



**HAL**  
open science

## Multi-session lake-shore monitoring in visually challenging conditions

Cédric Pradalier, Stéphanie Aravecchia, François Pomerleau

► **To cite this version:**

Cédric Pradalier, Stéphanie Aravecchia, François Pomerleau. Multi-session lake-shore monitoring in visually challenging conditions. 2020. hal-01867104v2

**HAL Id: hal-01867104**

**<https://hal.science/hal-01867104v2>**

Preprint submitted on 2 Jul 2020

**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.

# Multi-session lake-shore monitoring in visually challenging conditions

Cédric Pradalier, Stéphanie Aravecchia and Francois Pomerleau

**Abstract** Long-term monitoring of natural environments raises significant challenges due to the strong perceptual aliasing in trees, bushes and shrubs. This paper reports on the multi-session localization and mapping of a small lake shore using an autonomous surface vessel equipped with a 2D lidar and a camera. Our publicly available dataset includes 130 autonomous surveys of the 1 km shoreline while recording lidar, GPS and image data. We build our globally consistent multi-session map using ICP at multiple scale. The end result is evaluated qualitatively by superimposing all the lidar maps, and quantitatively by comparing images taken from the same pose at different times. The localization and mapping results, as well as the dataset of image pairs, are made available within our public dataset.

## 1 Introduction

This paper presents the results of a large-scale experiment on multi-session monitoring of a natural environment using a 2D laser scanner on-board an autonomous surface vessel. Our dataset (Symphony Lake Dataset [8]) was recorded over four years by autonomously circumnavigating a small reservoir, named Lake Symphony, in Metz France (see Fig. 1). A total of 130 surveys were realized between January 1<sup>st</sup> 2014 and November 1<sup>st</sup>, 2017, corresponding to approximately 130 km of autonomous operations. During these surveys, the vessel recorded 2D laser scans, side looking images, GPS and compass. From these data, our purpose is to build a consistent map of the lake shore over time.

---

Cédric Pradalier, Stéphanie Aravecchia  
GeorgiaTech Lorraine-UMI 2958, GeorgiaTech-CNRS, Metz, France,  
e-mail: [firstname.lastname@georgiatech-metz.fr](mailto:firstname.lastname@georgiatech-metz.fr)

Francois Pomerleau  
Laval University, Quebec, Canada

At a time where most mapping work are focused on 3D data (point clouds acquired from RGB-D sensors or multi-beam lidars) or visual(-inertial) SLAM, the originality of our 2D-mapping stems from the type of environment we are considering for multi-session mapping. The lake-shore we are considering consists mostly of trees, reeds, shrubs and lawn, with minimal structure and no corner or straight lines. Also, because we are observing trees and shrubs from a distance of 10 m, the laser beams stochastically penetrate the environment making it semi-transparent from the laser point of view. Regarding lawn, the low ground gradient makes for grazing laser beams, which means that small variations in height or attitude of the sensor lead to large measurement variations. This is amplified by the fact that this lake is a flood buffer so the water level can vary by up to two meters in a short amount of time. These conditions, combined with the large time span covered by our dataset, makes for a specific set of challenges that will be discussed in this paper.

Because we deal with a 2D laser in a natural environment, we do not enjoy the luxury of reliable feature detectors and descriptors. As a result, this work relies on matching and aligning 2D point clouds, from raw sensor measurements to local accumulated maps (keyframes) or global maps. This is the domain of Iterative Closest Point (ICP)[14]. In its principle, the ICP algorithm takes as an input a pair of point clouds and iteratively estimates the geometric transformation that project the first point cloud to the second one. The basic assumption is that the two point clouds have a reasonable overlap, the environment is not changing too much between the two scans and a reasonable estimate of the transformation is available to start from. The latter is particularly important if the two former assumptions cannot be guaranteed. In the right conditions, ICP can be a very fast algorithm, suitable for real-time. This paper relies on the implementation available as `libpointmatcher` [15].

The contribution presented in this paper are three-fold. First, we present a method for large scale multi-survey lake-shore mapping using 2D laser and we evaluate it on the publicly available Lake Symphony Dataset [8]. Second, the localization and 2D environment maps are added to the public dataset as a globally consistent baseline localization. Third, combining the localization result with the visual part of the dataset, we propose a 16k-image pair dataset for place recognition and change detection in a natural environment under significant seasonal changes.

