



HAL
open science

Hierarchical Sparse Dictionaries: how sparsity, topology and pooling can improve visual place recognition

Sylvain Colomer, Nicolas Cuperlier, Guillaume Bresson, Olivier Romain

► **To cite this version:**

Sylvain Colomer, Nicolas Cuperlier, Guillaume Bresson, Olivier Romain. Hierarchical Sparse Dictionaries: how sparsity, topology and pooling can improve visual place recognition. Human-like Behavior and Cognition in Robots - workshop at IROS, Oct 2021, Prague, Czech Republic. hal-03469399

HAL Id: hal-03469399

<https://hal.science/hal-03469399>

Submitted on 13 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hierarchical Sparse Dictionaries: how sparsity, topology and pooling can improve visual place recognition

Sylvain Colomer^{1,2}, Nicolas Cuperlier², Guillaume Bresson¹, Olivier Romain²

¹Institut de recherche Vedecom, France ; ²Laboratoire ETIS, CY Cergy Paris Université, ENSEA, CNRS, France

Context

- Visual place recognition (VPR) for self-driving vehicles [1].
- Large-scale environment and computational cost.
- Neurocybernetics approach.

Contributions

- HSD, a sparse model for visual information encoding.
- HSD+MP, a full visual neuro-inspired localization model using HSD.

HSD : Hierarchical Sparse Dictionaries

General overview:

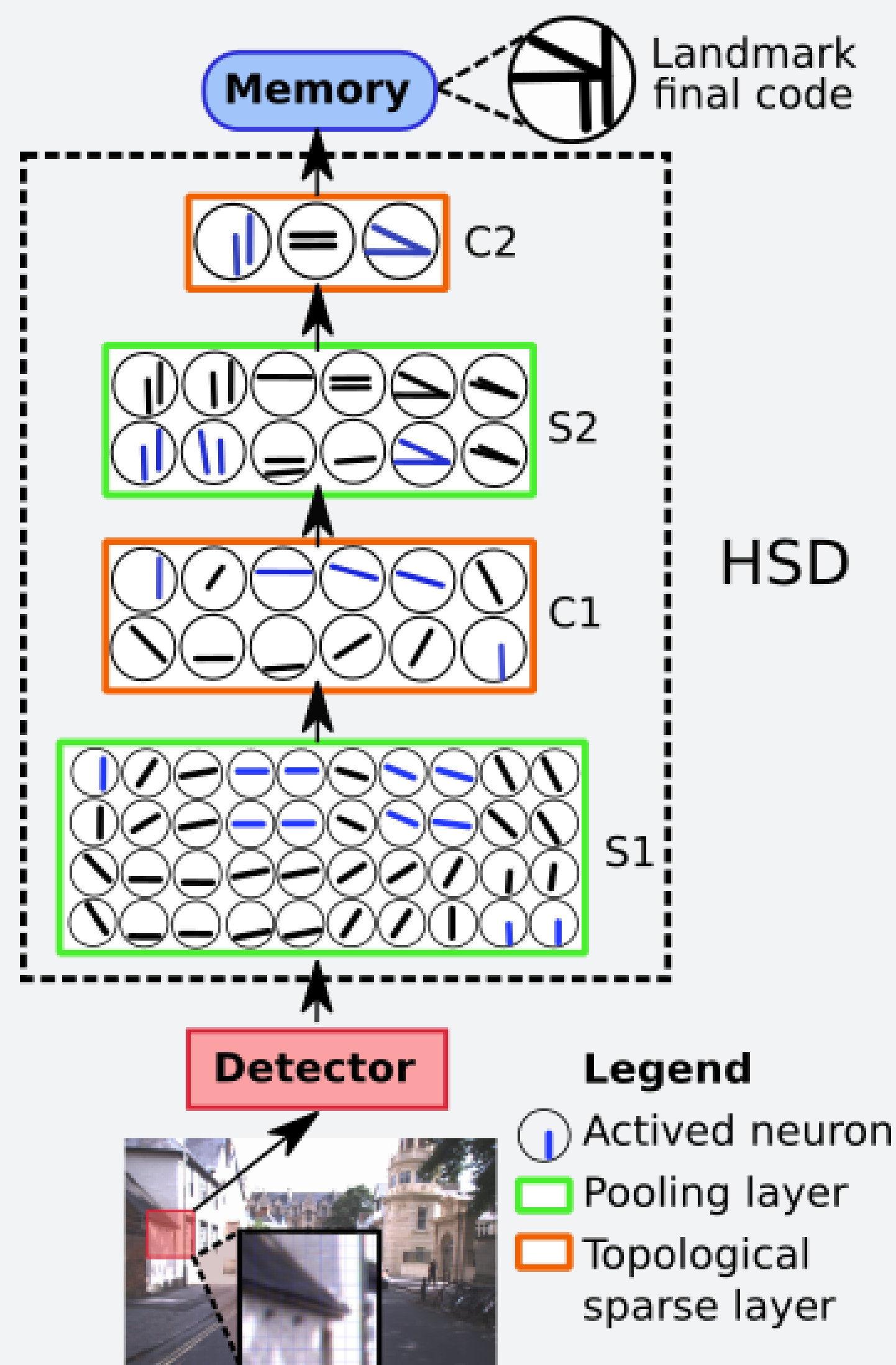
- Our unsupervised neural network aims to compress visual information.
- Alternates "Topologic Sparse Layers" and pooling layers.
- Inspired by visual cortex key properties: sparsity, topology and pooling [3].

