NeurIPS 2021

Qualcomm

NEURAL INFORMATION PROCESSING SYSTEMS

Modality-agnostic topology aware Localization

Farhad G. Zanjani, Ilia Karmanov, Hanno Ackerman, Daniel Dijkman, Simon Merlin, Max Welling, Fatih Porikli

Qualcomm AI Research

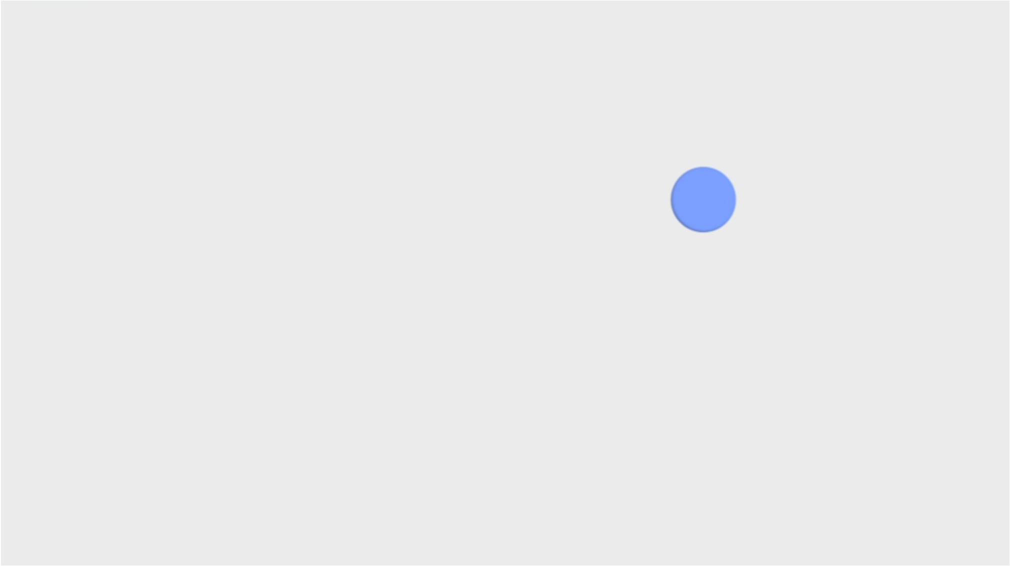
Indoor Localization Problem

Camera-based system

http://svl.stanford.edu/igibson/

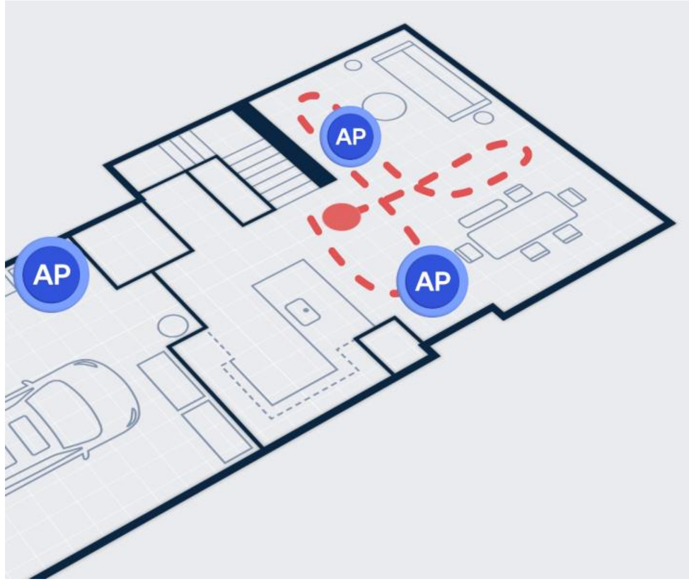
Localizing a person on the 2D map of a building

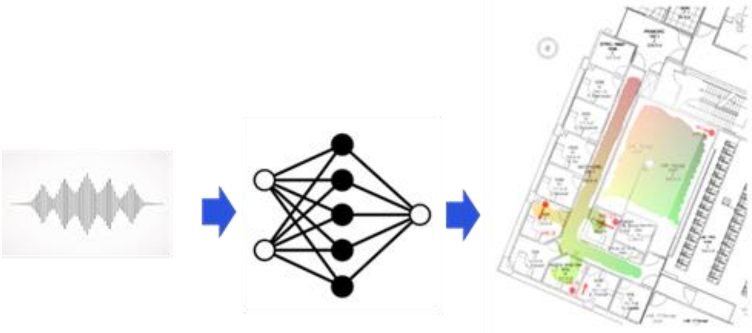
Indoor localization using RF signals (e.g. WiFi)