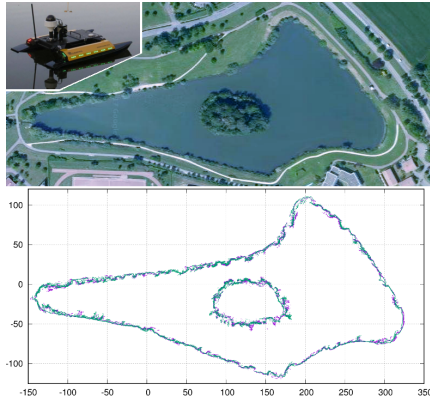


Fig. 1: Overhead imagery of the environment considered in this paper. *Insert*: Kingfisher robot. *Bottom*: Resulting lidar map (in meters).

## 2 Related Work

Our team initial work on the Symphony Lake Dataset has been presented in [9]. So far, our objective has been to focus on the visual data to find similar views of the lake shore between different surveys and use image deformations to align them. First, a 3D model was constructed using visual-SLAM assisted with GPS for scale consistency. Second, images were aligned using first SiftFlow [12] and finally our own Reprojection Flow [9]. Initial results have been promising, but getting consistently good alignments beyond four to six months has proven to be a very significant challenge. It has not yet been possible to bring the full 130 surveys into alignment with these methods. A review of lake shore monitoring can be found in [9]. [17] also addresses this problem but focuses on real-time short-term visual odometry from lake-shore images.

More recently, a review of the performance of off-the-shelf visual-SLAM solutions on this dataset was conducted in [3]. This showed that dense approaches relying on the minimization of photogrammetric errors are providing the best results on the very specific images of this dataset. In particular, DSO [6] proved to be able to build good quality maps on some sub-sequences of some of the surveys, mostly in summer where the semi-transparency is not so marked. An extension of DSO has been presented in [18] to use multiple baseline and make it more robust to the conditions observed in natural environment datasets. All these works have shown that single-survey visual mapping on this dataset is a very significant challenge and that we are still far from the large scale multi-survey mapping we are aspiring to. In comparison, this work use a 2D lidar, which is in general a much less challenging sensor, but focuses on building the necessary robustness to work on 130 surveys, from summer to winter, flood to drought, including interaction with local wild-life, fishermen and low-level mechanical issues. For surveying purposes, images will then be compared based on the pose computed from the laser data.

Laser-based solutions for multi-session mapping was investigated before using `libpointmatcher`, but in urban environments [16]. Only three sessions during a year were used to investigate the impact of dynamic elements on the representation of the surrounding. Application of the same algorithm was demonstrated for shore mapping [10], but only with few surveys. This paper extend on the same idea, but push the extent of the evaluation using two order of magnitude more surveys than prior work.

Finally, our approach relies on factor-graphs [4] and specific approaches for multi-session SLAM have been explored extensively (e.g. [13]). Similarly to those, we build an extensive network of relations between the surveys and include them in the factor-graph optimization. However, our evaluation is the only one that focuses on a mostly natural environment dominated with vegetation.

### 3 Method

At a high level, our method is a three-stage process as showed in fig. 2. Initially, every survey is processed to build an accumulated shore map as a trajectory and point cloud split into local keyframes (Section 3.1 and 3.2). In a second stage (Section 3.3), these survey maps are put in a common reference frame and at the third stage (Section 3.4 and 3.5), a fine-grained alignment is used to deform the individual maps into alignment. The following sections will provide details on every stages of this process as well as our evaluation method.

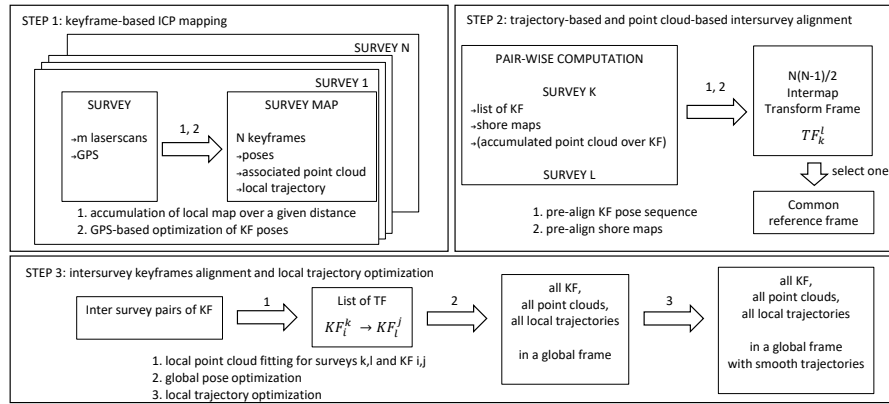


Fig. 2: Overview of our method.

