian \( i \) at \( t_i \) in \( x \) and \( y \) directions: \( v_i^x(t_i), v_i^y(t_i) \)

Mahalanobis assignment rule (M-3D Voronoi)

\[ d_M(p, p_i) = \sqrt{(p - p_i)^T M_i (p - p_i)} \]

\( M_i \) - symmetric, positive-definite matrix that defines how distances are measured from the perspective of pedestrian \( i \)
Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
4. Empirical analysis
5. Conclusion and future work
Voronoi-based traffic indicators

The set of all points in $V_i$ corresponding to a specific time $t$

$$V_i(t) = \{(x(t), y(t), t) \in V_i\} \sim [m^2]$$

Density indicator

$$k(x, y, t) = \frac{1}{|V_i(t)|}, \text{ for } x, y \in V_i(t)$$
Voronoi-based traffic indicators

The set of all points in $V_i$ corresponding to a given location $x$ and $y$

$$V_i(x) = \{(x, y, t) \in V_i\} \sim [ms]$$

$$V_i(y) = \{(x, y, t) \in V_i\} \sim [ms]$$

Flow indicator

$$\bar{q}(x, y, t) = \begin{pmatrix} q^x(x, y, t) \\ q^y(x, y, t) \end{pmatrix} = \begin{pmatrix} \frac{1}{|V_i(x)|} \\ \frac{1}{|V_i(y)|} \end{pmatrix}$$

Velocity indicator

$$\vec{v}(x, y, t) = \begin{pmatrix} \frac{q^x(x, y, t)}{k(x, y, t)} \\ \frac{q^y(x, y, t)}{k(x, y, t)} \end{pmatrix} = \begin{pmatrix} \frac{|V_i(t)|}{|V_i(x)|} \\ \frac{|V_i(t)|}{|V_i(y)|} \end{pmatrix}$$
Outline

1 Introduction

2 Related research

3 Methodology
   - Discretization framework
   - Definitions of the indicators

4 Empirical analysis

5 Conclusion and future work
Performance of the approach

Synthetic data - unidirectional flow

NOMAD simulation tool [Campanella, 2010]

**Scenario I**: low congestion, homogenous population

**Scenario II**: high congestion, heterogeneous population

Indicators

Robustness w.r.t. the aggregation

Robustness w.r.t. the sampling frequency
Characterization based on trajectories

Robustness with respect to the aggregation

- Ability of tolerating perturbations in data

- 100 sets of pedestrian trajectories synthesized per scenario
- Indicators $k$, $\nu$, and $q$ calculated for each set
Robustness with respect to the aggregation

Standard deviation (1000 points) - Scenario I
Robustness with respect to the aggregation

Standard deviation (1000 points) - Scenario II
Characterization based on sampled data

Robustness with respect to the sampling frequency

- Ability of tolerating missing data

- Synthetic trajectories sampled using different sampling frequencies

- Indicators calculated via
  1. 3D Voro applied to the interpolated trajectories
  2. 3D Voro applied directly to the samples

- Comparison of the indicators at 1000 randomly selected points to the corresponding values obtained utilizing true trajectories
Robustness w.r.t the sampling frequency - Scenario I

High sampling frequency: $3.33\text{s}^{-1}$

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>90% quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IT</td>
<td>SoP</td>
<td>IT</td>
<td>SoP</td>
</tr>
<tr>
<td>N-3DVoro</td>
<td>1.17E-02</td>
<td>/</td>
<td>0</td>
<td>/</td>
</tr>
<tr>
<td>TT$_1$-3DVoro</td>
<td>2.70E-03</td>
<td>6.70E-03</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TT$_2$-3DVoro</td>
<td>5.80E-03</td>
<td>3.50E-02</td>
<td>0</td>
<td>2.80E-03</td>
</tr>
<tr>
<td>TT$_3$-3DVoro</td>
<td>5.40E-03</td>
<td>4.34E-02</td>
<td>0</td>
<td>8.00E-03</td>
</tr>
<tr>
<td>P-3DVoro</td>
<td>8.20E-03</td>
<td>5.36E-02</td>
<td>0</td>
<td>6.10E-03</td>
</tr>
<tr>
<td>M-3DVoro</td>
<td>4.50E-03</td>
<td>5.65E-02</td>
<td>0</td>
<td>6.80E-03</td>
</tr>
</tbody>
</table>

Low sampling frequency: $0.5\text{s}^{-1}$

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>90% quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IT</td>
<td>SoP</td>
<td>IT</td>
<td>SoP</td>
</tr>
<tr>
<td>N-3DVoro</td>
<td>1.64E-01</td>
<td>/</td>
<td>0</td>
<td>/</td>
</tr>
<tr>
<td>TT$_1$-3DVoro</td>
<td>2.54E-01</td>
<td>1.27E-01</td>
<td>1.35E-02</td>
<td>9.00E-03</td>
</tr>
<tr>
<td>TT$_2$-3DVoro</td>
<td>1.64E-01</td>
<td>1.22E-01</td>
<td>1.44E-02</td>
<td>1.06E-02</td>
</tr>
<tr>
<td>TT$_3$-3DVoro</td>
<td>1.89E-01</td>
<td>1.24E-01</td>
<td>1.84E-02</td>
<td>1.09E-02</td>
</tr>
<tr>
<td>P-3DVoro</td>
<td>3.19E-01</td>
<td>1.21E-01</td>
<td>3.26E-02</td>
<td>6.20E-03</td>
</tr>
<tr>
<td>M-3DVoro</td>
<td>1.97E-01</td>
<td>1.24E-01</td>
<td>3.48E-02</td>
<td>9.90E-03</td>
</tr>
</tbody>
</table>
Robustness w.r.t the sampling frequency - Scenario II