Topologic Sparse Layers:

- Sparse dictionary sorted via an algorithm of Self-Organising Map.
- Put side by side neurons sensitive to similar orientations.
- Replicate the behaviour of visual cortex *simplex cells*.

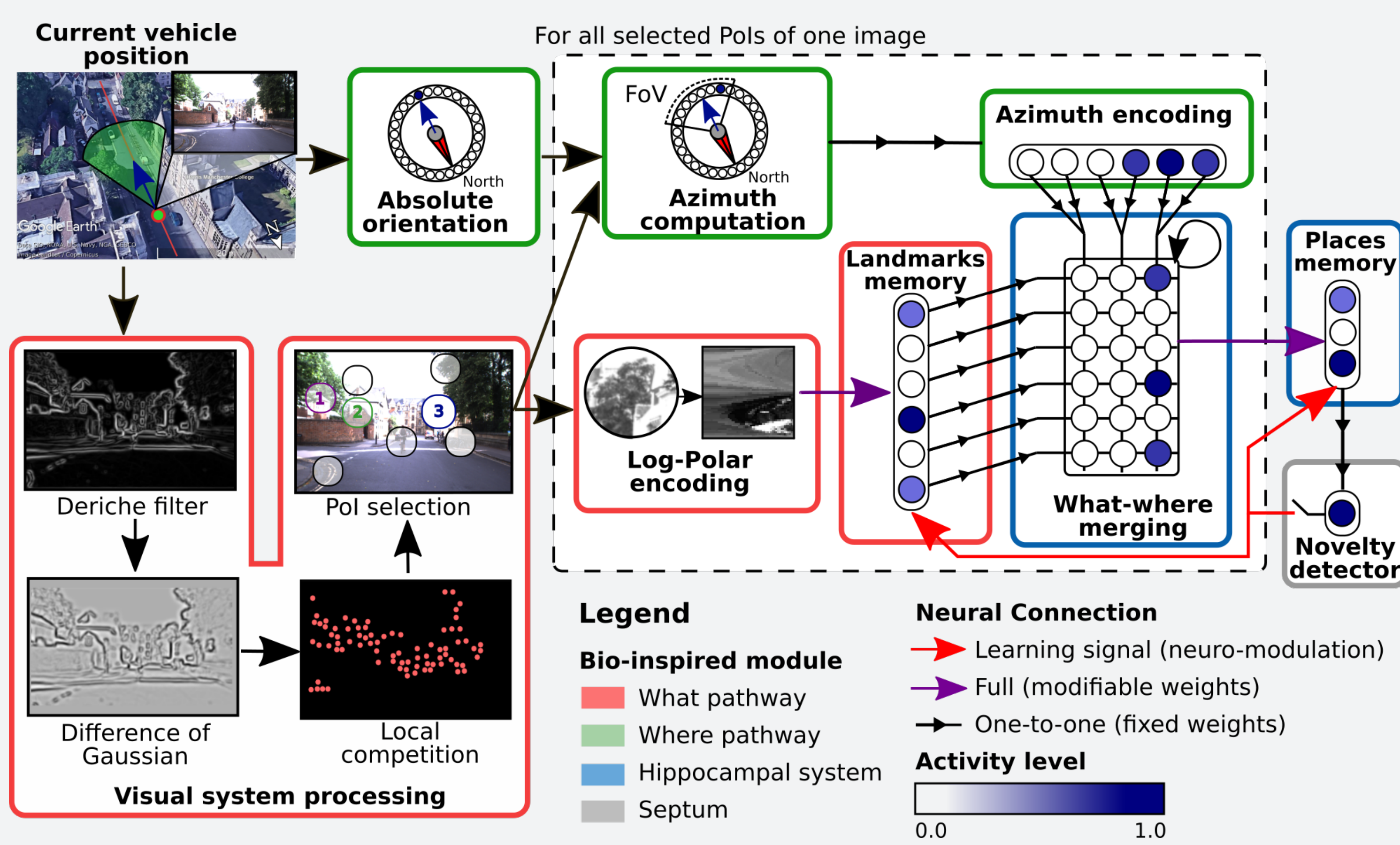
Pooling layers:

- Merge information from S1 and S2 layers to create spatial invariance.
- Replicate the way the visual cortex operates (*complex cells*).



HSD+MP : application to visual localization

- Modification of LPMP, a bio-inspired model of VPR [2], to integrate HSD.
- HSD replaces the *log-polar* block to compress landmarks representation.



Experiments

Evaluation methodology: Ability of models to recognize already visited places.

- Three distances of sampling on learning dataset : 2m, 3m and 5m.
- Six configurations of HSD : five balanced and one (HSD-15/30) unbalanced.
- Comparison with a state-of-the-art VPR model named CoHog.

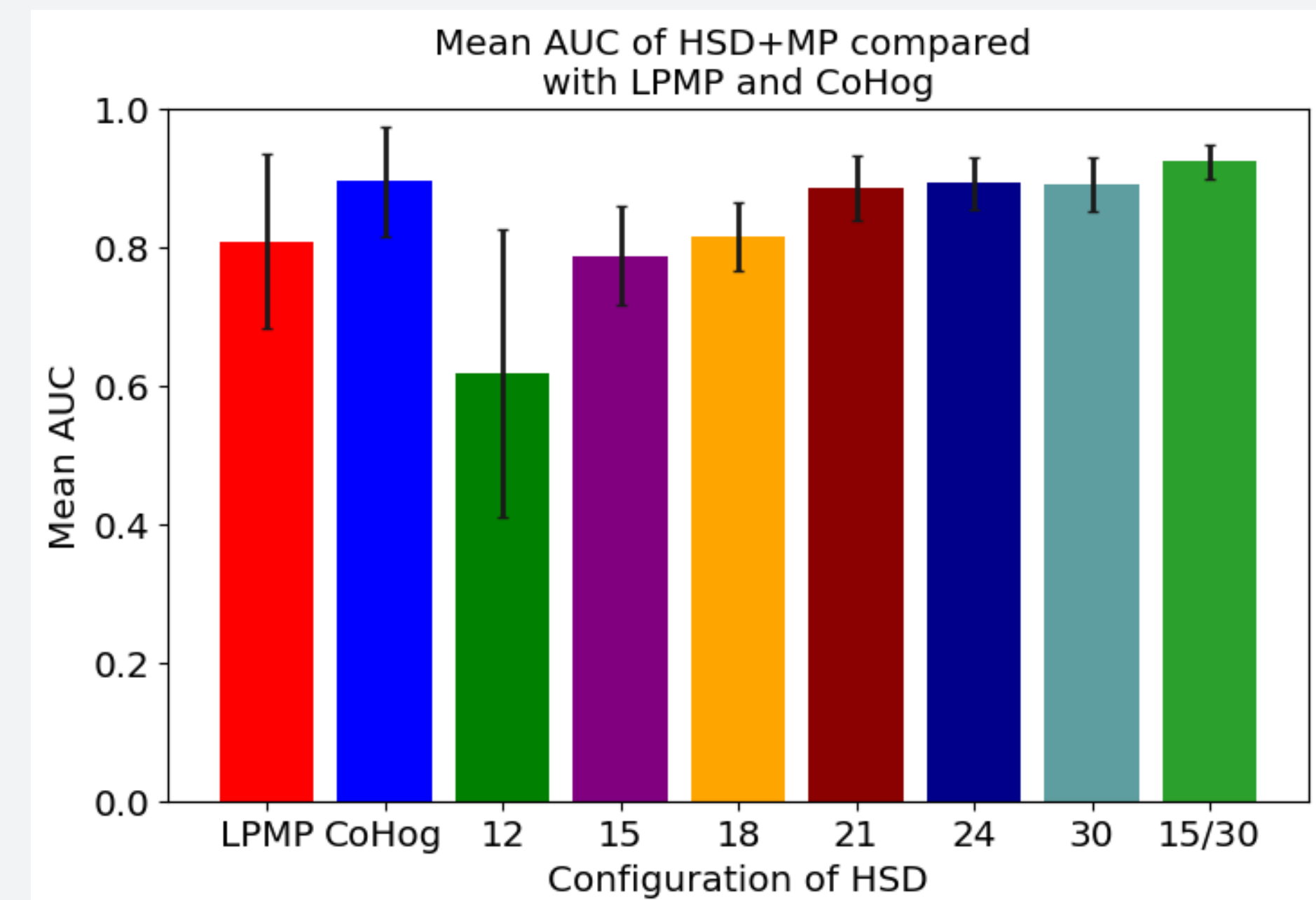
Metrics: AUC (area under the curve) of precision/recall curves and average processing frequency of an image.

Dataset: Three routes with two distinct trajectories extracted from the KITTI dataset.