#### 3.1 Intra-survey keyframe creation

The first step of our long-term mapping solution is the acquisition of local maps using a 2D lidar. This has been demonstrated many time in the past using lidar and ICP. In this case, ICP is applied between one laser scan and the next to estimate the 2D displacement between the laser scans. Contrary to approaches such as [2], our framework [15] is implementing simultaneous localization and mapping, hence accumulating lidar points into a local map. One of the challenges of boat-mounted lidar mapping is that it is extremely hard to measure odometry (i.e., local displacement with respect to the water surface) on a slow-moving vessel. When there are enough features in the laser scans, odometry is not really needed and laser odometry works precisely. However, in our natural settings and given the large variation in conditions we are considering, laser odometry fails occasionally and another odometry source can be used to bridge the gap until the lidar sees enough of the environment to recover. We resort to the simple expedient of using GPS information

as an estimate of position. At our speed (lower than 0.5 m/s), GPS-based velocity and heading estimation are rather unreliable. As a result, when it is available, we complement the GPS information with heading estimated from a low-cost compass.

Regarding the mapping process, we use the dynamic mapper presented in [16]. Its main advantage, beyond being a mapping framework and not an odometry as [2], is that it is able to estimate the dynamicity of laser points, hence focusing the alignment and localization on the most reliable and stable points in the environment. However, a lake is a specific environment because of its size with respect to the lidar perception. Lidar returns tend to come from only a small section of the shore on the side of the surveying vessel, and similarly to the “infinite” corridors from [1], the trajectory and in particular the vessel heading estimate tend to drift if no globally consistent correction is applied. This correction will be presented in the next section. However, to make sure the effect of this drift does not impact the quality of the final map, we resorted to the common solution of separating the map into keyframes. This keyframes correspond to the accumulation of a local map over a number of laser scan or a given distance. These cut-off point are chosen to make sure that the orientation drift within this window is minimal, guaranteeing that the map is locally structurally consistent within a keyframes. This is important from a computational perspective because all the laser scans that have been involved in the creation of a keyframes can be forgotten. All their information is considered summarized in the keyframe map and the latter can be considered immune to the local trajectory adjustments in the various step of optimization that will be presented in the following sections.

The output of this intra-survey mapping is, for each survey, a vessel trajectory, a list of keyframe coordinates and for each keyframe an associated point cloud. Combining the point clouds from all the keyframe leads to an environment map as shown in Fig. 3.

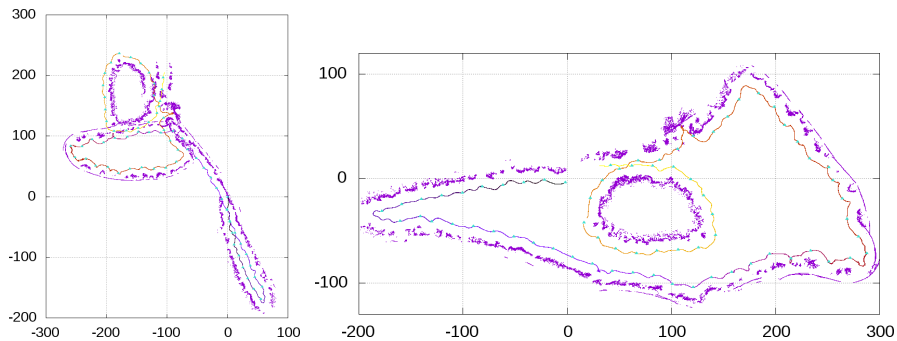


Fig. 3: ICP maps from the survey on February 2<sup>nd</sup>, 2016, axes in meters. *Left*: map without using GPS correction, which is locally accurate but accumulates drift over the 1 km path. *Right*: map with GPS correction showing global consistency because of the use of ICP and factor-graph optimization. The colored line is the trajectory, colored as a function of time.

### 3.2 Drift control from GPS and Factor-Graphs

Despite its local precision, lidar-based odometry tends to slowly drift when not provided with absolute measurements acting as anchors. This is illustrated in Fig. 3. Two types of measurements can play this role: either absolute information such as GPS or compass, or loop-closure detection. Our solution is built on a factor-graph [4] implementation using G2O [11]. The optimized variables are the poses of the keyframes ( $SE(2)$ ). The absolute observations are the GPS measurement ( $\mathbb{R}^2$ ), complemented with the compass observation to make it a full pose in  $SE(2)$ . The inter-keyframe transformations from lidar odometry ( $SE(2)$ ) are used to link the estimated poses. These are estimated a first time when a new keyframe is created by recording the position of the boat in its original keyframe at the time of the transition. The transformation is then refined when a keyframe is completed by using ICP between the keyframe point cloud and the point cloud from the previous keyframe.

