2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Automated Registration for Multi-year Robotic Surveys of Marine Benthic Habitats

Mitch Bryson, Matthew Johnson-Roberson, Oscar Pizarro and Stefan Williams

Abstract— **This paper presents recent developments in data processing of multi-year repeat survey imagery and precision automatic registration for monitoring long-term changes in benthic marine habitats such as coral reefs and kelp forests. Three different methods are presented and compared for precision alignment of imagery maps collected over a range of time-scales from 12 hours to two years between dives. The first method uses Scale Invariant Feature Transform (SIFT) features computed over imagery mosaics to compute the relative translational offset between repeat dives. The second method employs scan-optimisation using the bathymetry generated via structure-from-motion thus capturing more stable features in the environment, lending itself to larger timescale registration. The third method uses mutual information optimisation to register imagery maps, providing robustness to changes in the colour and brightness of objects in an underwater scene across multiple years. Results are presented from field data collected using an Autonomous Underwater Vehicle (AUV) in sites across the Australian coast between 2009 and 2011.**

I. Introduction

Benthic marine habitats such as coral reefs, seagrass meadows and kelp forests are environments that have significant economic value worldwide and are expected to face increasing pressures from human impacts such as urban development, fishing and climate change. Long-term monitoring of these habitats provides a means for detecting and quantifying changes in the distribution and abundance of different species, aiding our understanding of human impacts. The Australian Centre for Field Robotics operates an ocean-going Autonomous Underwater Vehicle (AUV) called *Sirius* capable of undertaking high-resolution, geo-referenced surveys which is currently used as part of Australia's Integrated Marine Observing System (IMOS) [13]. As part of the IMOS program, *Sirius* is deployed at several key locations along Australian coastal waters on a yearly basis to perform repeated surveys and collect data which can be used for long-term monitoring. In contrast to variable site selection in a given area, precision revisiting of exactly the same area of benthos across multiple years provides higher statistical power for temporal change detection [5]. Precision revisiting also provides the ability to answer questions about changes in individual organisms (i.e. a single coral polyp) enabling the possibility of novel lines of inquiry for marine biologists and ecologists. The AUV employs a variety of navigation sensors including Global Positioning Systems (GPS) at the surface and ship-borne Ultra-Short BaseLine (USBL) positioning while underwater to re-localise itself in

This work was based at the Australian Centre for Field Robotics, University of Sydney, NSW 2006, Australia `contact: m.bryson@acfr.usyd.edu.au`

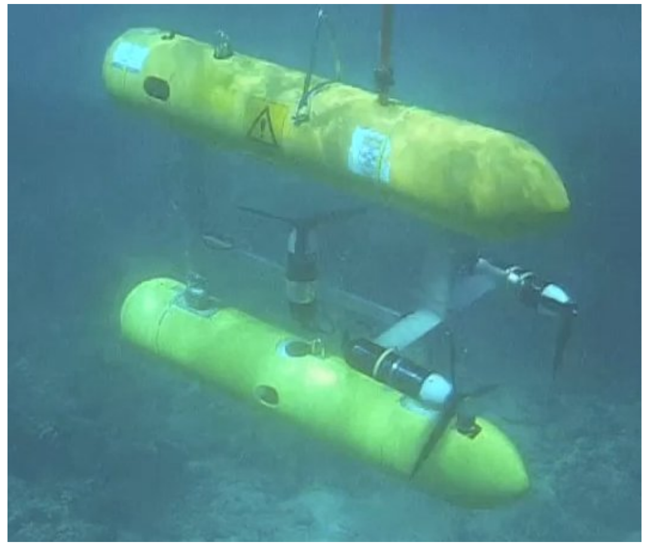

Fig. 1. The ocean-going Autonomous Underwater Vehicle (AUV) *Sirius*, capable of undertaking high-resolution, geo-referenced surveys which is currently used as part of Australia's Integrated Marine Observing System (IMOS) and used to collect the imagery data used in this study.

the same location across multiple years with an accuracy of approximately $\pm$5m. Reliable multi-year change detection requires significantly more precise registration to capture cm-level changes that can occur across multiple years in habitats such as coral reefs.

