Journal Title
XX(X):1–7

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Symphony Lake Dataset

Shane Griffith[1, 3], Georges Chahine[2,3], and Cédric Pradalier[1,3]

Abstract

This paper describes Symphony Lake Dataset, 121 visual surveys of an approx. 1.3 km lakeshore in Metz, France. Different from roadway datasets, it adds breadth to the space at a time when larger and more diverse datasets are desired. Over 5 million images from an unmanned surface vehicle captured the natural environment as it evolved over three years. Variation in appearance across weeks, seasons, and years is significant. Success on Symphony Lake Dataset could demonstrate advancements in perception, SLAM, and environment monitoring.

Keywords

dataset, unmanned surface vehicle, visual survey, lakeshore, natural environment

Introduction

A growing homogenous space of publicly available robotic vision datasets capture a roadway from a car, which the release of Symphony Lake Dataset can help to diversify. Interest in creating autonomous driving vehicles has contributed to the growth and availability of roadway data. Work on perception has benefitted from the fact that these images are captured outdoors, and sometimes over long–term time periods (e.g., Maddern et al. 2017; Geiger et al. 2013). Yet, a roadway is highly structured, which could simplify perception and lead to non-general algorithms. A long–term dataset of a large–scale natural environment would add breadth to this space to advance research in perception.

Simultaneously, advancements in deep learning have generated interest in massive datasets. Baseline performance in tasks like scene classification improve with the amount and the diversity of the training data (Zhou et al. 2016; Chen et al. 2017). With millions of exemplars, some basic labeling tasks have reached nearly human–level performance, while some advanced game AI have surpassed the best humans. Advancements seem to come in parallel with the availability of data. The release of Symphony Lake Dataset may contribute to this growth in results for perception in natural environments.

This paper proposes the release of Symphony Lake Dataset, 121 visual surveys of the shore and the island of Symphony Lake in Metz, France. The 1.3 km shore was surveyed using a pan–tilt–zoom (PTZ) camera mounted on an unmanned surface vehicle (see Fig. 1). The camera faced starboard (or port for the island) as the boat moved in parallel with the shore. We deployed the boat on average every 10 days from Jan 6, 2014 to April 3, 2017. Over 5 million images were captured.

The 600 GB dataset is released in two sets: 1) 4 GB full surveys and 2) 200 MB sub-sampled surveys. The surveys include GPS, IMU, and compass data, which are synchronized to the 704 × 480 @ 10 fps color images.

Figure 1. Our Clearpath Kingfisher as it circled the perimeter of Symphony Lake.

Readings from the 2D LiDAR are also included. Each survey is available for individual download on a dedicated website at `dream.georgiatech-metz.fr/?q=node/76`.

This paper uses a similar structure to Maddern et al. (2017) (with permission), which is an archetypal robotics dataset paper. Their autonomous–capable car captured a suburban neighborhood in Oxford, UK twice a week, on average, for over a year. The full view of street scenes around their vehicle is captured in 3D LiDAR and image data (among data from other sensors). In contrast, this paper captures a natural environment week-to-week as it evolved over three years and the data consists primarily of side-view images.

Platform

Our platform is the Kingfisher M200 unmanned surface vehicle (USV) from Clearpath Robotics (see Fig. 2). The

[1]College of Computing, Georgia Institute of Technology, Atlanta, USA
[2]College of Engineering, Georgia Institute of Technology, Atlanta, USA
[3]GeorgiaTech Lorraine, Metz, France

Corresponding author:
Cédric Pradalier, GeorgiaTech Lorraine, Metz, France
Email: cedric.pradalier@georgiatech-metz.fr

Figure 2. a) Front view of our Clearpath Kingfisher and side views of **b)** the PTZ camera, and **c)** the 2D LiDAR. A survey is initiated using a computer connected via the wifi antennae in the back. As the motor in each pontoon propels the robot, the camera pans starboard (or port for the island). The laser range–finder measures the ranges to obstacles, which are used to maintain a 10m distance to the shore. The GPS, the compass, and the IMU measure trajectory values while the computer inside the waterproof compartment records all of it.

USV has the style of a pontoon–boat. A 0.55 m × 0.80 m metal base connects the top of two 1.3 m-long pontoons. The back of each pontoon houses a jet thruster, which propels the boat up to 1.7 m/s. A power differential between the motors turns it.