An example of the resulting map is shown in Fig. 3 *Right*, while Fig. 4 gives an impression of all the maps generated at this stage.

In practice a new keyframe starts every 20 m and this optimization is fast enough to be run in real-time in the background every time a new keyframe is inserted. Because it is running continuously, most of the graph is already stabilized and G2O only requires 3-5 iterations to converge after each keyframe.

The output of this stage is the same as before but the keyframe poses have now been adjusted to account for the absolute measurement. Note that we make the assumption that the keyframes are adjusted by small increments which are not visible at the scale of the point cloud inside a keyframe. Hence, the adjustment result is not used to refine the local point cloud and the recorded trajectory inside the keyframe.

### 3.3 Global map alignment

From the previous sections, we have now constructed a set of 130 lake-shore maps (Fig. 4), all of them internally consistent up to the trade-off between the GPS and lidar-ICP precisions. It can be noted that some maps are incomplete. This is the results of the survey ending before a full circumnavigation, typically due to hardware errors. Our ultimate goal is to align all of these to a level where we can constitute pairs of images viewing the lake shore from the same pose. To this end, we intend to build a massive factor-graph optimization problem where the keyframes in all of the 130 surveys are linked with their counterparts in other surveys with estimate of the transform between the keyframes estimated by ICP between the keyframe point clouds. The detail of this process will be presented in Section 3.4. However, because the keyframe point clouds have been acquired at different time of the year, sometimes with more than three years of interval, and because the keyframes may have a minimal overlap, ICP does not converge reliably unless it has a reasonable initial estimate. This section presents how this estimate is obtained.

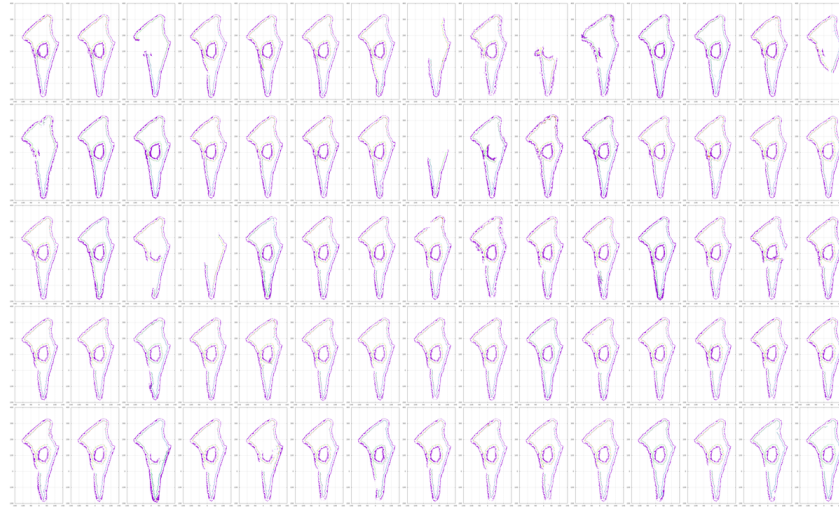


Fig. 4: All maps and trajectories after intra-survey ICP and GPS-based map stabilization

The map obtained from intra-survey mapping ([Section 3.2](#)) are optimized to be close to the GPS measurements, assuming a 10 m uncertainty on GPS. Hence, to build an initial estimate of keyframe pose, we need to estimate the precise offset between the maps from two surveys. Note that, because of the GPS constraint and the size of the environment we are considering, the optimization gives very little freedom in rotation and the map offset is mostly a translation. To obtain this offset, we again use a two-stage process: first we use ICP to align the boat trajectories considered as point clouds and starting from this offset estimate, we try to align the map point clouds. It is important to remember that the global maps are obtained by agglomerating the point clouds from all the keyframes so they may be large and noisy because of the natural environment. Running ICP directly on these maps, without the trajectory alignment, works sometimes but lacks robustness and precision. Nevertheless, the boat is running an approximate model-predictive control system to follow the lake shore at a fixed distance despite its significant dynamic. This means that trajectories tend to be very similar between surveys and very smooth. We take advantage of this by running a first alignment between the trajectories using ICP. We can then use the resulting transformation between trajectories as initial estimate for running ICP between the global point clouds.

After this level of alignment of the surveys, the map point clouds are globally aligned but they may have a section where the map orientation drifted and was not corrected by the GPS measurements. [Fig. 5](#) shows aligned keyframe trajectory and point clouds, but despite the global alignment local drift is still visible.