Play Video

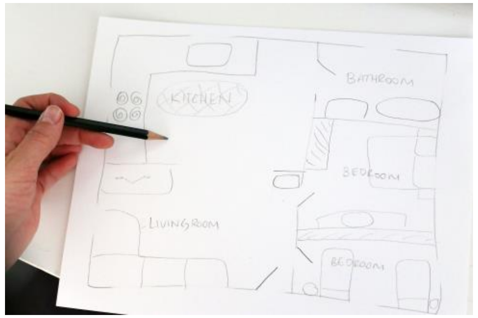
WiFi sensing

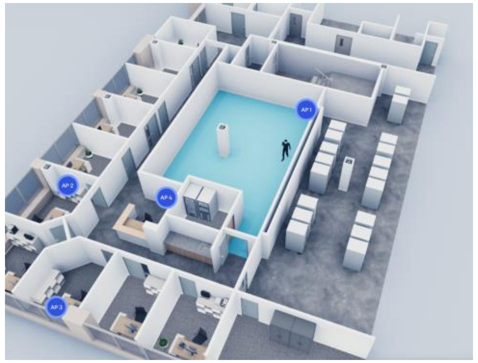




Assumptions

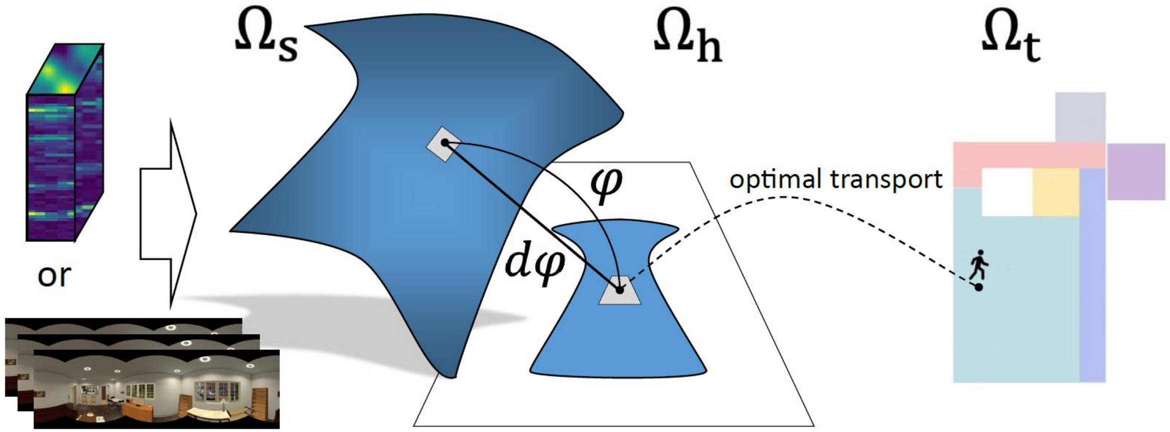
- Sequential data was recorded when the subject visited different locations in the environment
- The 2D/3D map of the environment is given
- The training set includes the (zone-) room-level labels





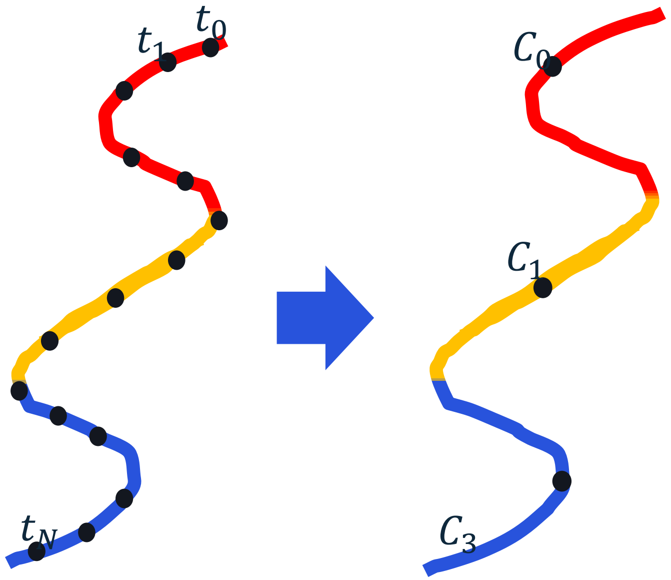
Overview of the proposed method

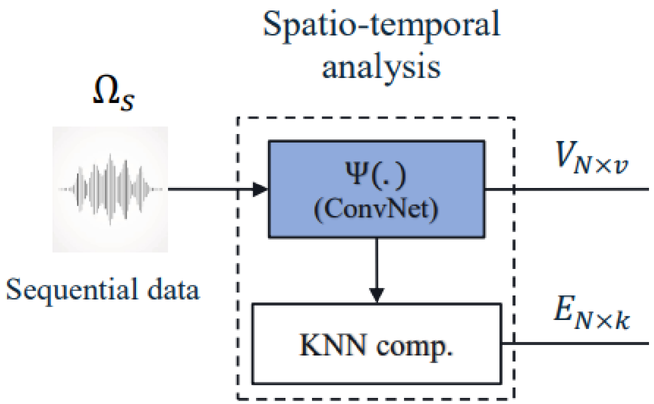
- Estimating the geodesic distances between samples on the data manifold in input space (Ω_s)
- Training a neural network (φ) to map the input samples into their 2D intrinsic space (Ω_s)
- Transporting the 2D representation into the floorplan (Ω_t)