A 40 Ah nickel–metal hydride battery powers the USV. The battery is secured inside a compartment in the metal base before each survey. It has enough charge to move the boat at nearly 0.35 m/s for over an hour. While stationary it can power the sensors and the onboard computer for up to 10 hours.

Our USV is equipped with four primary sensors:

PTZ Camera: Axis P5512-E. 360° Pan. 180° Tilt. 12x zoom. 704x480 @ up to 60 Hz. 3.8mm Lens. 51.6° HFoV.
2D LIDAR: SICK LMS111. 20m Range. 0.5° Resolution. 50 Hz. 270° HFoV.
GPS: U-Blox LEA-6. 5 Hz. 2.5m
IMU: CHR-UM6. 2° Pitch and roll accuracy. 5° Yaw accuracy.

The metal base of the USV has a waterproof electronics bay inside it and a platform bay on top for the sensors. The GPS and the onboard computer are housed within the electronics bay. The PTZ camera, the laser range–finder, and the IMU are mounted to the platform bay. The camera is mounted behind the laser, high enough for an unobstructed view. Because the laser range–finder is mounted facing forward, distances to objects behind the USV are not measured.

Sensors fed their data to an embedded computer. The computer has an Intel Atom Z530 CPU (1.6 GHz, 2 threads, 32-bit), 1 GB RAM, and a 16GB SanDisk SSD U100. In addition to planning the robot's motion using LiDAR data, the computer processed and stored sensor readings. The harddrive is large enough to store over a survey of data.

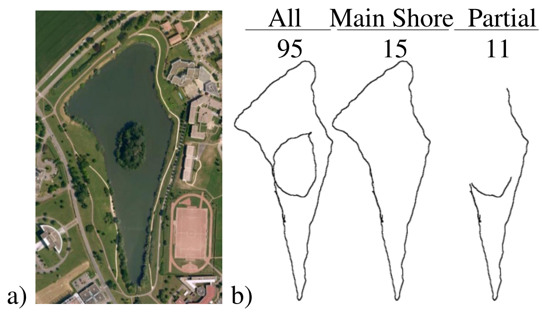

Figure 3. a) Symphony Lake from the perspective of Google Maps Satellite View, and b) depictions of the different trajectories of the robot with their number of occurrences. The boat approximately circled the entire perimeter (95 cases), missed the full island (15 cases), or otherwise partially traversed its route (11 cases). That is, *Main Shore* includes surveys with partial island coverage. *Partial* here is illustrated with one example of a partial route; each one was different.

Symphony Lake

Symphony Lake is 2 km south east of Metz, France, across the street from GeorgiaTech–Lorraine (see Fig. 3a). It is approximately 400 m at its longest point and 200 m at its widest point. The total area of the lake and its surroundings spans 6 ha. It also has an 80 m-wide island in the middle. The lakeshore perimeter, including the island, is about 1.3 km.

The lake was created in 1986 to prevent floods in Metz. One main inlet and a single outlet control the flow of water down a creek. During periods of heavy rain the lake's water level can increase several meters. The bank of the lakeshore is fairly steep, which keeps the water contained in the basin.

The nature of the lakeshore is varied. Some areas are surrounded with shrubs, bushes, and 20 m-tall trees. There are areas with boulders, sand, and grass. Buildings loom in much of the background. They are closer to the shore on the north east side.

The land around the lake is used to promote recreation. The grass is periodically mowed, and the other flora are sometimes trimmed or removed. A 1.35 km fitness path and a nature trail encircle the lake. Fishing, tanning, biking, jogging, and walking are common.

Behavior

The trajectory of the boat on Symphony Lake is shown in Fig. 3b. The boat was typically deployed from the west side of the shore and was pulled out at the same location after one complete run. It was sometimes pulled out at other locations in order to reset automation or to end the survey. Surveys could be started anywhere along the shore.

The USV circled roughly all of the approx. 1.3 km lakeshore, which took it nearly 70 minutes. Surveys occasionally took longer (due to e.g. wind). Several surveys captured less of the perimeter. Fifteen captured all of the main shoreline and were ended without the full island. Eleven captured parts of the main shore and the island. Rain, battery charge, control errors, and swan interference were typical limiting factors. A survey was sometimes cut short if multiple issues occurred.

