Martech 2023.
Marine Technology Workshop
19,20th June. Castellón de la Plana, València, Spain

Towards Hash Similarity and Geometric-based Image Filtering to Lighten Underwater Photo-Mosaicing

Pol Baños Castelló[1], Francisco Bonin-Font [2]

[1,2] Systems, Robotics and Vision Group, Department of Mathematics and Computer Sciences, University of the Balearic Islands (UIB), Ctra. Valldemossa Km. 7.5, 07122 Mallorca, Illes Balears, Spain

***Abstract* – This paper presents a novel strategy to filter out images that contain redundant information when building photo-mosaics from the sea bottom. The goal is to alleviate computational resources and processing time in their construction. The algorithm operates in a set of sequential steps, extracting images from video sequences following camera translation and rotation pre-defined constraints and removing redundant data using hash similarity and SIFT feature matching techniques.**

***Keywords* – *Photo-Mosaicing, Image Hash, SIFT features*.**

I. MOTIVATION, TECHNICAL BACKGROUND AND EXPERIMENTAL RESULTS.

The interest in exploring the sea bottom with advanced technology, such as, towed UWTV systems, ROVs or AUV equipped with cameras, combined with image post-processing techniques, has been growing in the last years since it permits to extend underwater observation missions in depth, space and time, without risks on humans. Photo-mosaics offer a unique, colored and detailed view of the recorded seabed area, as a single big spot, and they can be build either from just visual information [1] or combining it with the camera localization estimated by means of navigation instruments resident on the carrying vehicles [2]. However, a specific criteria to select the images for the mosaic that includes a trade off between complete informativeness of the whole recorded area with a minimum of redundant information remains still unformalized, leaving a wide open field for further improvements: a) How many images shall be added to the mosaic to cover the whole area minimizing the gaps, b) The acceptable percentage of overlap between images that view partially the same area, and c) How many opperlapping images can be accepted to avoid areas with repeated views or unrealistic or distorted textures due to color blendings on zones where various images overlap. This paper presents a novel strategy to select images from a dataset that will be used to form a mosaic. The process has been designed to run on the ROS platform [3] because this is the middleware that has major impact on field robotics nowadays. The successive steps work and have been tested as follows: 1) Mosaics are build from images grabbed *in situ* by a moving platform, and stored in ROS *bag* files [3] that also contain the platform navigation data (global pose, velocity, altitude, depth, etc..); the platform needs to have a reliable self-localization module based on its navigation instruments to provide the camera pose and velocity, 2) consecutive images contained in the *bag* file are decimated by either 2 or 4 and stored in the hard disk together with their corresponding localization, at a maximum rate of 2.5 fps, considering a maximum overlap of 75%, and a minimum platform velocity, 3) besides, the image storage is also constricted to a minimum and a maximum altitude to guarantee that images grabbed during the platform immersion and emersion are rejected, and a minimum variation in yaw in consecutive frames to avoid including many images of the same position when the camera turns essentially in yaw with no traslation, 4) for each stored image, SIFT keypoints and descriptors [4] and the HALOC [5] image hash are also computed and stored for further comparison and filtering, 5) each individual image, called from now on a *query*, is related with all images (from now on called *candidates*) already stored and that were grabbed inside a circular Region of Interest (ROI), centered in the *query* global pose and with a radius parametrized between 2.5 and 3 meters, 6) afterwards, a double layer filter is applied, namely, a first coarse one, to select *candidates* inside the ROI with a difference (L1-norm) between their hash and the hash of the *query* (so called hash similarity) lower than a certain threshold, and a second and fine, to keep only those pre-filtered *candidates* that match with the *query* a number of features (rejecting outliers with RANSAC) higher than another predefined threshold; filtering first by hash similarity permits to select *candidates* that potentially and most likely overlap with the *query* saving time and computer resources [5], while requiring a minimum robust feature matchings confirms solidly that the *query* and the finaly selected *candidates* overlap *(*the greater the number of matchings, the higher should be the overlap), and, finally, 7) for each *query*, the algorithm discards for the mosaic all *candidates* that survived the second filter, that is, those that presumably view an area most in common with the *query*, and, in consequence, those that represent redundant information that can be omitted; since this process is applied continuously, while the *bag* file is running, discarded images are not taken again as possible *candidates*, been removed from the whole process. Finally, the mosaic is constructed placing the remaining images in the mosaic frame according to the camera global pose (in horizontal coordinates [x,y] with respect to the origin of the mission) converted from meters to pixels thanks to the known calibrated focal length and the vehicle navigation altitude observed at each frame timestamp. Thresholds have to be adjusted depending on the site, the texture of the sea bottom and a relation between overlap and number of matchings observed in each different type of environment. Test include, for a certain dataset, checking how many images are discarded changing the radius of the ROI, the hash similarity and the feature matching thresholds, evidence visually that the discarded *candidates* have indeed an important overlap with the query, and verify the quality, coherency and reliability of the resulting mosaic. Some datasets used to assess the procedure are lawn-mower shaped surveys conducted in the south coast of Mallorca with an AUV model SPARUS II [6] equipped with an IMU, a DVL and a stereo rig, at depths between 9 and 15 meters, and tentative navigation altitudes between 3 and 4 meters, at constant lineal speed between 0.5 and 1 knots, mostly is areas where the sea bottom was densely colonized with sea grass. Table 1 shows some results of one mission: a relation of number of discarded images over the total number of images that would have been stored for the mosaic in case the double layer filter had not been applied, depending on the radius of the ROI, the hash similarity and the features matching mask. Ongoing work includes to verify the adequacy of the discarded frames and to generate sets of true and false positives and negatives that will be used to calculate the quality metrics of the process. Figure 1 shows the mosaic corresponding to combination 6 (left), which is not the one with more discards, but it turns out to be the most reliable in terms of minimum repeated structures and maximum similarity with the real environment. Figure 1 also shows two images that overlap (one discarded) nearly the 70%, with a hash similarity of 35.54 and 200 feature matchings.