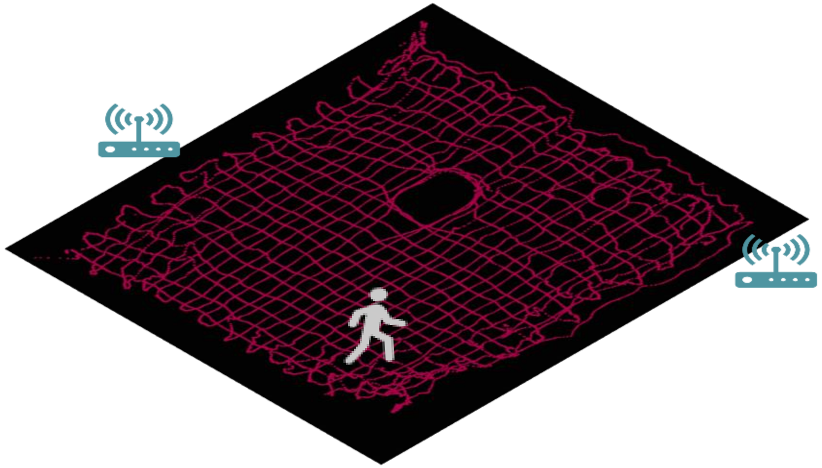


Spatio-temporal analysis

- Estimating the geodesic distances by computing the shortest path on the KNN graph
- Detecting the KNN samples by minimizing the triplet margin loss







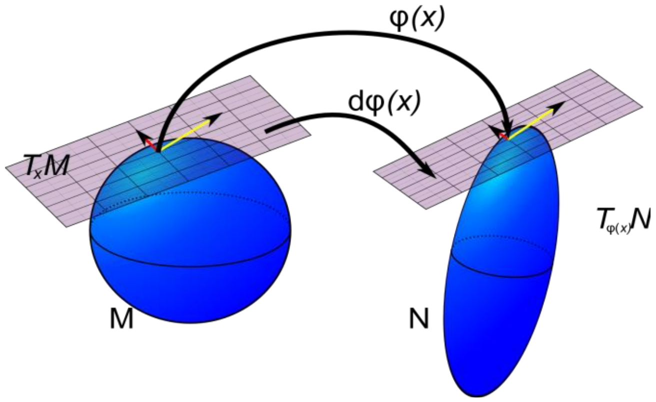
$$\mathcal{L}(x_i^a, x_i^p, x_i^n) = max\Big(0, d(h_i^a, h_i^p) - d(h_i^a, h_i^n) + \alpha\Big) \quad \text{where} \quad h_i = \Psi(x_i).$$

Isometric embedding

Estimating the pushforward around each vertex in the graph

$$\varphi: \mathbb{R}^m \to \mathbb{R}^n \quad , \; m > n$$

$$d\varphi: T_x M \to T_{\varphi(x)} N, \qquad \mathcal{M}(y_i) = \left[\frac{d\varphi}{dx}(x_i)\, \frac{d\varphi}{dx}(x_i)^T\right]^\dagger$$

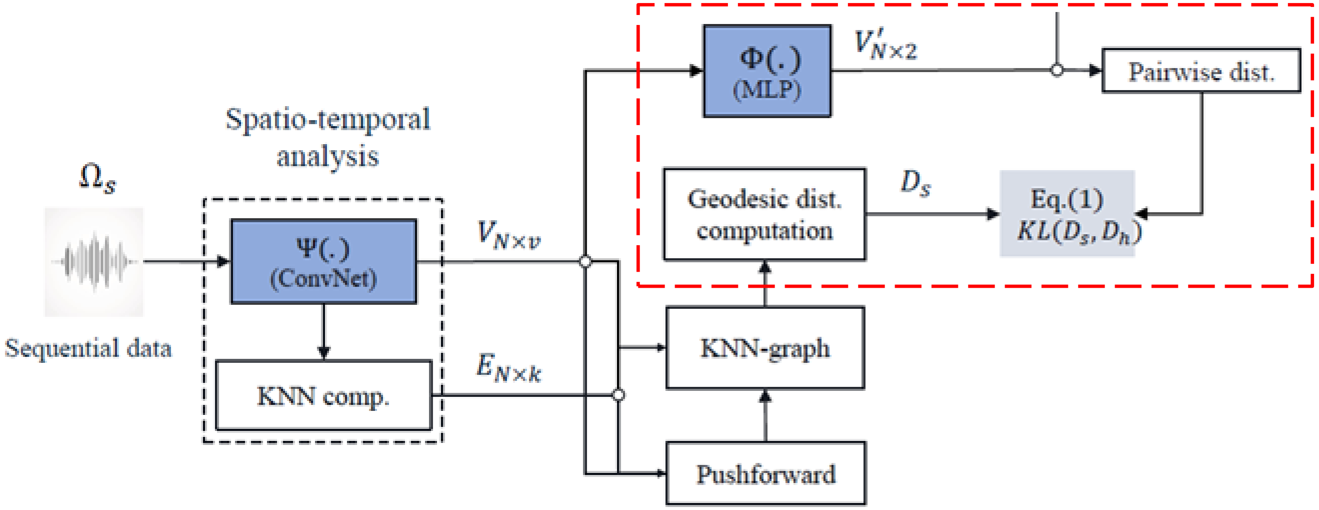


The pushforward at the location of each sample is proportional to the covariance matrix of a Gaussian distribution, centered at that sample [Dsilva et al. 2015]

$$\|\Phi(x_i) - \Phi(x_j)\|^2 \approx \frac{1}{2}[x_j - x_i]^\top \cdot [C^\dagger(x_i) + C^\dagger(x_j)] \cdot [x_j - x_i]$$