This paper focuses on developments in data processing of multi-year robotic repeat survey imagery and other map data for precision automatic registration for change detection. Due to the lack of fixed navigation infrastructure or control points, registration of data across multiple years must be performed using the map and image data itself. Registration is made challenging by variations in the water column properties and vehicle perspective that change between years and have significant effects on the colour and brightness in collected images. Furthermore, changes in the benthic coverage and assemblages themselves occur over multiple temporal and spatial scales. In this paper, we present and compare three different methods for precision alignment of imagery maps collected over multiple years. The first method uses Scale Invariant Feature Transform (SIFT) features computed over imagery mosaics to compute the relative translational offset between repeat dives. The second method employs scan-optimisation using the bathymetry generated via structure-from-motion thus capturing more stable features in the environment, lending itself to larger timescale registration.

The third method uses mutual information optimisation to register imagery maps, providing robustness to changes in the colour and brightness of objects in an underwater scene across multiple years.

Results are presented from pairs of field data collected using an Autonomous Underwater Vehicle (AUV) in sites across the Australian coast between 2009 and 2011. The time between revisiting sites varies from 12 hours to 24 months in different habitats including coral reefs and underwater boulderfields. The difference in timescales is used to assess the performance of various registration strategies to different degrees of change in the environment.

II. Related Work

In a previous study [14], SIFT image features were used to co-register multiple 3D image maps collected by the AUV comparing the same area from daytime to nighttime (over a 12 hour period). Similarly, the authors of [4] use SURF feature to register, and subsequently detect changes in images taken 10 months apart in a small coral reef scene. These studies focussed on small areas; our subsequent investigation into using local feature point techniques (i.e. SIFT, SURF) to register multiple dives conducted over large timescales (i.e. multiple years) and in a range of different benthic habitats found that these features were not always reliably matched, motivating the exploration of alternative registration techniques presented in the current paper.

The registration of time series images collected over terrestrial landscapes via satellites and aircraft in remote sensing is a related problem which has received significant attention in the literature. In [2], the authors present a method for detecting and matching tie points in overlapping imagery mosaics using intensity correlation and a multi-resolution matching scheme. More typically, registration algorithms are used to register images from several different sensor modalities including LiDAR and multi-spectral images. In [3], the authors discuss methods based on mutual information to register satellite images recorded at different wavelengths and the authors in [8] extend this concept to detect changes in satellite imagery of a city over four years tracking urban development.

Mutual information as a tool for registration has also been used extensively in medical imaging [6], [10], [11] and has shown to be robust to structural changes in the imaged-object over time series and to changes in image modality. In [11], the authors compare techniques based on mutual information for registering scans of the brain using both Computed Tomography (CT) scans and Positron Emission Tomography (PET) scans. In [10], the authors develop an efficient optimisation algorithm based on spanning graphs that uses mutual information to register PET and Magnetic Resonance (MR) images.

III. Methodology

The following subsections describe the AUV repeat surveying process and data processing procedures and outline three different methods considered for registering multiple AUV dives.

A. AUV-based Benthic Surveying

At each reference site, the AUV *Sirius* is deployed at the surface and uses GPS to navigate to a starting location. The vehicle then dives and performs a pre-programmed trajectory along the seafloor collecting stereo-image pairs, mutlibeam sonar and water column data. Different survey trajectories are used including long transects, broad-scale sparse grids and dense grids where overlapping imagery is used to provide a small-scale patch (ranging anywhere from 10x10 to 50x50m) of contiguous coverage, suitable for re-localisation within the accuracy limits of the AUV navigation system. Once the survey is complete, the vehicle returns to the surface and is recovered. Post-processing of the stereo-imagery and other navigation data via Simultaneous Localisation and Mapping (SLAM) and 3D reconstruction [9], [7] is used to provide a 3D photo-textured bathymetric reconstruction of the seafloor. An orthographic projection of this 3D model is then used to create a geo-referenced mosaic map. The data collection process is repeated after a time interval determined by the application (from 12 hours to 24 months in the results shown in this paper) by programming the AUV to dive along the same trajectory.