A finite state machine generated the robot's trajectory. First, the USV navigates to a position 10 m from the shore and the camera pans starboard. The USV maintains its 10 m distance as it circles the perimeter in the counter–clockwise direction. The main shore survey continues until the USV crosses a virtual transition line, which extends west from the GPS position of the island's center. The boat surveys the island after it aligns itself 10 m from the island's shore and the camera pans port. The same transition line is used to shift back to the main shore survey.

The boat's trajectory was replanned at a rate of 5 Hz. A local lattice planner with a 10 m horizon provided the set of behaviors to choose from. Each one was evaluated using ranges from the LiDAR. The planner chose smooth trajectories that also kept the USV 10 m from the shore and at least 2 m away from obstacles. If the USV got closer to obstacles, however, the planner diverted its course more abruptly to avoid collisions.

We monitored the boat for the duration it was deployed, except while it circled the island. We intervened if necessary to keep the USV moving in the right direction or to completely reset automation. A control error sometimes occurred at the south end of the lake where the sharp turn caused the USV to oversteer and spin in place. The GPS position sometimes fluctuated enough that the transition to the island occurred at the wrong places. We also intervened to avoid fishing lines and to maneuver around swans.

Survey Data

Data Collection

Each survey consists of image, LiDAR, pose, and state data. There is one file per image, a set of files for the LiDAR readings, and one file of all the pose and state information. Thousands of images and LiDAR readings are saved per survey. Each 704×480 image is stored in a jpeg format with a slight compression. (Lossy compression was unavoidable due to the limited choice of formats available from the camera. The compression level was set to the minimum.) Each LiDAR reading provided 541 range measurements across the 270°arc in front of the robot. New readings were recorded at a rate of 50 Hz.

Readings from the other sensors are saved to an auxiliary file. Pose data includes the 2D position (m) from the GPS, the heading (deg) from the compass, and the angular velocity (deg/s) from the IMU. The auxiliary file has one set of pose data per line. Each line in the file corresponds to one image.

State information is saved with the pose data to guide data processing. The camera state includes its dynamic pan and tilt values, as well as its static intrinsic parameters. A pan value of approx. ± 1.57 identifies when the survey is occurring. Other values indicate transitions. A positive value indicates the USV is surveying the main shore, a negative value the island. Other information includes the time, the image number, the battery charge, and the RC controller state (i.e., whether the USV was operating in autonomous or manual control mode).

Table 1. The order of values of one line of image_auxiliary.csv.

1	timestamp	seconds	image time
2	image number		the image file index
3	UTM E	meters	GPS position in UTM 32 N
4	UTM N	meters	GPS position in UTM 32 N
5	compass	degrees	NED frame
6	camera pan	degrees	positive for starboard
7	camera tilt	degrees	
8	fx	pixels	
9	fy	pixels	
10	cx	pixels	
11	cy	pixels	
12	image width	pixels	
13	image height	pixels	
14	omega	degrees/second	IMU angular velocity
15	battery	voltage	
16	RC state	enum	1 - in range. 2 - in use.

Survey Package

The dataset is packaged according to its size. Thus, 4 GB surveys are available for individual download rather than as one large chunk. We also provide a 20x downsampled, 200 MB version for uses that require images with less overlap. The lidar data is made available in its own package. Using the May 2, 2014 survey as an example (referenced as 140502), the files for one survey are:

- *140502f.tar.gz*
  - Around 41,000 jpeg images in a hierarchical directory format with 1000 images per directory. Files are referenced using the value of an image counter. Images are numbered between *0000.jpg* to *0999.jpg* per directory, and directory names typically span approx. *0000/* to *0041/*.
  - *image_auxiliary.csv* - lines of image timestamp, image number, pose readings, and state information, in that order, with a new line of values every 0.1 s. The full list is shown in Table 1.
- *140502d.tar.gz*
  - Contains the same files as *140502f.tar.gz*, except with $1/20$ of the readings.
- *140502l.tar.gz* - a tar file of LiDAR data for a survey. LiDAR data is saved to a set of time–ordered *csv* files, each with a timestamp in the first column, followed by 541 range readings (in meters) in the following columns. Roughly 110 *csv* files per survey, each with around 1,800 scans, make a total of about $110 \times 1,800 = 200,000$ scans per survey.

