LARGE: A Length-Aggregation-based Grid Structure for Line Density Visualization

Tsz Nam Chan
Shenzhen University
edisonchan@szu.edu.cn

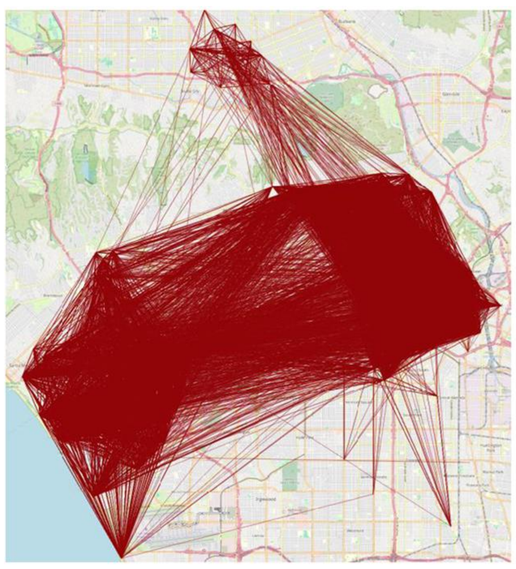
Bojian Zhu
Hong Kong Baptist University
csbjzhu@comp.hkbu.edu.hk

Dingming Wu
Shenzhen University
dingming@szu.edu.cn

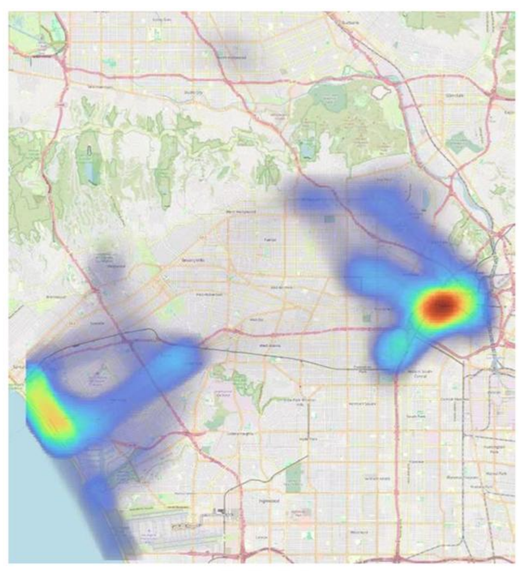
Yun Peng
Guangzhou University
yunpeng@gzhu.edu.cn

Leong Hou U
University of Macau
ryanlhu@um.edu.mo

What is Line Density Visualization (LDV)?



Line segment dataset



LDV

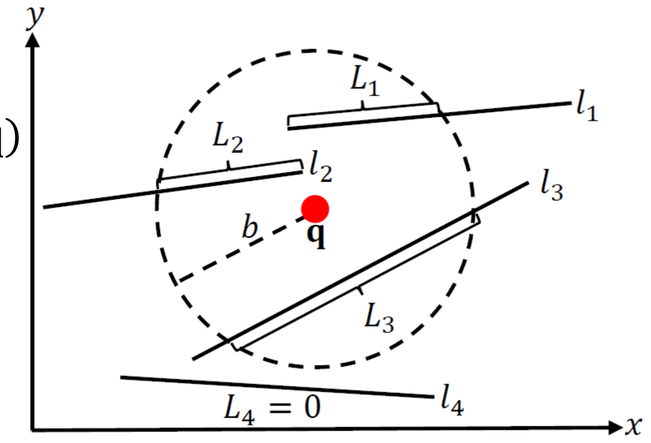
Applications: (i) traffic flow analysis and (ii) mobility analysis

Software packages that support LDV: ArcGIS and QGIS

Given a set of line segments with size n, generating a $X \times Y$-resolution LDV involves the computation of the line density value $\mathcal{L}(\mathbf{q})$ for each pixel $\mathbf{q}$.

$$\mathcal{L}(\mathbf{q}) = \frac{1}{\pi b^2} \sum_{i=1}^{n} L_i$$

where L_i denotes the length of line segment l_i that is within the range b of $\mathbf{q}$.