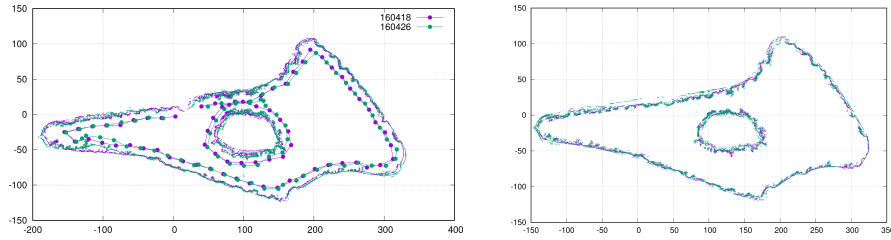


Fig. 5: Left: visualization of the keyframe trajectory and point clouds for two surveys (i.e., map 160418 and 160426) after global alignment. Right: Visualization of map 140106 (in cyan) compared with 171030 (in purple) after inter-survey map alignment. The lines on the top part of 140106 are shore lines visible because of the high water level on this day.

At this end of this stage, we have obtained all our keyframes in a common reference frame where they can be compared. The trajectory and the point clouds have not been modified.

### 3.4 Iterative Multi-Session SLAM

At this stage, all the keyframes from all the surveys are now in a common reference frame. In order to align all the surveys and build a common map combining information from all the surveys, we will resort once again to a large factor-graph structure. The variables to be optimized will be the keyframe poses in the common frame. The intra-survey information will provide estimate of the transform between successive keyframes within a survey, as well as GPS estimates. What remains to be estimated are transforms between pairs of keyframes in different surveys. Non-consecutive pairs from a given survey are also considered here. They correspond to situations where the boat revisit the same place at two different times in the survey.

To estimate the transform between a pair of keyframes, we will take advantage of our ICP framework and try to align the keyframe point clouds. To achieve a satisfactory robustness, we need to be very strict and require at least 70 % of overlap between the aligned point clouds. In our setup, we have 130 surveys, each containing of the order of 80 keyframes, resulting in approximately  $10^4$  poses to optimize, and naively  $10^8$  ICP pairs to consider.

In order to make the estimation of ICP links tractable, we take advantage of the common reference frame and for each keyframe we use a nearest neighbor search to find keyframes on which point clouds we can apply ICP. Note that all these ICP instance are run using the estimated keyframe pose in the common reference frame as initial guess of the transformation.

Once all the keyframe links have been added to the factor-graph problem, the optimization can be driven to its minimum. This optimization locally deforms the

map from every survey to match all the constraints from intra- and inter-survey ICP matches while accounting for the GPS measurements. The tuning of the relative weights is tricky here but we consider that the intra-survey transformation from consecutive frames are the most reliable, inter-survey transformations are a bit less reliable but still much better than GPS.

However, because the survey map are locally modified, these deformations can lead to a new set of neighbors to every keyframe. Among these new neighbors may be some for which the transformation can be estimated by ICP. To account for this new information, the neighbor search and optimization are run iteratively: using the new pose estimates, a new KD-tree[5] is built, out of which new set of neighbors are extracted. ICP is then applied to those that have not been considered yet and a new factor-graph problem is built and optimized. The iterations stop when no new neighbors can be found. In our setup, every iteration adds approximately 10 % of the number of neighbors found by the previous iteration, leading to a bit more than  $10^7$  ICP evaluation among which  $1.4 \times 10^6$  are found to have more than 70 % of overlap and are used in the final factor-graph problem.

Fig. 5 shows an example of pairs of aligned maps reconstructed from the optimized keyframes. Fig. 8 shows the consistency of the reconstruction in a map containing all the point clouds from the 130 surveys.

The output of this stage is the optimized poses of all the keyframes in a common reference frame, as well as the point cloud map resulting from every surveys.

### 3.5 Local trajectory optimization

As mentioned earlier, the optimization stages from the previous sections only affect the keyframe poses but do not adjust the boat trajectory within the keyframes. In this section, we detail how the trajectory can be adjusted to make sure it smoothly goes from the origin of a keyframe to the origin of the next one.

The ICP-based odometry from Section 3.1 provides a set of poses originating at the origin of a keyframe and reaching the next keyframe. These are the variables we intend to optimize in this section. In addition to making sure that the end of the local trajectory reach the next keyframe, we also take the opportunity of smoothing the trajectory during this stage. The lidar poses are acquired at 40 Hz on a vessel moving at most at 0.3 m/s, so there is less than 1 cm of displacement between lidar scans. The boat has very little control authority in rotation due to the water resistance and propeller arrangement so the rotation between lidar scans is rarely larger than 0.01 rad. Despite these mechanical constraints, lidar measurement noise and in particular the semi-transparency of the environment, make the ICP somewhat noisy at the 5-10 cm scale.