To assist in the selection of surveys, a summary video for each survey is available on our website. The summary is a subset of images, taken every 1.5 m of the USV's motion, compiled into a video.

Additional Files

Additional files are included in Symphony Lake Dataset that apply to all the survey data:

- *ParseSurvey* - C++ code to interface with the survey data. The code reads an *image_auxiliary.csv* file and can provide the file paths to the images. It also converts raw sensor data into the camera pose.

Figure 4. Montage of images of one scene of the lakeshore from 118 surveys, inspired by Maddern et al. (2017). Consecutive surveys are in row–major order. This scene primarily has features from an unstructured environment, captured over three years. In the montage of Maddern et al. (2017), in contrast, the structured, street environment has some features whose appearance is more static (e.g., the sign post they used as a reference), which can simplify data association.

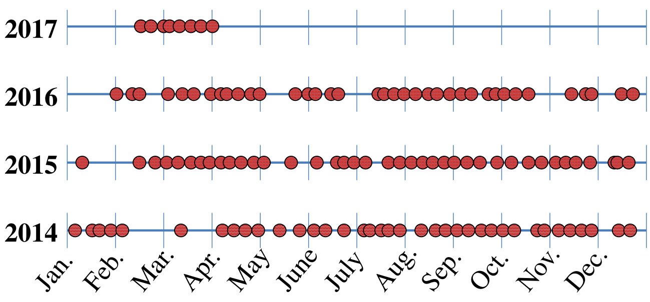

Figure 5. Timeline of surveys of Symphony Lake between Jan. 6, 2014 and Apr. 3, 2017. The lake was surveyed 37 times in 2014, 39 in 2015, 37 in 2016, and 8 (currently) in 2017.

- *camera_calibration.txt* - The full set of calibration values for the PTZ camera. A sequence of checkerboard images was used to obtain the calibration parameters.
- *sensor_positions.xls* - A spreadsheet of sensor positions for GPS, the PTZ camera, and the 2D LiDAR.
- *catalogue.xls* - A catalogue that collates survey attributes like those visible in Fig. 4. Each entry consists of survey duration, distance traveled, weather pattern, presence of noise, and more. The attributes for a survey were manually populated while viewing its summary video.

The following section characterizes the dataset using the catalogue.

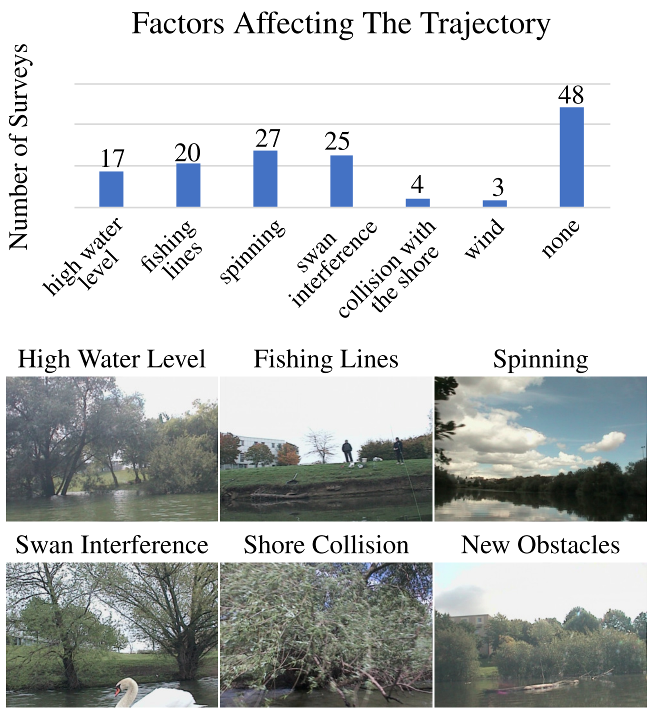

Figure 6. Factors affecting the boat's trajectory and their occurrence in the dataset. Although some factors were present throughout a survey (high water level, wind), others were localized to specific places (fishing lines, automation error, collisions, and new obstacles). A combination of these factors affected several surveys.