Learning the isometric embedding

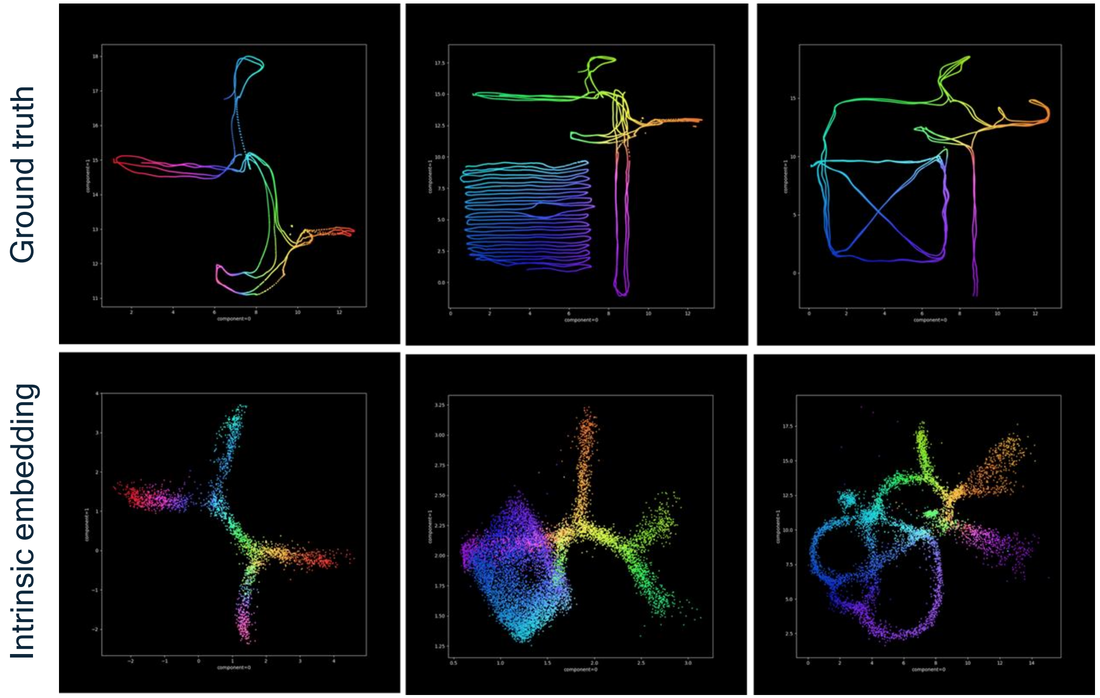
- Training a neural network (Φ) to map the input samples into their 2D intrinsic space (Ω_h)



$$\min_{\Phi} D_{KL}(D_s||D_h), \qquad \text{where } D_s, D_h \in \mathbb{R}^{N_s \times N_s}$$

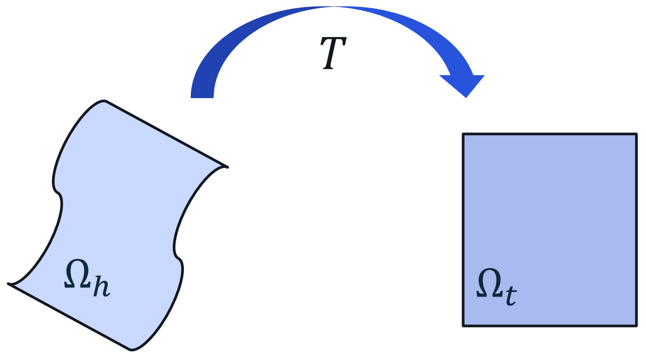
D_s : Geodesic distance matrix of samples in $\mathbb{R}^M$

D_h : Euclidean distance matrix of samples in $\mathbb{R}^2$



Discrete Optimal Transportation using Sinkhorn Distance

- Optimal transportation is used for solving the correspondence problem between the intrinsic space and the given topological map



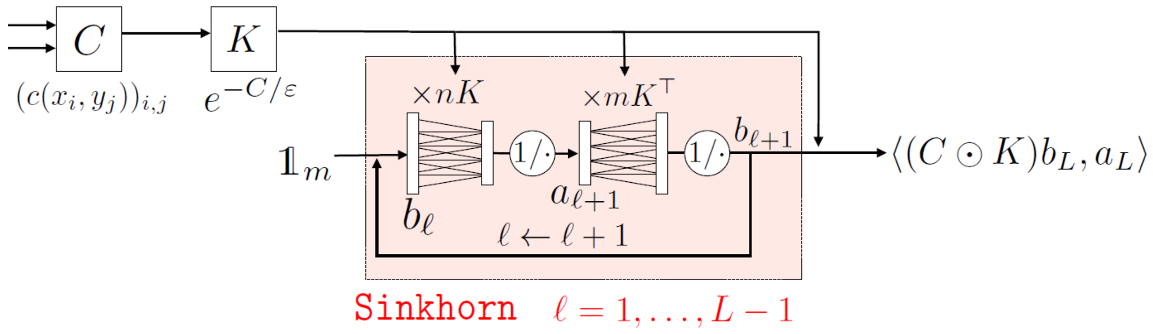
$$T(C,p,q) = \underset{T \in \gamma(p,q)}{argmin} \langle T, C \rangle - \frac{1}{\lambda} H(T)$$