To smooth the ICP trajectory, we took inspiration from [7] and decided to first approximate the local trajectory by a cubic B-spline with regularly sampled control points every 1 s (i.e., consistent with the low dynamic of our system). The B-Spline approximation can be conducted on every *SE2* pose parameters independently and

then a linear correction can be added to make sure the end pose reach the next keyframe. Fig. 6 illustrates this process. Note that because these are the trajectories of a boat instead of a wheeled vehicle, the boat heading is not necessarily exactly aligned with the local tangent to the trajectory. At the end of this stage, we now have all the keyframes, point cloud maps and boat poses in a globally consistent frame with smooth trajectories.

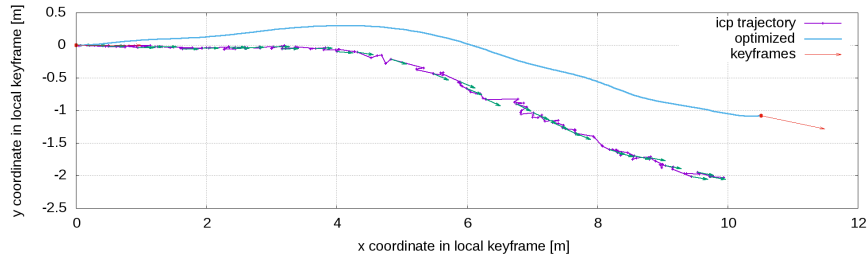


Fig. 6: Example of local trajectory optimization between two keyframes identified by the red dot and arrows. The purple sequence of pose is characteristic of unfiltered ICP result while mapping.

### 3.6 Image pair selection

One use of the trajectories from the multi-session localization and mapping framework described above is to select images looking at the same shore point at different times. From a computer vision point of view, it also helps if the images are taken from the same viewpoint or at least from the same optical axis in order to reduce the effect of parallax. Given two optimized and smooth trajectories, it is straightforward to select pair of nearest positions. Intersecting the optical axis of a camera with the trajectory in the second survey gives us neighboring poses with near identical optical axes. It is then possible to use a homography to virtually rotate the two viewpoints so as to make the image planes parallel.

An example of the resulting images can be seen in Fig. 7. The perspective transformation of the selected image is made visible by the black pixels which were outside of the camera field of view in the untransformed images.



Fig. 7: Matching images from different days. *Left*: image of reference taken in January 6<sup>th</sup>, 2014. *Middle and Right*: matched and rotated images taken respectively on June 13<sup>th</sup> and June 25<sup>th</sup>, 2014.

## 4 Results and experiments

### 4.1 Global map alignment

The main challenge of multi-session mapping over a large time span is to maintain a consistent map. We illustrate the performance of our approach by superimposing the point clouds of a map pair in Fig. 5 and of all the 130 maps in Fig. 8. As can be observed, the local details are clearly represented and aligned precisely in the pairwise figure. With that many maps, the level of noise increase significantly. A lot of dynamic features in open water can be attributed to wild-life interacting with the boat. On the shore, the noise level is due to uncertainty in the ICP output combined with change of appearance resulting from variation in the water level.

Our ICP library (`libpointmatcher`) also estimates if a point is dynamic or static based on how often a laser beam goes “through” it (see [16] for details). In our particular environment, most points belongs to reeds, bushes or shrubs which are semi-transparent from the laser point of view. As a result of this complexity (often resulting in “ghost” points), most points are considered dynamic using the method cited above. When plotting the histogram of the log likelihood of a point being dynamic, we observe a distribution between  $-4$  and  $-14$ . Selecting only the point with a dynamic likelihood below  $10^{-9}$  results in a qualitative map improvement depicted in Fig. 8. The thinness of the shore lines across seasons highlight the precision of the performance of the iterative global alignment proposed in this paper.