Dataset Characteristics

Symphony Lake Dataset has 5,031,232 images from 121 visual surveys. Figure 5 shows the timeline of surveys, which span from January 2014 to April 2017. We endeavored to deploy the robot every week, but we averaged about one survey every 10 days. Surveys were missed during weeks of heavy rains, if we were traveling, or if the lake was frozen.

Although winter surveys were captured with less regularity than surveys during other seasons, multiple surveys were captured during each seasonal period. The winter of 2016-2017 appears the most sporadic due to the absence of surveys between December 23, 2016 and February 17, 2017. Yet, seven surveys were captured during this winter period. Because the flora has minor changes during winter, and because there is seldom snow, the loss of information due to the gap in surveys is minimal.

The evolution of one scene across all three years is shown in Fig. 4. There are large changes in appearance. Different changes are more apparent across different time scales. Changes in weather, illumination, viewpoint, and water reflectivity are apparent in many comparisons week-to-week and year-to-year. Large changes in foliage become apparent season-to-season. The montage also shows some cases of noise (e.g., sun glare).

Perspective Differences

Although surveys often captured the same scene, perspective differences occurred due to the fact that the images are captured from a mobile robot. The camera trajectory was typically at least slightly different between surveys, while sometimes factors were present that contributed to more substantial variation in viewpoint, as shown in Fig. 6. For example, the entire trajectory changed in times of high water. Aside from the fact that the camera has a fixed height from the lake's surface, more water meant the boat moved more inland. Strong winds also skewed the boat's trajectory because power to the boat's motors was set to a constant value. Fortunately, in those cases the boat could still capture a survey automatically. Perspective differences also occurred when manual control was required.

The boat's trajectory was also effected during the variable amounts of time it was in the company of a swan. A pair of swans occupied Symphony Lake. They were always peaceful towards the boat. Often during nesting season (late March through early May) the male exhibited its dominance nearby. It learned how to divert the boat from its path (swim up a side of the boat), which it typically did near the island (the annual location of the swans' nest). On these occasions we manually steered the boat on its path, at a comfortable distance from the swan, but we were unsuccessful if the boat was beyond the line of sight.

Variation in Appearance

The fact that images are captured outdoors adds to the variation in appearance caused by perspective differences. Illumination is, for example, non-uniform and varying, and a function of the sun's position in the sky and the particular weather pattern. The sun's position varied, in turn, with the time of the day and the day of the year. The more sun, the

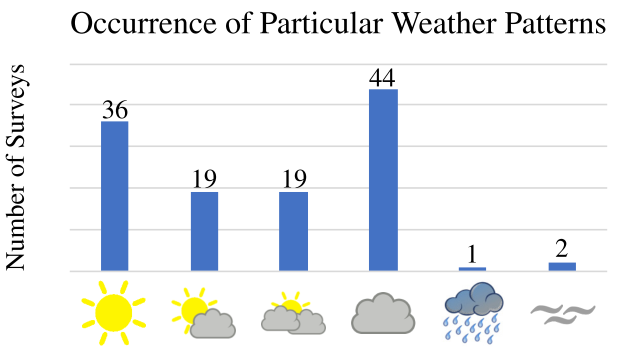

Figure 7. Occurrence of particular weather patterns in the dataset. The surveys spread well between sunny and overcast. In general, we avoided deploying the robot on rainy days. Two surveys captured fog.