The resulting maps have a high orientation accuracy (approx. $\pm 0.5^o$) due to the use of tilt and magnetic sensors and a high vertical accuracy (approx. $\pm$5cm) due to the use of a pressure-based depth sensor, but low-horizontal accuracy globally (approx. $\pm$5m).

B. Mosaic-SIFT Registration

A registration technique that used SIFT features across mosaic imagery maps was developed. SIFT features were extracted from the imagery mosaic tiles for both the original and repeated survey, using the implementation provided by [12]. We extracted SIFT features from the reconstructed mosaic tiles themselves rather than the raw images. Once a set of features was extracted, one set for the original survey and one set for the repeated survey, we extracted a set of matching pairs by computing closest descriptor distance using a kd-tree [1] implementation. For each matched feature pair, the two-dimensional world-coordinate in each mosaic is extracted, based on the geo-referenced image coordinates. A 2D RANdom SAmple Consensus (RANSAC) algorithm was used to compute a robust horizontal translation between the two sets of points. The algorithm iteratively selected ten random pairs, computed the mean 2D offset from the selected points and computed the number of other pairs in the whole set that were inliers to the computed offset (i.e. pairs that were within 20cm of each other after the offset was applied were considered inliers). The process was repeated $k = \frac{\log(1-p)}{\log(1-w^n)}$ times (i.e. to ensure at least one set of n selected points was all inliers with probability p, assuming w is the ratio of inliers to the total number of matched features ($p = 0.99$, $w = 0.5$, $n = 10$)). The offset that resulted in the largest set of inliers was selected as the optimal registration of the two mosaics.

C. Terrain Scan-optimisation Registration

A second registration method using scan-optimisation of the 3D bathymetry underlying the imagery mosaics was developed. Digital Elevation Maps (DEMs) derived from the stereo image data were compared between dives along a fixed grid of different 2D horizontal offset values over the North-South (x) and East-West (y) directions. The grid spacing was at a resolution of 10cm reflecting the average spacial point resolution achieved in the structure-from-motion mapping pipeline. At each potential offset, the overlapping section of DEMs were extracted and the difference in height (Δz at each point in the overlapping grids) was computed. At each potential offset value, a miss-registration cost $c_{x,y}$ was computed:

$$c_{x,y} = \frac{\frac{1}{N}\sum_{x,y}\Delta z^2}{\sigma_{z_1}\sigma_{z_2}} \quad (1)$$

where σ_{z_1} is the standard deviation of heights in the overlapping area of the first DEM, σ_{z_2} is the standard deviation of heights in the overlapping area of the second DEM and N is the total number of overlapping grid cells for a given offset in (x, y). For different offsets, the amount of overlapping area used in the comparison varied (due to small differences in the area covered by each survey, gaps in the coverage etc). The normalising term $\sigma_{z_1}\sigma_{z_2}$ served to increase the cost associated with a registration that only considered flat, low vertical variance terrain overlap (that typically occurred when the area of overlap for the considered offset was very low). The offset which resulted in the lowest miss-registration cost $c_{x,y}$ was chosen to register the maps into a single corrected coordinate system.