Results

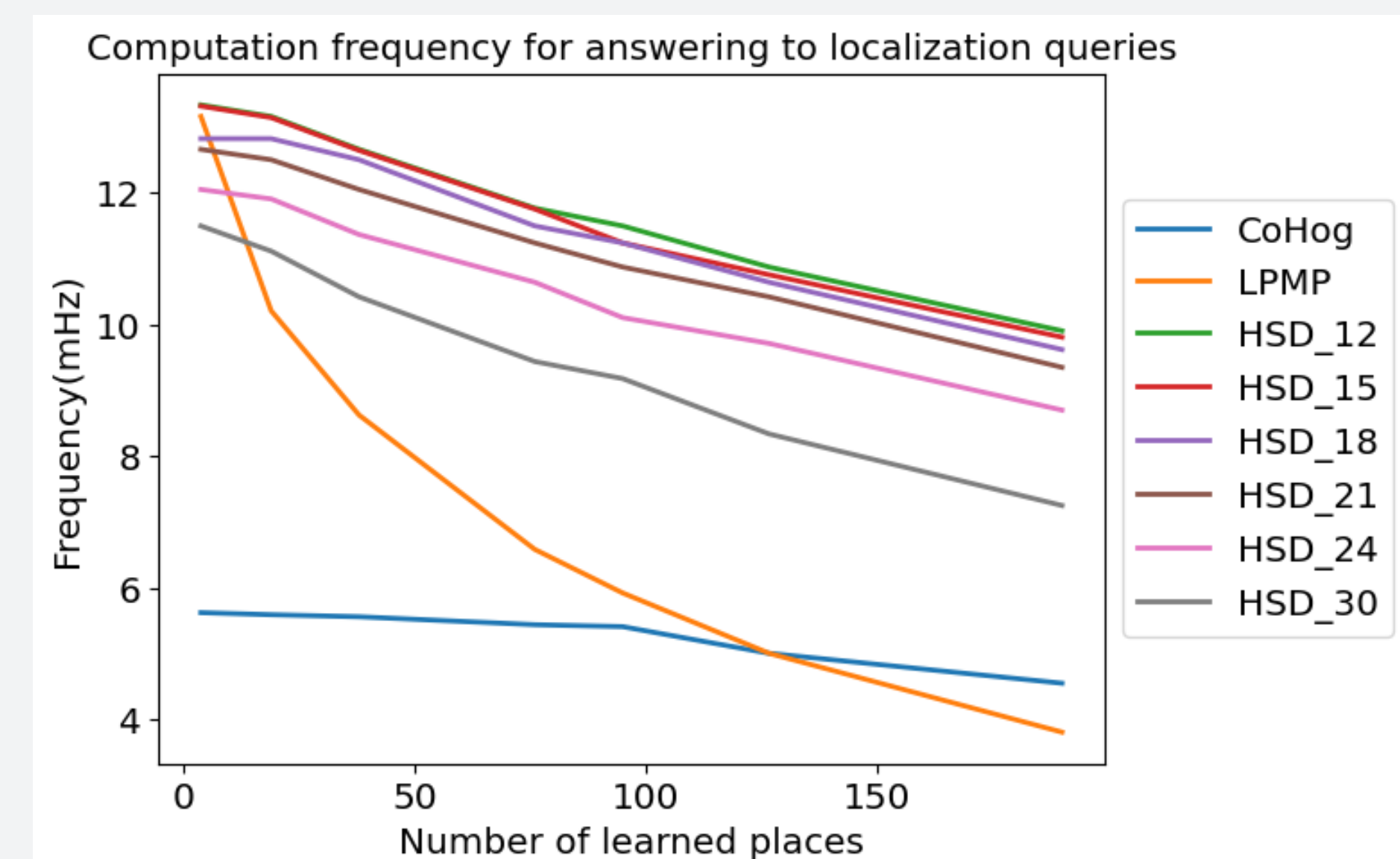
AUC by configuration:

- Mean AUC increases with the size of HSD.
- Gave better localization performances than LPMP with HSD-18.
- Gave better localization performances than CoHog with HSD-24.
- Best performances with HSD-15/30.