### 4.2 Image pairs dataset

As mentioned earlier, the Symphony Lake Dataset was not acquired with a high precision localization system, so there is no direct ground truth to compare our multi-session SLAM results to. In order to provide a quantitative metric anyway, we elected to work on the lake shore images. To this end, we built a dataset of image pairs by selecting randomly a pair of surveys, a reference view in the first

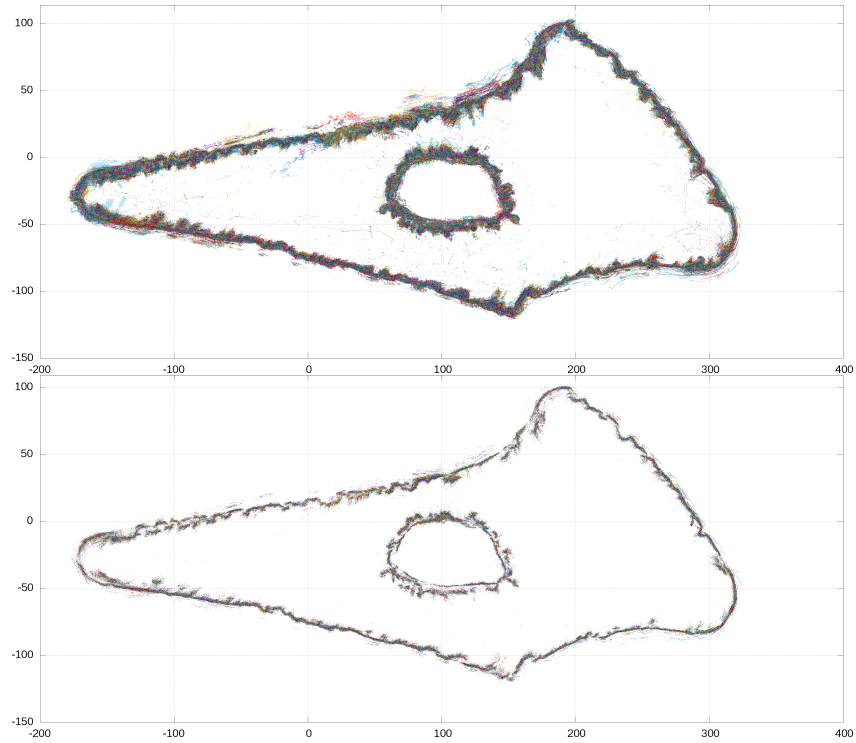


Fig. 8: Final survey combining all maps. *Top*: unfiltered point clouds. *Bottom*: filtered maps keeping only points with a low likelihood of dynamicity. The consistency and level of details is particularly visible on the island where the visible shoreline is mostly vegetation.

survey and then using the approach presented in [Section 3.6](#) to find the corresponding rotation-compensated image in the second survey. The two resulting images are then presented to a user who needs to click on a point in each image representing the same physical location in the world. Because the work presented here is only concerned with 2D, we can get an estimate of the alignment error by comparing the horizontal coordinates of the clicked points. With a perfectly consistent pose across surveys, and neglecting translation inaccuracies, the horizontal coordinates of physical objects should not change after compensating for the boat rotation with the homography. Note that the pixel coordinate error can be converted to angular error using the camera intrinsics. For reference, our camera has 704 pixels to cover a  $42^\circ$  field of view.

The histogram of the distribution of the alignment error is depicted in [Fig. 9](#). This was computed by hand-labeling 265 image pairs, out of which four were considered too uncertain to be included. These are images with very heavy seasonal changes, lighting changes or sun glares, for which the authors were not able to iden-

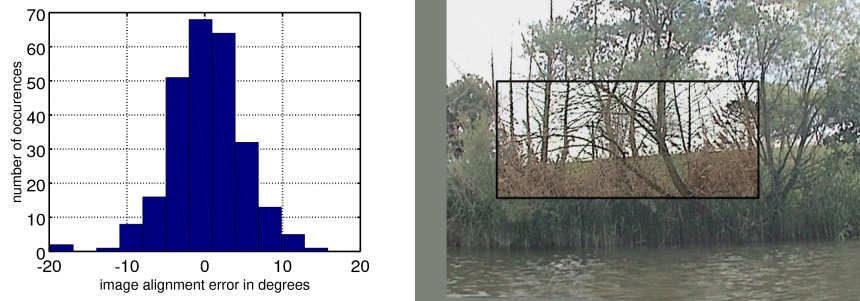


Fig. 9: Left: Image alignment error distribution in degree for 265 random image pairs across multiple seasons. Right: Example of a perfect alignment with an automatically generated winter insert on a rotation-compensated summer image.

tify precisely a matching physical place. This does not necessarily mean that the alignment is incorrect so these four images were removed from the statistics. Out of the remaining 261 images, the median error is  $0.00^\circ$ , the mean  $-0.25^\circ$  with a standard deviation of  $4.7^\circ$ . Overall, 90% of the pairs exhibit an absolute alignment error lower than  $8^\circ$ . The composite picture in Fig. 9 shows how a high quality of localization can be used to highlight seasonal changes in this natural environment dataset.