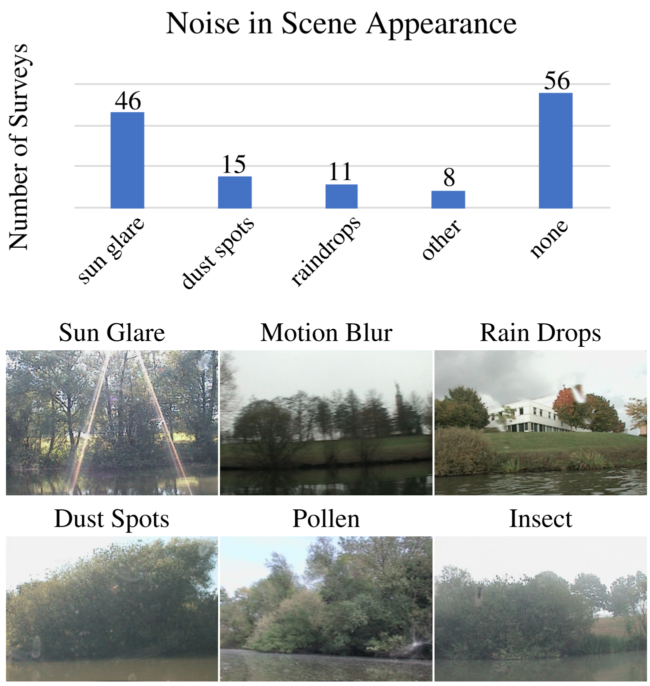

Figure 8. Significant noise was present in many surveys. Sun glare was worse than other types of noise in terms of how much it changed images, how many surveys it was present in, and how many images per survey it affected. Specular reflections on the camera dome are apparent in many images on days of strong illumination. Occasionally, other types of noise obstructed the camera view (raindrops, pollen, insects).

stronger the illumination, yet the stronger the shadows. The more direct sun, the more intense the sun glare.

Figure 7 shows that the surveys varied well from sunny to overcast. Other weather patterns were harder to capture. Rainy days were avoided because raindrops on the dome of the PTZ camera blurred the images. Fog occurred infrequently. Snow rarely accumulated in more than trace amounts. Nevertheless, the frozen lake precluded its capture when it was there. Also, none of the surveys captured the spectacular visuals of inclement weather. Thus, the dataset is without some common weather patterns, which might be part of some roadside datasets.

Variation in appearance in our dataset is perhaps stronger than in street datasets because our surveys captured a natural environment. Most images captured flora, which changed significantly across seasons. In the winter the background can be seen through trees and bushes, but in the summer and the fall it is occluded. The structures of some plants are occluded by their own foliage, which makes their recognition and association across seasons difficult. Foliage also often lacks strong features, and resembles nearby plants.

Being on a lake means that the bottom 18% of each image captures water, which varies from murky, to wavy, to reflective. Although a comparison of two images may be a success if the water is disregarded, it can interfere with the process. The flora, the shoreline, and the water blend together on days when the water is reflective. Water does, however, add scene context.

Several kinds of noise also add to the variation in appearance of the images (see Fig. 8). They typically show up as sun glare, distortions, or occlusions. Sun glare was the most prominent (per image and for how long it affected the images). It reduced image contrast and also caused other lens flare artifacts. Dust spots were also often visible in surveys with strong illumination. Debris like pollen and insects sometimes occluded the scene.

Discussion

Symphony Lake Dataset is novel as a robotic vision dataset because it captures an unstructured environment as it evolved week-to-week over three years (see e.g., Geiger et al. 2013; Glover et al. 2010; Milford et al. 2014; Naseer et al. 2017; Valada et al. 2016; Milford and Wyeth 2012; Sunderhauf et al. 2015; Dong et al. 2016; Skrede 2013, for examples of more structured, less frequent, and/or shorter time span datasets). A natural environment is dynamic, which means more can change in a smaller amount of time. Our environment has trees, water, birds, and other flora and fauna of a lakeshore, with some buildings in the background. Sometimes a lot of variation occurred between weeks.

Our dataset is interesting for the challenge it brings to perception. Data association across surveys would have to address the variation in appearance of a natural environment. Many different approaches have been proposed to improve condition–invariance in different environments (see e.g. Lowry et al. (2016) for a review of methods for place recognition; see e.g., Roy and Isler (2016) for a method designed for surveying apple orchards). In contrast to indoor environments and suburban streets, the most persistent feature of a natural environment may be its 3D structure (Griffith and Pradalier 2016).

Our dataset also presents a challenge to the SLAM community. The size and the number of surveys requires scalable optimization. Each survey potentially has hundreds of thousands of landmarks and thousands of poses. Because standard local image features lack robustness to the variation in appearance, multiple sets of landmarks may have to be used to represent the environment. Optimization must also deal with incorrect correspondences and loop closures (as in Sünderhauf and Protzel 2012; Olson and Agarwal 2013; Latif et al. 2013; Pfingsthorn and Birk 2016).

Success in these spaces could enable work towards identifying and characterizing changes in natural environments.