Fixed-point Sinkhorn iterations:

$$T(C,p,q) = \mathrm{diag}(a) K \mathrm{diag}(b)$$

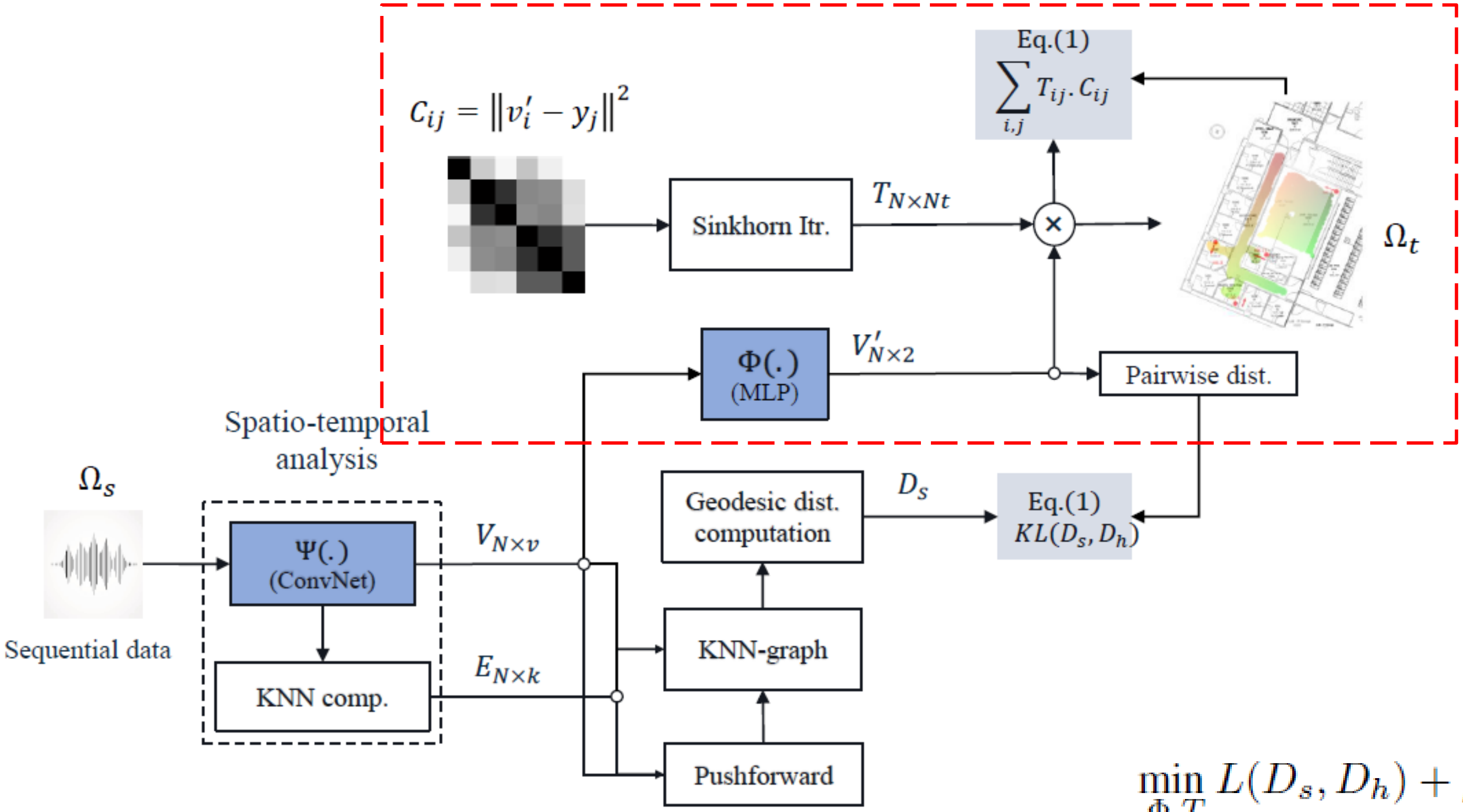
$$K = e^{-\lambda C} \in \mathbb{R}_+{}^{Ns \times Nt}$$

$$a \leftarrow \frac{p}{Kb} \quad \text{and} \quad b \leftarrow \frac{q}{K^\top a}$$



Sinkhorn divergence layer [Genevay et al., 2017]

