Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells

Gengchen Mai¹, Krzysztof Janowicz¹, Bo Yan², Rui Zhu¹, Ling Cai¹, Ni Lao³ ¹STKO Lab, UC Santa Barbara; ² LinkedIn Corporation; ³ SayMosaic Inc.



Embedding clustering of different location encoding models: (a)-(d) baselines (e)-(f) **Space2Vec**

Unsupervised Text Encoding

Position Encoding: encode word positions with sinusoid functions of different frequencies



Transformer (Vaswani et al., 2017) BERT (Devlin et al., 2019)

Unsupervised Location Encoding

1. Radial Basis Function (RBF)

$$K(\mathbf{x},\mathbf{x}') = \exp{\left(-rac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}
ight)}$$

- choosing the correct scale is challenging
- Need to memorize the training samples

2. Tile-based approaches (Berg at al. 2014): discretize the study area into regular grids

- choosing the correct scale is challenging
- does not scale well in terms of memory

3. Directly feed the coordinates into a FFN (inductive single-scale location encoder)

• hard to capture fine grained distributions



Geo-aware Image Classification (Mac Aodha et al., 2019)

Key challenge for location encoding

- Joint modeling distributions with very different characteristics
- => multi-scale location representations





Women's Clothing (Clustered Distribution) Education (Even Distribution)



Stensola et al. (2012) Gao et al. (2019)

Grid Cell Based Multi-Scale Location Encoding



- Grid cells in mammals provide a multi-scale periodic representation that functions as a metric for location encoding.
- It can be simulated by summing three cosine grating functions oriented 60 degree apart (a simple Fourier model of the hexagonal lattice).





Point of Interest Type Classification

POI Groups	Clustered	Middle	Even
	$(r \leq 100m)$	(100m < r < 200m)	$(r \ge 200m)$
direct	0.080 (-0.047)	0.108 (-0.030)	0.084 (-0.047)
wrap	0.106 (-0.021)	0.126 (-0.012)	0.122 (-0.009)
tile	0.108 (-0.019)	0.135 (-0.003)	0.111 (-0.020)
rbf	0.112 (-0.015)	0.136 (-0.002)	0.119 (-0.012)
theory	0.127 (-)	0.138 (-)	0.131 (-)
# POI	16,016	7,443	3,915
Root Types	Restaurants; Shopping; Food;	Beauty & Spas; Health & Medical;	Home Services;
	Nightlife; Automotive; Active	Local Services; Hotels & Travel;	Event Planning
	Life; Arts & Entertainment;	Professional Services;	& Services;
	Financial Services	Public Services & Government	Pets; Education



Geo-Aware Image Classification



⁽Mac Aodha et al., 2019)

Arxiv paper: https://arxiv.org/abs/2003.00824

GitHub Repo: https://github.com/gengchenmai/space2vec

wrap (Mac Aodha et al., 2019) (location only)

 $grid (\lambda_{min}=0.0001, \lambda_{max}=360, S=64)$

theory (λ_{min} =0.0001, λ_{max} =360, S = 64)

 $rbf(\sigma=1k)$

BirdSnap[†]

70.07

77.76

77.98

78.65

77.19

78.65

78.56

79.44

79.35

NABirds[†]

76.08

79.99

80.79

81.11

79.58

81.15

81.13

81.28

81.59