D. Mosaic Mutual Information Registration

A third registration method using a mutual information optimisation [11] of imagery mosaic intensity histograms was developed. Mutual information provides a generalised measure of the consistency of a proposed alignment and is robust to non-linear changes in the colour of objects across years when compared to direct comparison/matching of mosaic image intensity values. Offset values were compared between dives along a fixed grid of offset values over each of the x and y directions (i.e. as performed in Section III-C). The grid was evaluated at a resolution of 5cm; due to the mechanics of the structure-from-motion process, image mosaics possessed a higher resolution than structural point clouds used to construct the DEM described in Section III-C. This allowed the imagery mosaic registration method to be performed at a higher resolution than the DEM registration. At each potential offset, the overlapping section of image mosaics were extracted, converted to grey-scale and a joint histogram ($\mathbf{J}$) of the greyscale intensity values computed. The two-dimensional joint histogram was constructed with greyscale intensity values between 0 and 255 with a bin size of 1; for each overlapping grid cell, a value was added to the histogram with the coordinates provided by the 8-bit grey-scale values of the original survey mosaic and the repeat survey mosaic. The mutual information between the image intensity distributions was then calculated using:

$$M_{x,y} = -\sum_{m2}\left(\sum_{m1}\mathbf{J}\log\sum_{m1}\mathbf{J}\right) - \ldots \sum_{m1}\left(\sum_{m2}\mathbf{J}\log\sum_{m2}\mathbf{J}\right) - \sum_{m1,m2}(\mathbf{J}\log\mathbf{J}) \quad (2)$$

where $\sum_{m1}$ is the summation over bins corresponding to the first map, $\sum_{m2}$ is the summation over bins corresponding to the second map and $\sum_{m1,m2}$ is the summation over all bins of the joint histogram. The offset which resulted in the highest mutual information $M_{x,y}$ was chosen to register the maps into a single corrected coordinate system.

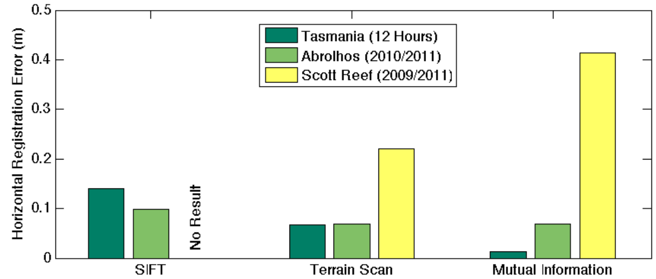

Fig. 2. Multi-year registration errors (difference from hand selected feature registration) of the three sites using the three different registration methodologies.

Computed Repeat Survey Offset from Geo-referenced Solution		
Tasmania (12 Hours)	Abrolhos (2010/2011)	Scott (2009/2011)
5.2113m	0.7708m	2.4851m
Average Residual Errors		
Tasmania (12 Hours)	Abrolhos (2010/2011)	Scott (2009/2011)
0.2258m	0.0725m	0.2095m

TABLE I

COMPUTED OFFSETS AND AVERAGE RESIDUAL ERRORS IN HAND SELECTED FEATURE REGISTRATION FOR THE THREE SITES.

IV. EXPERIMENTAL SETUP

A. Study Sites

Three different repeat surveys performed over a variation of revisit times were used to assess and compare the three different registration methods presented in the previous section. The first study site was an underwater boulder field site located off the coast of Tasmania, in southeast Australia. The dives were performed approximately 12 hours apart (once during the day and once during the night) as part of a study of the behaviour of sea-urchins, which were known to be predominantly active at night. The second study site was over a section of low-relief reef at the Abrolhos islands off the coast of Western Australia where two AUV dives were performed over subsequent years (April 2010 and April 2011). The reef was found to have undergone a bleaching