Weakness of LDV:

(1) LDV takes $O(XYn)$ time, which cannot be scalable to a large resolution size and a large number of line segments. ☹

(2) Commonly used software packages, ArcGIS and QGIS, merely adopt the naïve implementation, which cannot handle large-scale datasets. ☹

Approximate LDV

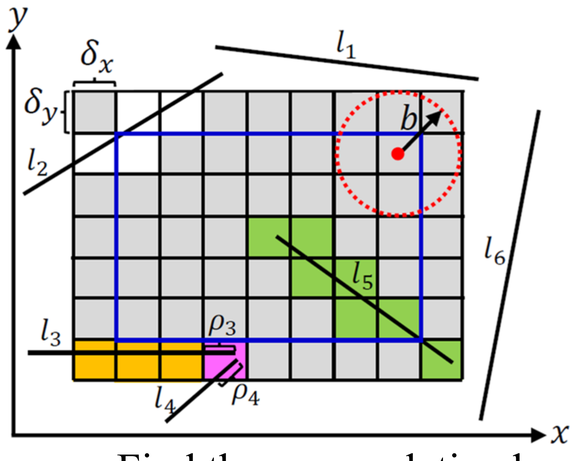
ϵLDV: Given a relative error ϵ, we need to compute $A(\mathbf{q})$ for each pixel $\mathbf{q}$ so that

$$(1-\epsilon)\mathcal{L}(\mathbf{q}) \leq A(\mathbf{q}) \leq (1+\epsilon)\mathcal{L}(\mathbf{q})$$

τLDV: Given a set of D thresholds, $\tau_1, \tau_2, \ldots, \tau_D$, we need to classify $\mathcal{L}(\mathbf{q})$ to be different color levels $C(\mathbf{q})$

$$C(\mathbf{q}) = \begin{cases} 0 & \text{if } \mathcal{L}(\mathbf{q}) < \tau_1 \\ 1 & \text{if } \tau_1 \leq \mathcal{L}(\mathbf{q}) < \tau_2 \\ \vdots & \\ D & \text{if } \mathcal{L}(\mathbf{q}) \geq \tau_D \end{cases}$$

LARGE: A Length-Aggregation-based Grid Structure



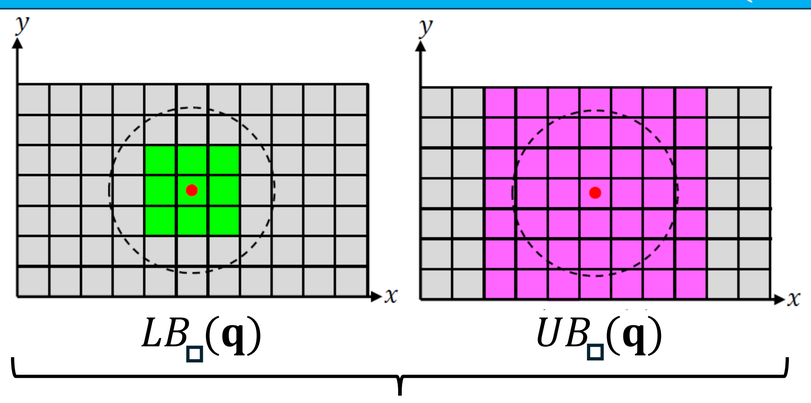
Find the accumulative length for each entry.

0	2.6	2.8	0	0	0	0	0	0
6.6	0.4	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	1.2	1.2	0	0	0
0	0	0	0	0	2.8	2.8	0	0
0	0	0	0	0	0	1.2	6.3	0
5	5	5	8.6	0	0	0	0	5.7

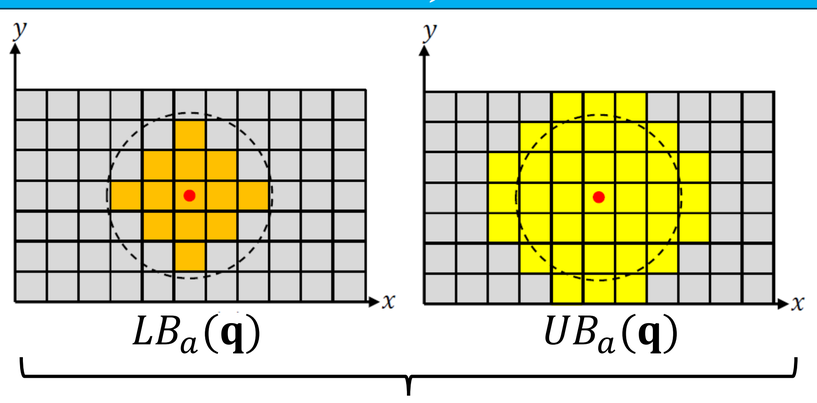
11.6	19.6	27.4	36	37.2	41.2	45.2	51.5	57.2
11.6	17	22	30.6	31.8	35.8	39.8	46.1	51.8
5	10	15	23.6	24.8	28.8	32.8	39.1	44.8
5	10	15	23.6	24.8	28.8	32.8	39.1	44.8
5	10	15	23.6	23.6	26.4	30.4	36.7	42.4
5	10	15	23.6	23.6	23.6	24.8	31.1	36.8
5	10	15	23.6	23.6	23.6	23.6	23.6	29.3

Construct LARGE.

