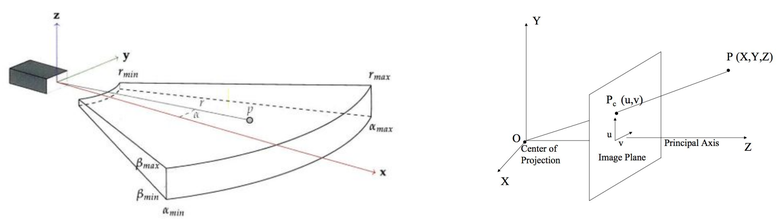
Open Set Opti-Acoustic Semantic Mapping for Underwater Vehicles

Kurran Singh
singhk@mit.edu

Prof. John Leonard[1]
1. Massachusetts Institute of Technology

Motivation

- **Semantic mapping** (object-based mapping) enables higher level autonomous behavior
  - Identify and more closely map assets of interest
  - Identify and avoid high-risk assets
  - More understandable map for human operator
- **Open set** semantic mapping means that the vehicle is not constrained by only being able to detect and map previously seen and selected objects
- **Opti-acoustic** mapping aggregates camera and sonar data to mitigate the shortcomings of each individual sensor



Above left: The geometry of a sonar sensor, where we lose the elevation angle β.
Above right: The geometry of a monocular camera, where we lose the range or depth out of the image plane z.
Only by combining the two sensor modalities is it possible to localize an object in 3D space.

Research Challenges

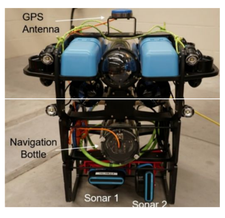
- Discrete-continuous optimization problem
  - Semantic mapping has discrete aspects (object classes, data associations *D*) and continuous aspects (vehicle poses *X* and object locations *L*)
  - Formulate this as a *maximum a posteriori* problem below with V_k all measurements

$$X^*, L^*, D^* = \underset{X,L,D}{\operatorname{argmax}} \prod_k \phi_k(\mathcal{V}_k)$$
$$= \underset{X,L,D}{\operatorname{argmin}} \sum_k -\log \phi_k(\mathcal{V}_k).$$

- Underwater object detection
  - Lack of labeled datasets for training
- Camera-sonar correspondences
- Reflections
- Integrating several individual research contributions into single system

Experimental Platform

Implementation and testing of full system on vehicle below



References

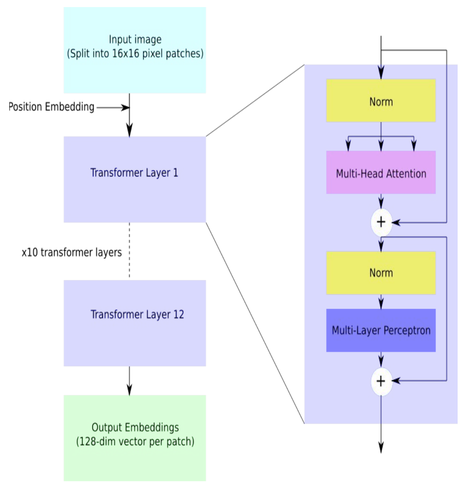
Singh, K., Rypkema, N., & Leonard, J. (2023). Attention-based Self-Supervised Hierarchical Semantic Segmentation for Underwater Imagery. IEEE OCEANS 2023.

Doherty, K. J., Lu, Z., Singh, K., and Leonard, J.J. Discrete-continuous smoothing and mapping. IEEE Robotics and Automation Letters, 7(4):12395–12402, 2022.

Methods

Object detection

- Self-supervision method through a student-teacher multicrop method addresses the lack of labeled underwater datasets
- Outputs semantically meaningful features that are clustered into objects *without* needing labels → open set mapping

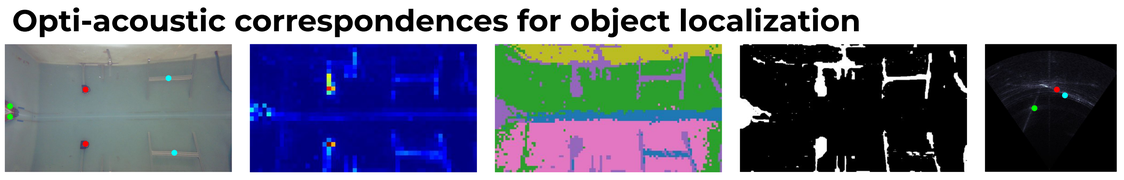


The network architecture used for embedding the images into the latent space

$$\min_{\theta_s} \sum_{x \in \{x_1^g, x_2^g\}} \sum_{\substack{x' \in V \\ x' \neq x}} P_t(x) log(P_s(x'))$$

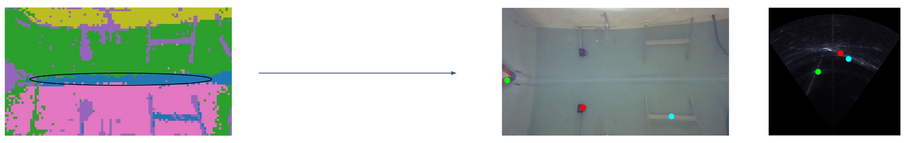
A cross entropy loss between the student (s) and teacher (t) outputs is calculated to update the network weights. By feeding smaller crops of the image (x') into the student network, local to global correspondences are learned.

Opti-acoustic correspondences for object localization



Left to right: 1) The raw input image, with detected object centroids colored by class 2) A visualization of the latent space 3) The object segmentations from KNN clustering 4) A saliency mask to remove background, as calculated by examining attention-head weights in network 5) The sonar image and corresponding object centroid detections used to determine range to object

Reflection Removal



Left: By segmenting and detecting the water line, we can remove all object detections above that line.
Right: Now only the actual objects are used for mapping, rather than including reflected objects.

Data association

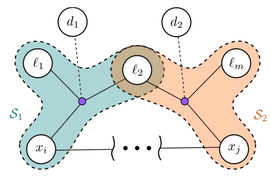
- Cosine similarity of object centroids in feature space
- Multiple hypotheses tracked through a mixture of Gaussian approach

$$\frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \cdot \sqrt{\sum_{i=1}^{n} B_i^2}},$$

Calculation of cosine similarity between two latent vectors A and B

Discrete-continuous smoothing and mapping

- Probabilistic graphical framework to tie together and optimize above system by finding *maximum a posteriori* estimates of object locations and classes, as well as vehicle poses



Co-funded by:

Fundação para a Ciência e a Tecnologia

MIT Portugal

under the Flagship Project: K2D | Earth Systems: Oceans to Near Space