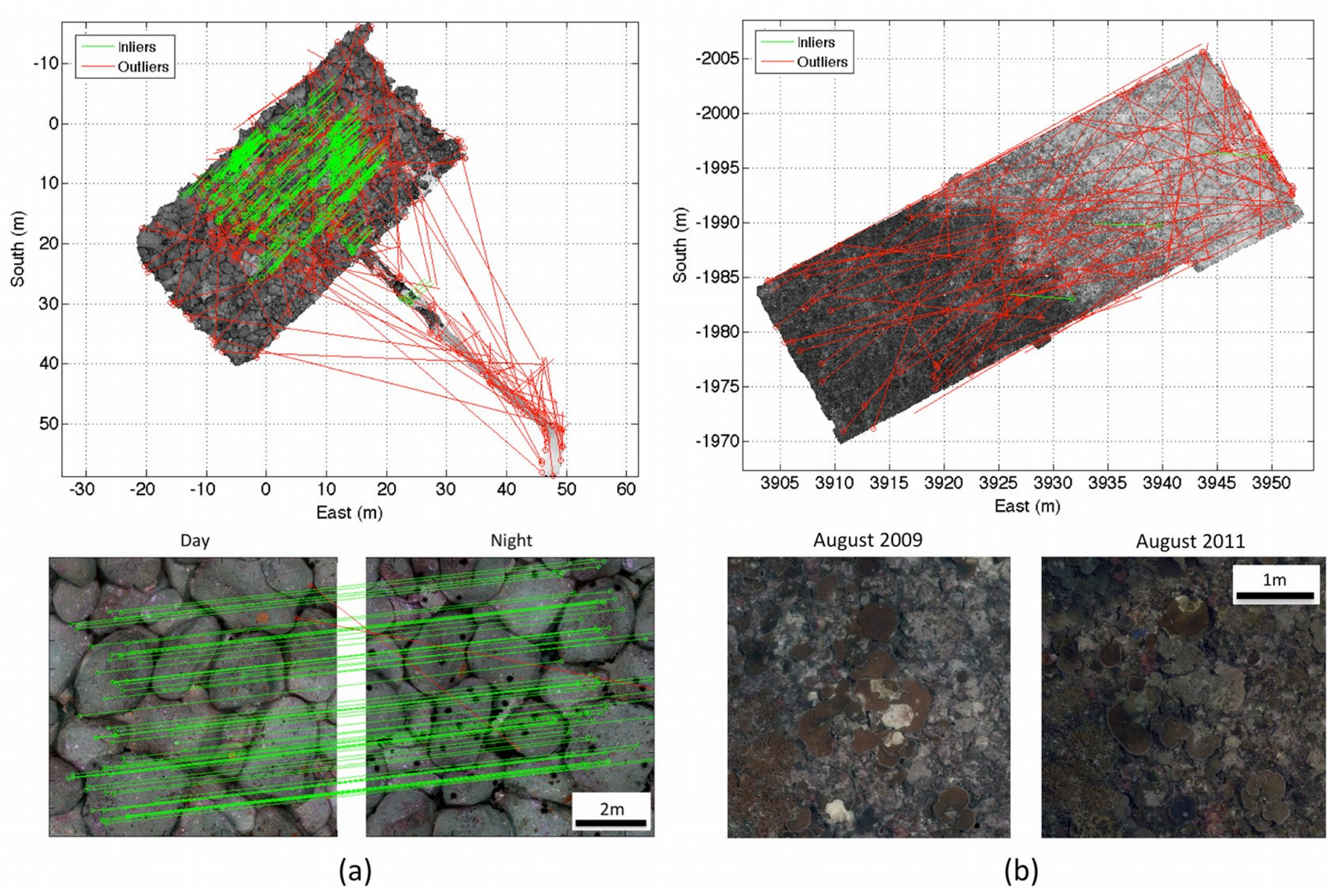

Fig. 3. Results of SIFT feature based registration: (a) The upper figure illustrates the day-time imagery mosaic for the Tasmania boulderfield mission with overlaid matched feature tracks representing the estimated translation of features after map registration (RANSAC inliers are displayed in green and outliers in red). The lower figure displays magnified sections of the mosaic for both the day and night time dives with overlaid matching SIFT features. (b) The upper and lower figures display the same information in (a) but for the Scott Reef 2009/2011 mission. SIFT features were not matched consistently over large timescales (i.e. 2 years for (b)) owing to the small-scale changes in the environment.

event in which a large proportion of the coral had turned white. The third study site was over a section of reef/sand interface at Scott Reef off the coast of Western Australia. Repeat dives of the area were performed 24 months apart in August 2009 and August 2011.