Jointly learning the embedding and transportation



$$\min_{\Phi, T} L(D_s, D_h) + \mu \sum_{ij} T_{ij} \cdot C_{ij} \quad \text{with} \quad C_{ij} = \|\Phi(x_i) - y_j\|^2$$

distance matrices $D_s, D_h \in \mathbb{R}^{N \times N}$

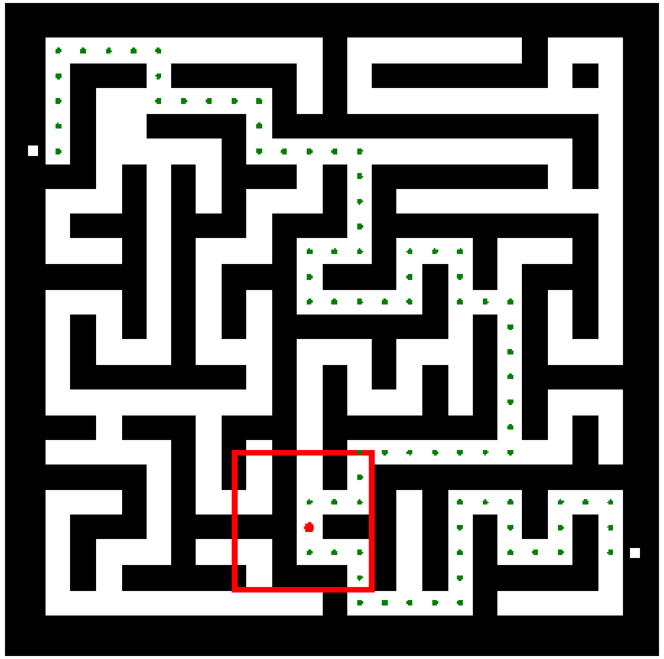
Cost matrix

(Densely) sampled points from $\Omega_t \in \mathbb{R}^2$

Experiments

- Walking through a synthetic 2D Maze environment

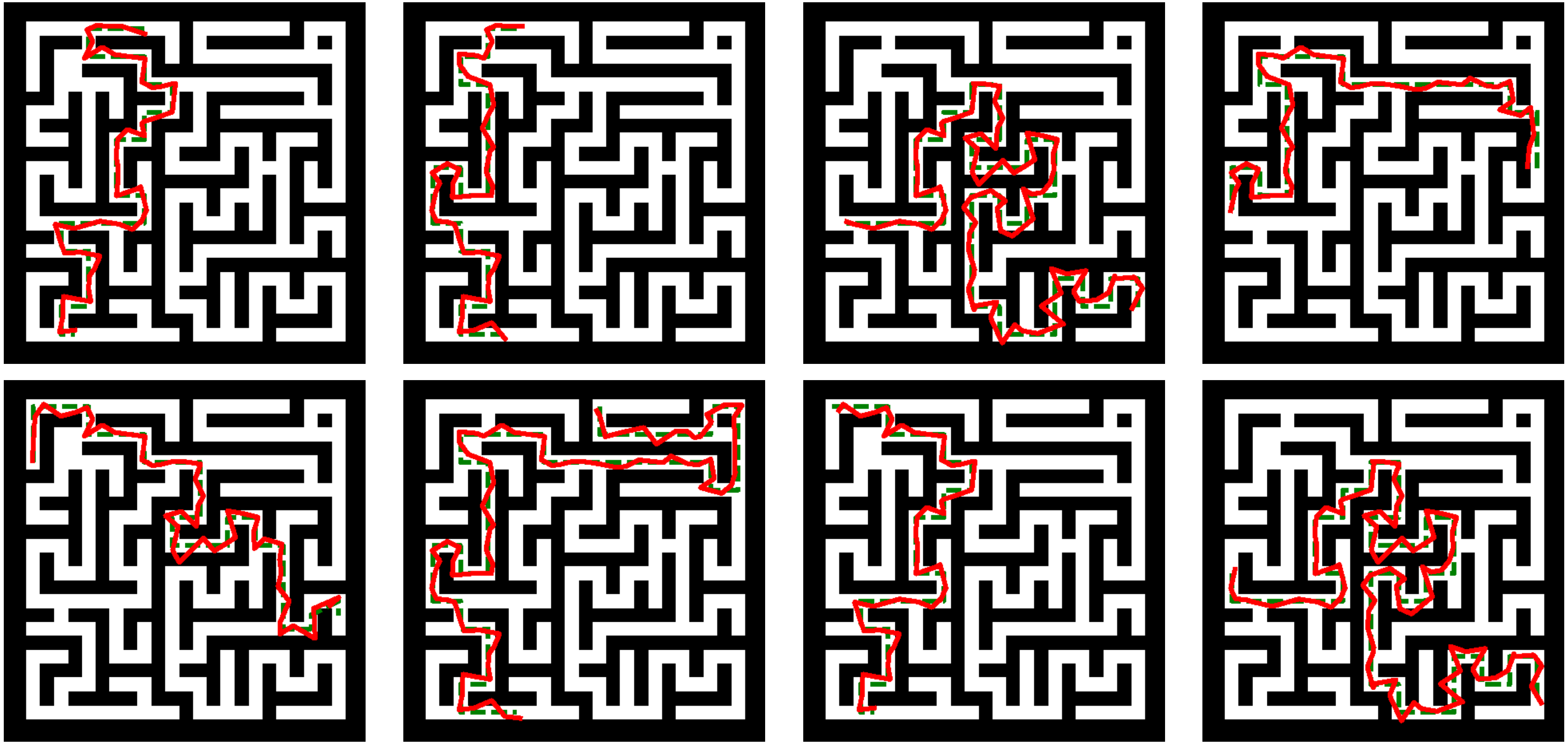
Example of a trajectory



Sampled patches

Experiments

- Synthetic 2D Maze environment



Experiments

- Camera-based localization

iGibson 3D dataset

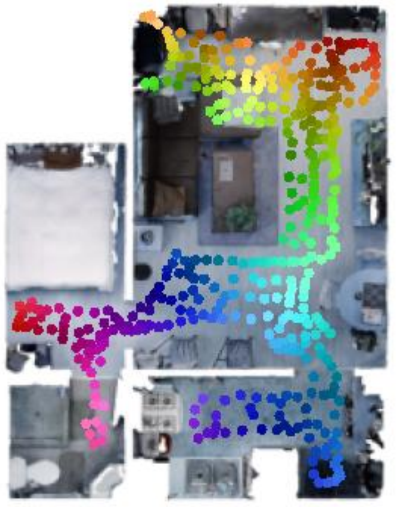