Because of the significant challenges observed when hand-labeling these image pairs, a dataset of 16k image pairs generated using the above approach has been added to the Symphony Lake Dataset[8] web page<sup>1</sup>. This dataset covers a very wide range of appearance change of this natural environment over a span of four years and it could be particularly useful to evaluate place recognition algorithms. Additionally, after considering the image alignment precision evaluated in this section, the maps and boat trajectories have also been added to the Symphony Lake Dataset web page and can be used as a reference localization system by others interested in working with this challenging dataset.

## 5 Conclusion

In this paper, we presented our approach to large scale multi-session SLAM using a 2D laser and ICP in a natural environment. The end-to-end approach uses ICP at multiple scales from local keyframe construction from individual laser scans to inter-survey keyframe alignment. The end results produce a set of globally consistent 2D maps and trajectories which are precise enough to select rotation-compensated images of the lake shore over multiple seasons. These results have been used to add reference maps and localization to the publicly available Lake

<sup>1</sup> <http://dream.georgiatech-metz.fr/?q=node/76>

Symphony Dataset as well as to create a new dataset of pairs of images of the lake shore that may be used for the evaluation and training of natural place recognition algorithms. From the stepping stone presented in this paper, we intend to keep exploring Visual-SLAM algorithms that can work with our images. Automated image alignment (without laser), change detection and change segmentation are also topics that may be investigated with the help of this visually challenging environment monitoring dataset.

## References

1. Bosse, M.C.: ATLAS: a framework for large scale automated mapping and localization. Ph.D. thesis, Massachusetts Institute of Technology (2004)
2. Censi, A.: An ICP variant using a point-to-line metric. In: Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on, pp. 19–25. IEEE (2008)
3. Chahine, G., Pradalier, C.: Survey of monocular SLAM algorithms in natural environments. In: Conference on Computer and Robot Vision (2018)
4. Dellaert, F., Kaess, M., et al.: Factor graphs for robot perception. *Foundations and Trends® in Robotics* **6**(1-2), 1–139 (2017)
5. Elseberg, J., Magnenat, S., Siegwart, R., Nüchter, A.: Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration. *Journal of Software Engineering for Robotics (JOSER)* **3**(1), 2–12 (2012)
6. Engel, J., Koltun, V., Cremers, D.: Direct sparse odometry. *IEEE transactions on pattern analysis and machine intelligence* **4** (2017)
7. Furgale, P., Tong, C.H., Barfoot, T.D., Sibley, G.: Continuous-time batch trajectory estimation using temporal basis functions. *The International Journal of Robotics Research* **34**(14), 1688–1710 (2015)
8. Griffith, S., Chahine, G., Pradalier, C.: Symphony lake dataset. *International Journal of Robotics Research* **36**, 1151–1158 (2017)
9. Griffith, S., Pradalier, C.: Survey registration for long-term lakeshore monitoring. *Journal of Field Robotics* (2016)
10. Hitz, G., Pomerleau, F., Colas, F., Siegwart, R.: State estimation for shore monitoring using an autonomous surface vessel. In: *Experimental Robotics*, pp. 745–760. Springer (2016)
11. Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: G2O: A general framework for graph optimization. In: Robotics and Automation (ICRA), 2011 IEEE International Conference on, pp. 3607–3613. IEEE (2011)
12. Liu, C., Yuen, J., Torralba, A.: SIFT Flow: Dense correspondence across scenes and its applications. *PAMI* **33**(5), 978–994 (2011)
13. Mühlfellner, P., Bürki, M., Bosse, M., Derendarz, W., Philippsen, R., Furgale, P.: Summary maps for lifelong visual localization. *Journal of Field Robotics* (2015)
14. Pomerleau, F.: Applied registration for robotics: Methodology and tools for ICP-like algorithms. Ph.D. thesis, ETH Zurich (2013)
15. Pomerleau, F., Colas, F., Siegwart, R., Magnenat, S.: Comparing ICP variants on real-world data sets. *Autonomous Robots* **34**(3), 133–148 (2013)
16. Pomerleau, F., Krüsi, P., Colas, F., Furgale, P., Siegwart, R.: Long-term 3D map maintenance in dynamic environments. In: Robotics and Automation (ICRA), 2014 IEEE International Conference on, pp. 3712–3719. IEEE (2014)
17. Terzakis, G., Polvara, R., Sharma, S., Culverhouse, P., Sutton, R.: Monocular visual odometry for an unmanned sea-surface vehicle. *arXiv preprint arXiv:1707.04444* (2017)
18. Wu, X., Pradalier, C.: Multi-scale direct sparse visual odometry for large-scale natural environment. In: *International Conference on 3D Vision* (2018)