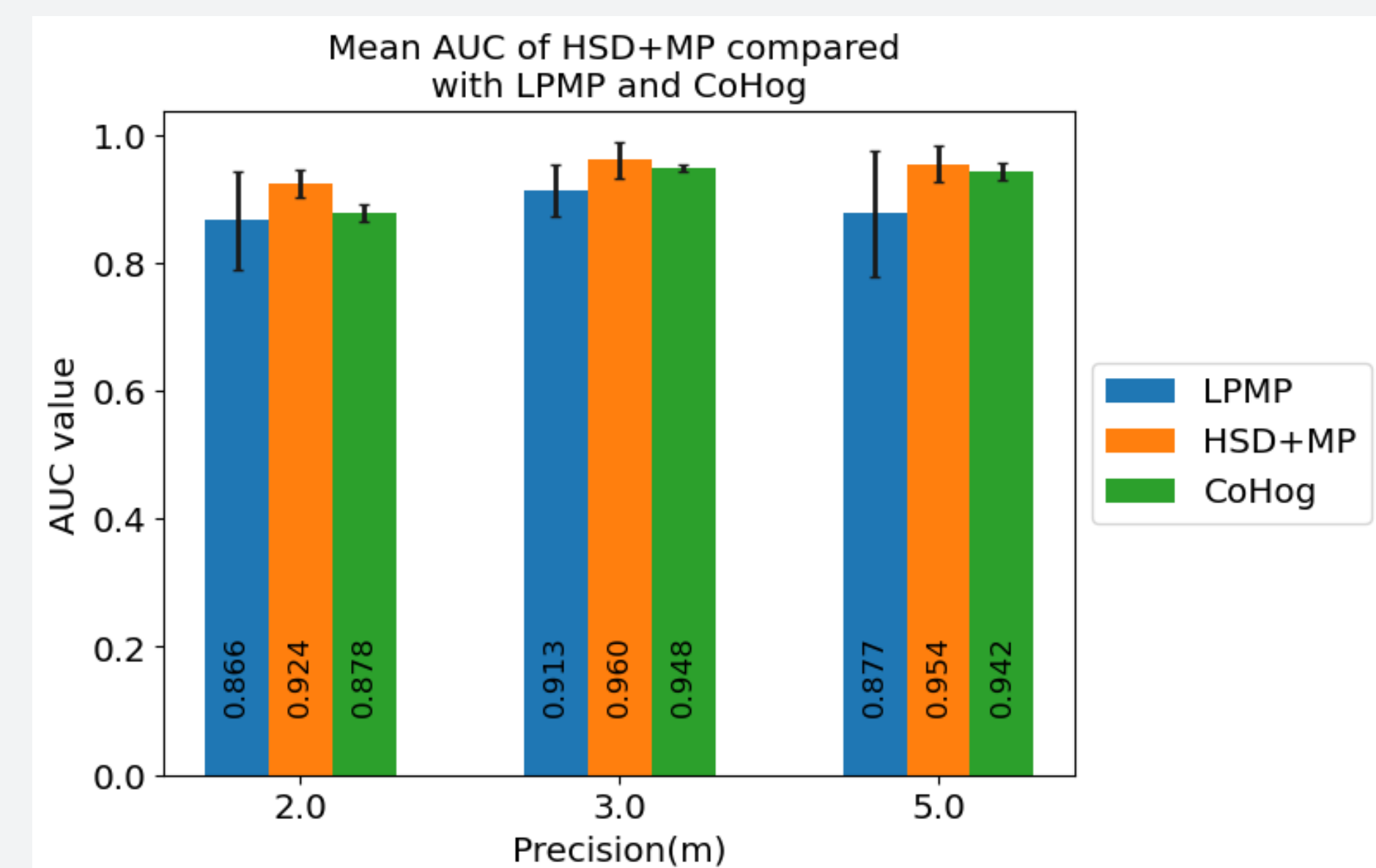
Frequency by configuration:

- Mean computation frequency decreases with the size of HSD.
- HSD-18 is $\times 0.54$ faster than LPMP at equal performance.
- HSD-24 is $\times 0.45$ faster than CoHog at equal performance.



Mean AUC by place sampling:

- HSD improved the mean AUC of LPMP by 7% for all the place sampling rates.
- Give similar results to CoHog for all the place sampling rates, with a slight advantage for HSD+MP.



Learning time by configuration:

- Learning time of HSD increases according to the size of HSD.
- Configuration HSD-15/30 reduces learning time with similar performances.

Model	Learning time	Pol descriptor size
LPMP	.	2916
HSD-15	262s	64
HSD-18	313s	81
HSD-24	555s	144
HSD-30	1527s	225
HSD-15/30	673s	225

Conclusion

- HSD improves computation cost and localization accuracy of LPMP.
- Allows a trade-off between localization and time performance.
- Relatively quick learning.

Future work

- Optimisation of HSD to achieve maximum performance.
- Complexification of the model : more layer, different pooling size etc.
- Evaluation on a more difficult dataset (Oxfordcar dataset).

References

- [1] Y. Espada, N. Cuperlier, G. Bresson, and O. Romain. From Neurobotic Localization to Autonomous Vehicles. *Unmanned Systems*, 07(03):183–194, July 2019.
- [2] P. Gaussier and S. Zrehen. Perac: A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, 16(2–4):291–320, Dec 1995.
- [3] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(3):411–426, Mar 2007.