http://svl.stanford.edu/igibson/

Experiments

- Camera-based localization

	Environments														
index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ϵ	0.62	1.05	0.73	0.64	0.69	0.83	1.02	1.03	0.73	0.62	0.74	0.71	0.98	1.04	1.14

Embedding representation



Ground-truth positions

Test trajectory

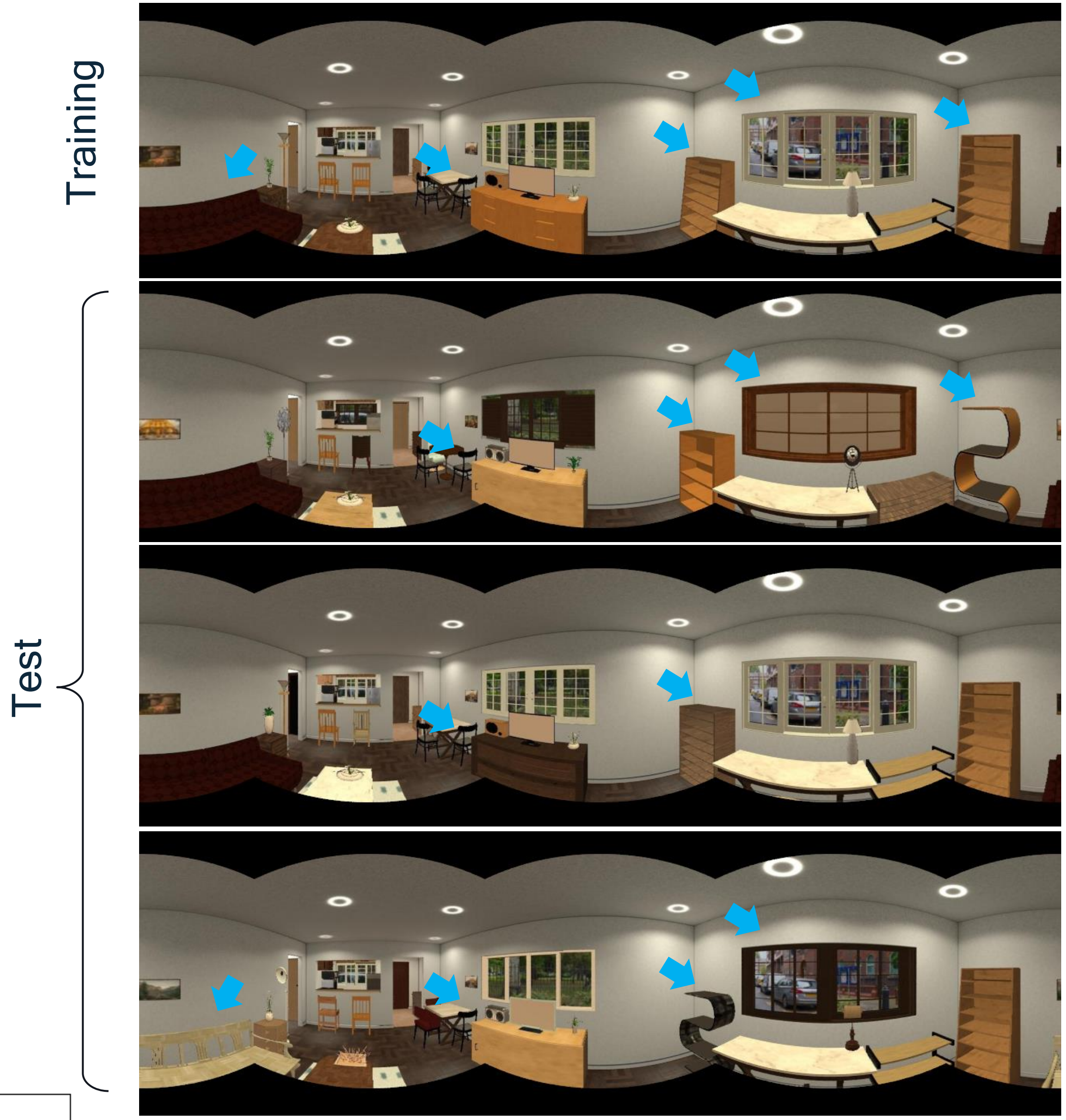
prediction

Experiments

- Altering object appearance in the environment

Demo of iGibson dataset

http://svl.stanford.edu/igibson/



Mean and standard deviation of error

	Perturbation rate					
Rate	0%	10%	20%	30%	50%	70%
Mean of error	61.9	61.87 ± 0.09	62.2 ± 0.09	62.82 ± 0.17	63.77 ± 0.14	64.07 ± 0.26