Bound Functions (based on LARGE)

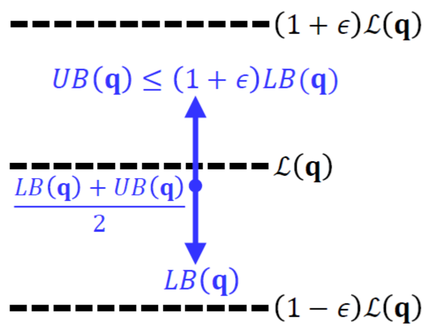


$O(1)$ time

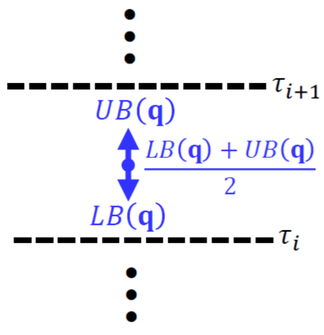


$O(\min(X, Y))$ time

Filtering Conditions



ϵLDV

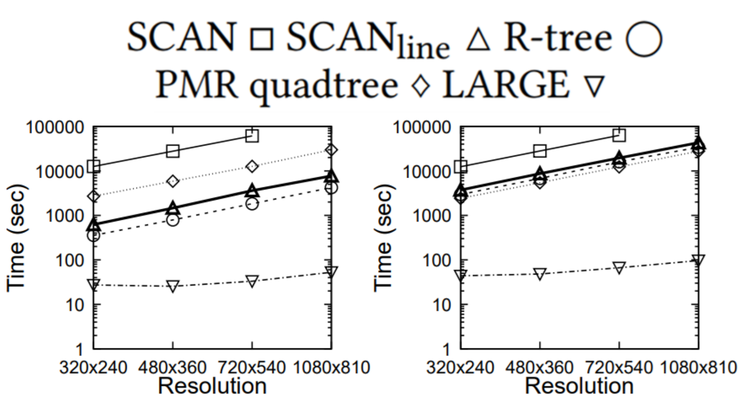


τLDV

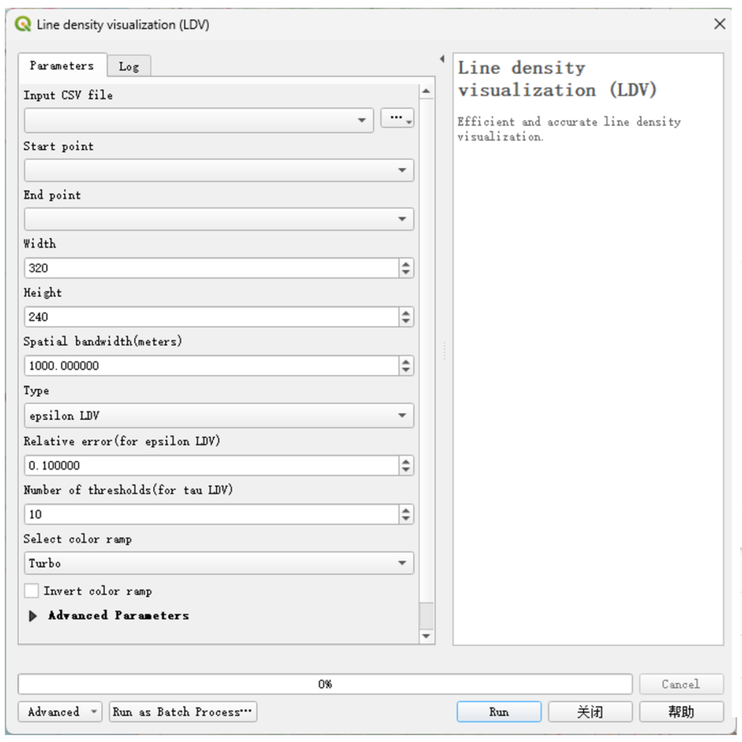
Efficiency Results

Details of datasets

Dataset	n	Category
Los Angeles	402,171	Bicycle mobility
San Francisco	402,602	Taxi mobility



User Interface



Number of Users