Through a manual comparison of images from 10 consecutive surveys (Griffith and Pradalier 2017), we know of several changes that occurred. For example, after a flood a large tree fell into the water. Automated methods for change detection would likely find more and more subtle changes. This work would help automate search and rescue along forest trails, scouting for threats in natural environments, surveilling secured sites, environment monitoring, disaster response, and precision agriculture.

Acknowledgements

This work was funded in part by the Lorraine Region, France as well as the Rhine–Meuse Water Agency.

References

Chen Z, Jacobson A, Sunderhauf N, Upcroft B, Liu L, Shen C, Reid I and Milford M (2017) Deep learning features at scale for visual place recognition. *arXiv preprint: 1701.05105* .

Dong J, Burnham JG, Boots B, Rains GC and Dellaert F (2016) 4d crop monitoring: Spatio-temporal reconstruction for agriculture. *arXiv preprint arXiv:1610.02482* .

Geiger A, Lenz P, Stiller C and Urtasun R (2013) Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research* 32(11): 1231–1237.

Glover A, Maddern W, Milford M and Wyeth G (2010) FAB-MAP + RatSLAM: Appearance-based SLAM for Multiple Times of Day. In: *ICRA*. Anchorage, USA.

Griffith S and Pradalier C (2016) Reprojection flow for image registration across seasons. In: *British Machine Vision Conference (BMVC)*.

Griffith S and Pradalier C (2017) Survey registration for long-term natural environment monitoring. *Journal of Field Robotics* 34(1): 188–208.

Latif Y, Cadena C and Neira J (2013) Robust loop closing over time for pose graph slam. *The International Journal of Robotics Research* 32(14): 1611–1626.

Lowry S, Sunderhauf N, Newman P, Leonard JJ, Cox D, Corke P and Milford MJ (2016) Visual place recognition: A survey. *IEEE Transactions on Robotics* 30(1): 1–19.

Maddern W, Pascoe G, Linegar C and Newman P (2017) 1 year, 1000 km: The Oxford RobotCar dataset. *The International Journal of Robotics Research* 36(1): 3–15.

Milford M, Firn J, Beattie J, Jacobson A, Pepperell E, Mason E, Kimlin M and Dunbabin M (2014) Automated sensory data alignment for environmental and epidermal change monitoring. In: *Australasian Conference on Robotics and Automation 2014*. Australian Robotic and Automation Association, pp. 1–10.

Milford MJ and Wyeth GF (2012) SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights. In: *International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1643–1649.

Naseer T, Oliveira GL, Brox T and Burgard W (2017) Semantics-aware visual localization under challenging perceptual conditions. In: *IEEE International Conference on Robotics and Automation (ICRA)*.

Olson E and Agarwal P (2013) Inference on networks of mixtures for robust robot mapping. *The International Journal of Robotics Research* 32(7): 826–840.

Pfingsthorn M and Birk A (2016) Generalized graph slam: Solving local and global ambiguities through multimodal and hyperedge constraints. *The International Journal of Robotics Research* 35(6): 601–630.

Roy P and Isler V (2016) Surveying apple orchards with a monocular vision system. In: *Automation Science and Engineering (CASE), 2016 IEEE International Conference on*. IEEE, pp. 916–921.

Skrede S (2013) Nordlandsbanen: minute by minute, season by season. http://nrkbeta.no/2013/01/15/nordlandsbanen-minute-by-minute-season-by-season/. Accessed: 2017-04-17.

Sünderhauf N and Protzel P (2012) Switchable constraints for robust pose graph slam. In: *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, pp. 1879–1884.

Sunderhauf N, Shirazi S, Dayoub F, Upcroft B and Milford M (2015) On the performance of convnet features for place recognition. In: *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, pp. 4297–4304.

Valada A, Oliveira G, Brox T and Burgard W (2016) Deep multispectral semantic scene understanding of forested environments using multimodal fusion. In: *The 2016 International Symposium on Experimental Robotics (ISER 2016)*. Tokyo, Japan. URL http://ais.informatik.uni-freiburg.de/publications/papers/valada16iser.pdf.

Zhou B, Khosla A, Lapedriza A, Torralba A and Oliva A (2016) Places: An image database for deep scene understanding. *arXiv preprint arXiv:1610.02055* .