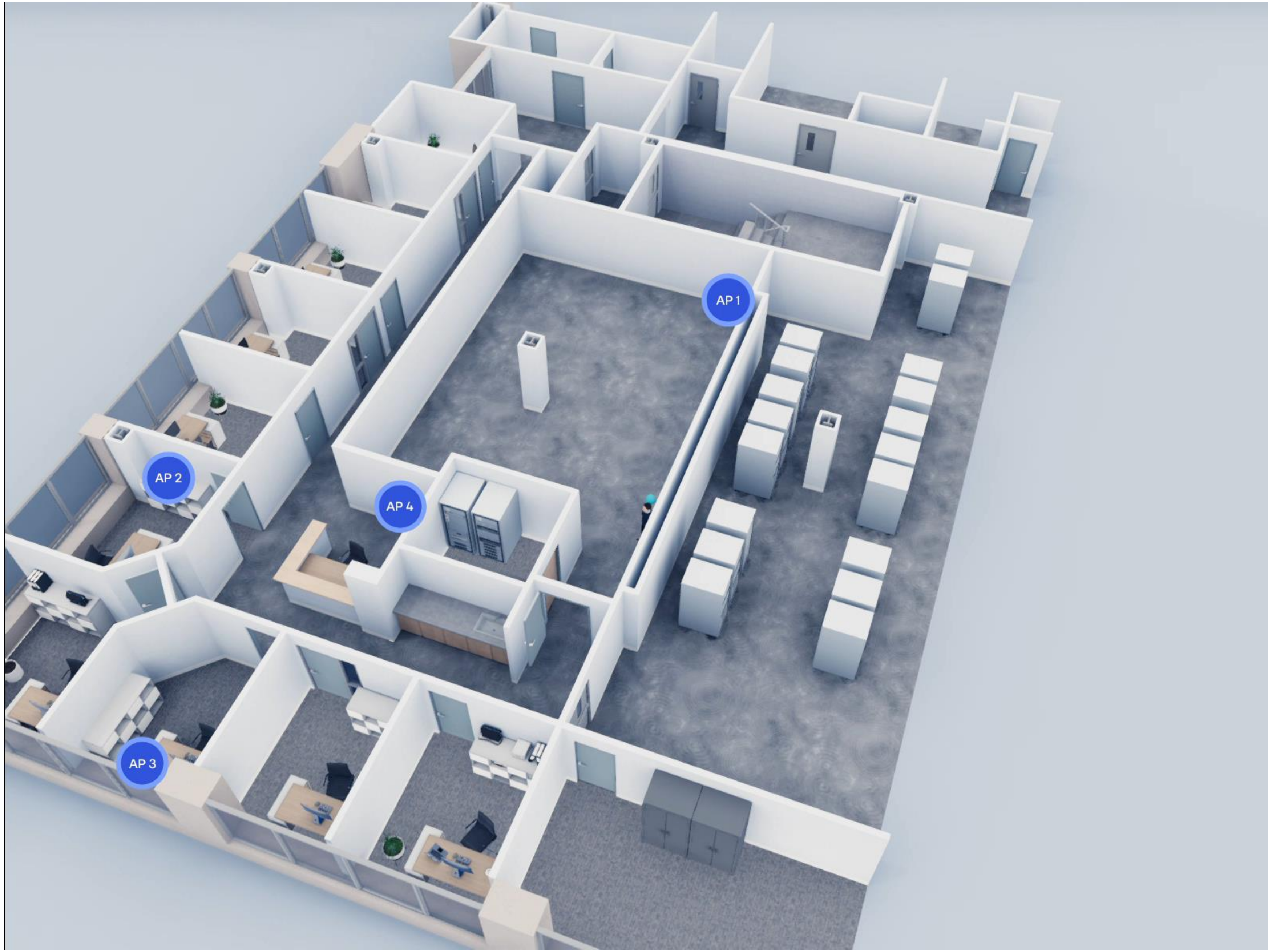
Experiments

- Localization in WiFi

- Commercial IEEE 802.11 access points (AP), 5GHz band

- Size of sensing environment 14×20 meters

- Mean error: 1.2 m



Summary

- We present a learning method for localization problem
- The proposed method is based on parametric manifold learning and optimal transportation
- Our method does not require the 2D coordinates of the moving target during training
- The proposed method does not make any assumption about the data modality in use and in that sense, it is modality agnostic and can be applied on a large family of sensory system.

Qualcomm

Thank you

Follow us on:

For more information, visit us at:

www.qualcomm.com & www.qualcomm.com/blog