High sampling frequency: 3.33s$^{-1}$

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>90% quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IT SoP</td>
<td>IT SoP</td>
<td>IT SoP</td>
<td>IT SoP</td>
</tr>
<tr>
<td>N-3DVoro</td>
<td>1.43E-02</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>TT$_1$-3DVoro</td>
<td>8.00E-03</td>
<td>4.55E-02</td>
<td>0 0</td>
<td>8.00E-04 1.75E-02</td>
</tr>
<tr>
<td>TT$_2$-3DVoro</td>
<td>1.49E-02</td>
<td>1.07E-01</td>
<td>0 0</td>
<td>3.20E-03 5.72E-02</td>
</tr>
<tr>
<td>TT$_3$-3DVoro</td>
<td>1.24E-02</td>
<td>1.60E-01</td>
<td>0 0</td>
<td>3.50E-03 9.62E-02</td>
</tr>
<tr>
<td>P-3DVoro</td>
<td>2.10E-02</td>
<td>1.66E-01</td>
<td>0 0</td>
<td>4.20E-03 1.16E-01</td>
</tr>
<tr>
<td>M-3DVoro</td>
<td>1.31E-02</td>
<td>2.40E-01</td>
<td>0 0</td>
<td>2.50E-03 1.75E-01</td>
</tr>
</tbody>
</table>

Low sampling frequency: 0.5s$^{-1}$

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>90% quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IT SoP</td>
<td>IT SoP</td>
<td>IT SoP</td>
<td>IT SoP</td>
</tr>
<tr>
<td>N-3DVoro</td>
<td>4.02E-01</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>TT$_1$-3DVoro</td>
<td>4.06E-01</td>
<td>2.90E-01</td>
<td>3.10E-01 2.48E-02</td>
<td>2.64E-01 1.65E-01</td>
</tr>
<tr>
<td>TT$_2$-3DVoro</td>
<td>3.92E-01</td>
<td>4.58E-01</td>
<td>2.85E-01 2.34E-01</td>
<td>2.48E-01 2.34E-01</td>
</tr>
<tr>
<td>TT$_3$-3DVoro</td>
<td>4.41E-01</td>
<td>5.07E-01</td>
<td>2.89E-01 5.89E-02</td>
<td>2.37E-01 3.06E-01</td>
</tr>
<tr>
<td>P-3DVoro</td>
<td>4.31E-01</td>
<td>3.71E-01</td>
<td>1.40E-03 0</td>
<td>2.58E-01 1.80E-01</td>
</tr>
<tr>
<td>M-3DVoro</td>
<td>4.34E-01</td>
<td>5.01E-01</td>
<td>3.16E-01 1.36E-01</td>
<td>2.75E-01 3.52E-01</td>
</tr>
</tbody>
</table>
Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
4. Empirical analysis
5. Conclusion and future work
Conclusion and future work

Conclusion

• A novel approach to pedestrian traffic characterization: data-driven discretization via 3D Voronoi diagrams
• Superior to existing methods w.r.t. robustness to the aggregation
• Robustness to the sampling frequency
  – $TT_1$-3DVoro: high sampling frequency or higher congestion
  – P-3DVoro: low sampling frequency and lighter traffic conditions

Future work

• Analysis of the performance for different scenarios
• Weighted assignment rules
Thank you

9th TRIENNIAL SYMPOSIUM ON TRANSPORTATION ANALYSIS (TRISTAN IX), Aruba: 
Data-driven characterization of pedestrian traffic
Marija Nikolić, Michel Bierlaire

Help by S. S. Azadeh and F. Hänseler is appreciated.

- marija.nikolic@epfl.ch
Robustness with respect to the aggregation

Spread/point - Scenario I
Robustness with respect to the aggregation

Spread/point - Scenario II
Robustness with respect to the aggregation

Smoothness/point - Scenario I
Robustness with respect to the aggregation

Smoothness/point - Scenario II
Mahalanobis distance

**Directions of interest**

\[ p_is = (x_is, y_is, t_s), \quad v_i(t_s) = \frac{1}{t(s+1) - t_s} \begin{pmatrix} x_i(s+1) - x_is \\ y_i(s+1) - y_is \\ 1 \end{pmatrix} \]

\[ d^1(t_s) = \frac{v_i(t_s)}{||v_i(t_s)||}, \quad ||d^1(t_s)|| = 1 \]

\[ d^2(t_s) = \begin{pmatrix} d^1_x(t_s) \\ d^2_y(t_s) \\ 0 \end{pmatrix}, \quad d^1(t_s)^T d^2(t_s) = 0, \quad ||d^2(t_s)|| = 1 \]

\[ d^3(t_s) = \begin{pmatrix} 0 \\ 0 \\ t(s+1) - t_s \end{pmatrix}, \quad ||d^3(t_s)|| = t(s+1) - t_s \]
Mahalanobis distance

Change of coordinates

\[ S_1(t_s, \delta) = p_is + (t_{s+1} - t_s)v_i(t_s) + \delta d^1(t_s) \]
\[ S_2(t_s, \delta) = p_is - (t_{s+1} - t_s)v_i(t_s) - \delta d^1(t_s) \]
\[ S_3(t_s, \delta) = p_is + \delta d^2(t_s) \]
\[ S_4(t_s, \delta) = p_is - \delta d^2(t_s) \]
\[ S_5(t_s, \delta) = p_is + \delta d^3(t_s) \]
\[ S_6(t_s, \delta) = p_is - \delta d^3(t_s) \]

\[ d_M = \sqrt{(S_j(t_s, \delta) - p_is)^T M_{is} (S_j(t_s, \delta) - p_is)} = \delta, \ j = 1, \ldots, 6 \]