B. Validation using Hand Selected Feature Registration

In order to validate the accuracy of the registration results, registration was also performed via hand-selected features to provide an evaluation of the automatic registration techniques. A human-expert provided 25 pairs of registered points selected across the image mosaics which were used to compute a horizontal registration using the 2D RANSAC method described above in Section III-B. The hand-registered offsets were then used as a "ground-truth" to which the automatic registration techniques were compared.

V. Results

Figure 2 presents the final horizontal offset errors (difference between the hand-registered and automatically registered offsets) computed for each of the registration techniques on each of the three datasets. Table I shows the computed horizontal registration offsets of the hand-selected feature registrations for each mission that was used as a ground-truth comparison for the automated techniques. Table I also shows the average residual errors across the 25 hand-selected features indicating the inherent accuracy of the registrations owing to internal errors or inconsistencies in each individual map. Figures 3, 4 and 5 present example results of the three different registration methods across the different datasets.

Overall the Scott Reef dataset exhibited larger registration errors than in the other datasets. This is primarily related to the increase in time between the repeat dives and thus the degradation of structural and visual similarity. The SIFT feature method failed to produce a registration for the Scott Reef dataset (24 months apart) and upon inspection it was found that none of the features had been correctly matched between the mosaics. The terrain scan-optimisation and mutual information registration methods were able to produce registrations that had comparable accuracy to SIFT feature matching over short periods of time (i.e. for the Tasmania dataset with a 12 hour difference) and also provided valid registrations over large timescale where SIFT features failed.