Combination	**Radius (m)**	**Hash similarity threshold**	**Feature matchings threshold**	**Total N_{br} of images**	**N_{br} of Discarded Images**
1	2.5	60	50	1265	664
2	2.5	60	60	1265	490
3	2.5	50	50	1265	543
4	2.5	50	60	1265	430
5	3	60	50	1265	709
6	3	60	60	1265	560
7	3	50	50	1265	622
8	3	50	60	1265	464

Tab 1. Discard metrics

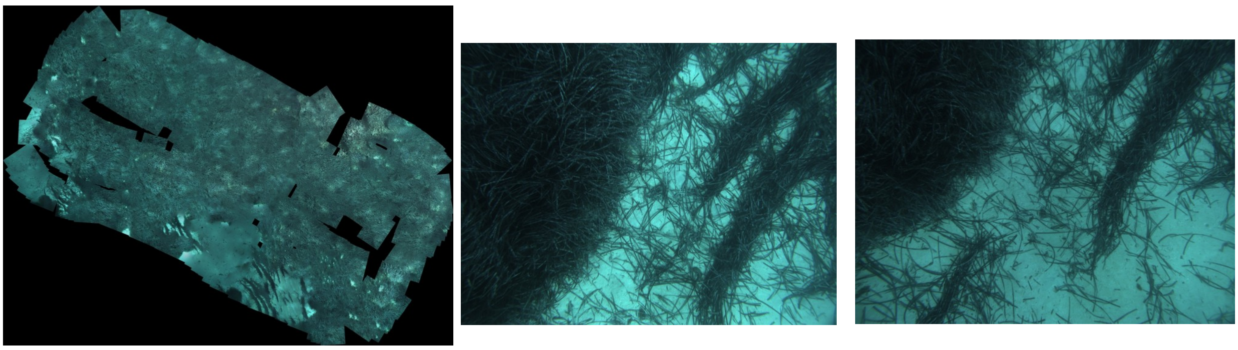

Fig 1. Best mosaic for the sample dataset and two images that overlap nearly a 70%

II ACKNOWLEDGEMENTS

This work is partially supported by by Grant PID2020-115332RB-C33 funded by MCIN/AEI/10.13039/501100011033, by "ERDF A way of making Europe" and by Grant PLEC2021-007525/AEI/10.13039/501100011033 funded by the Agencia Estatal de Investigación, under Next Generation EU/PRTR.

REFERENCES

[1] A. Elibol, J. Kim, N. Gracias and R. Garcia, "Fast Underwater Image Mosaicing through Submapping", *Journal of Intelligent and Robotic Systems*, vol. 85, pp. 167-187, 2017.
[2] F. Bonin-Font, J. Lalucat, G. Oliver, M. Massot, E. Guerrero, P. L. Negre. Evaluating the Impact of Sewage Discharges on the Marine Environment with a Lightweight AUV. In *Marine Pollution Bulletin,* Elsevier, vol. 135, pp. 714-722, October, 2018.
[3] Stanford Artificial Intelligence Laboratory et al. (2022). Robotic Operating System. URL https://www.ros.org.; http://wiki.ros.org/Bags
[4] Lowe, David G. (2004). "Distinctive Image Features from Scale-Invariant Keypoints". *International Journal of Computer Vision*. 60 (2): 91–110
[5] P. L. Negre, F. Bonin-Font, G. Oliver. Global Image Signature for Visual Loop-Closure Detection. In *Autonomous Robots,* Springer, vol. 40, pp. 1403-1417, December, 2016
[6] M. Carreras, J. D. Hernández, E. Vidal, N. Palomeras, D. Ribas and P. Ridao, "Sparus II AUV—A Hovering Vehicle for Seabed Inspection," in *IEEE Journal of Oceanic Engineering*, vol. 43, no. 2, pp. 344-355, April 2018.