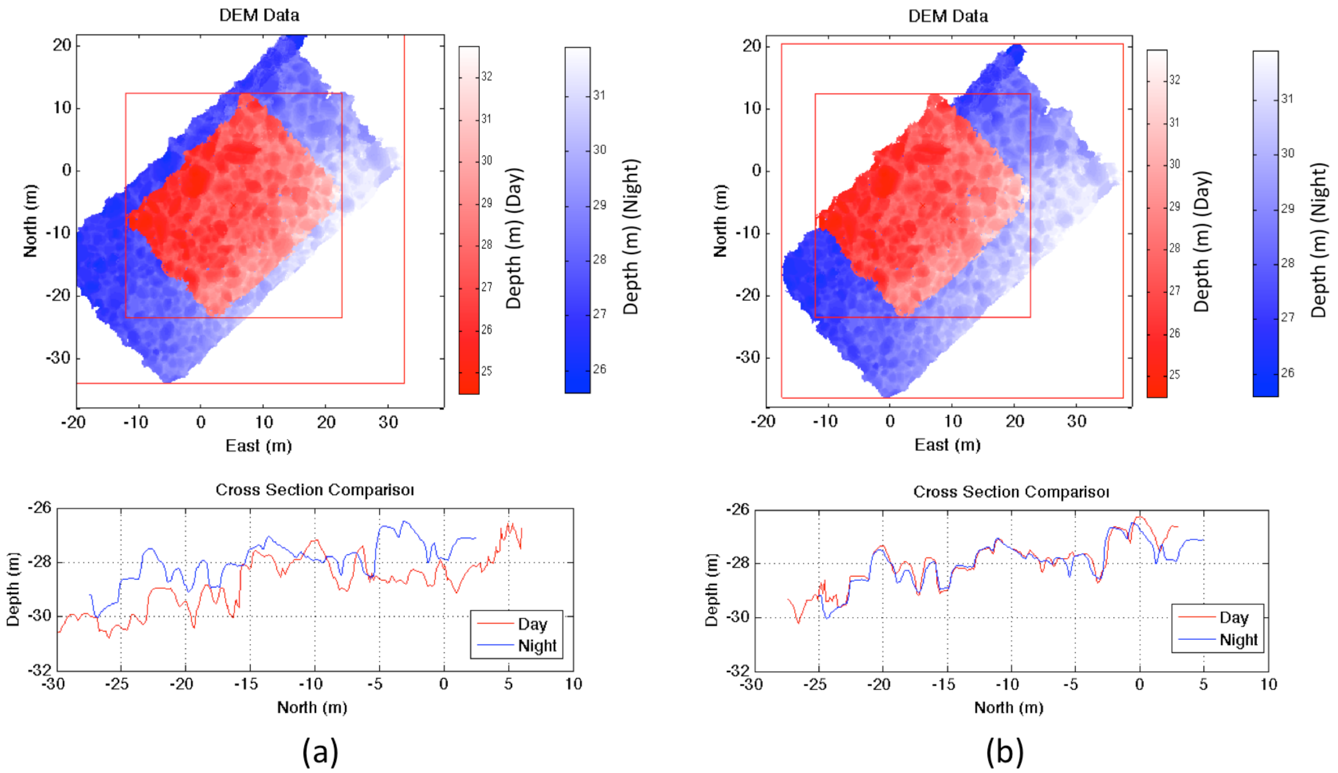

Fig. 4. Digital Elevation Map (DEM) registration of day (shown in red) and night (shown in blue) repeat surveys: (a) Original terrain map alignment based on navigation data and, (b) final terrain maps after scan optimisation alignment. Lower figure displays a north-south cross-section taken through both DEM models at 5m east of the map origin before and after alignment.

The SIFT feature method was successful at registering the Tasmania boulderfield dataset (12 hours difference) with 72% of the features computed as inliers. The Abrolhos islands dataset (12 months difference) was also successfully registered however only a very small proportion of features were successfully matched (12%). Examples of the SIFT feature matching for the Tasmania and Scott Reef datasets are show in Figure 3. From Figure 3 (b) it can be seen that although the overall position of individual corals can be tracked across the mosaic, small-scale features in the data exploited by SIFT (such as coral edges) have changed subtly across the years due to coral growth, death and weathering.

Figure 4 shows the results of the scan-optimisation matching algorithm when applied to the DEM models generated from the Tasmanian day/night dives. Figure 4 (a) illustrates the DEM models overlaid in their original geo-referenced coordinate systems as provided by the navigation sensor on the AUV. Figure 4 (b) illustrates the final alignment of the two models after applying scan-optimisation and a corresponding cross-section of the two 3D models. The cross-section has been taken in the north-south direction at 5m east of the map origin before and after alignment. The terrain scan-optimisation method worked well in missions with a high-degree of terrain complexity or relief (such as in the boulderfield environment).

For the Tasmania and Abrolhos islands datsets, the mutual information registration method provided accuracies that were equivalent to the other two methods. The results of the mutual information registration technique for the Abrolhos islands dataset are shown in Figure 5. Figure 5 (a) shows both the overall 2010/2011 mosaic maps compared prior to alignment. Shown in the lower figures are zoomed-in sections of the mosaics with the same geo-graphical coordinates after alignment, illustrating the registration of coral across the two dives. The mutual information optimisation is able to co-register the mosaics even in the presence of large colour changes (e.g. bleached coral) between years. Figure 5 (b) shows the calculated mutual information as a function of the considered offset between the mosaic maps, illustrating a global maximum corresponding to the optimal alignment.

VI. Conclusions and Future Work

This paper has discussed recent developments in automatic registration of multi-year AUV imagery surveys performed as part of a long-term benthic habitat monitoring program. Three different methods using SIFT features, terrain scan-optimisation and mutual information were presented and compared for precision alignment of imagery maps collected over a range of time-scales from 12 hours to two years between dives. Using hand-labelled registrations as a validation, the results indicate comparable registration accuracies can be achieved using all three methods over short time periods and that terrain scan-optimisation and mutual

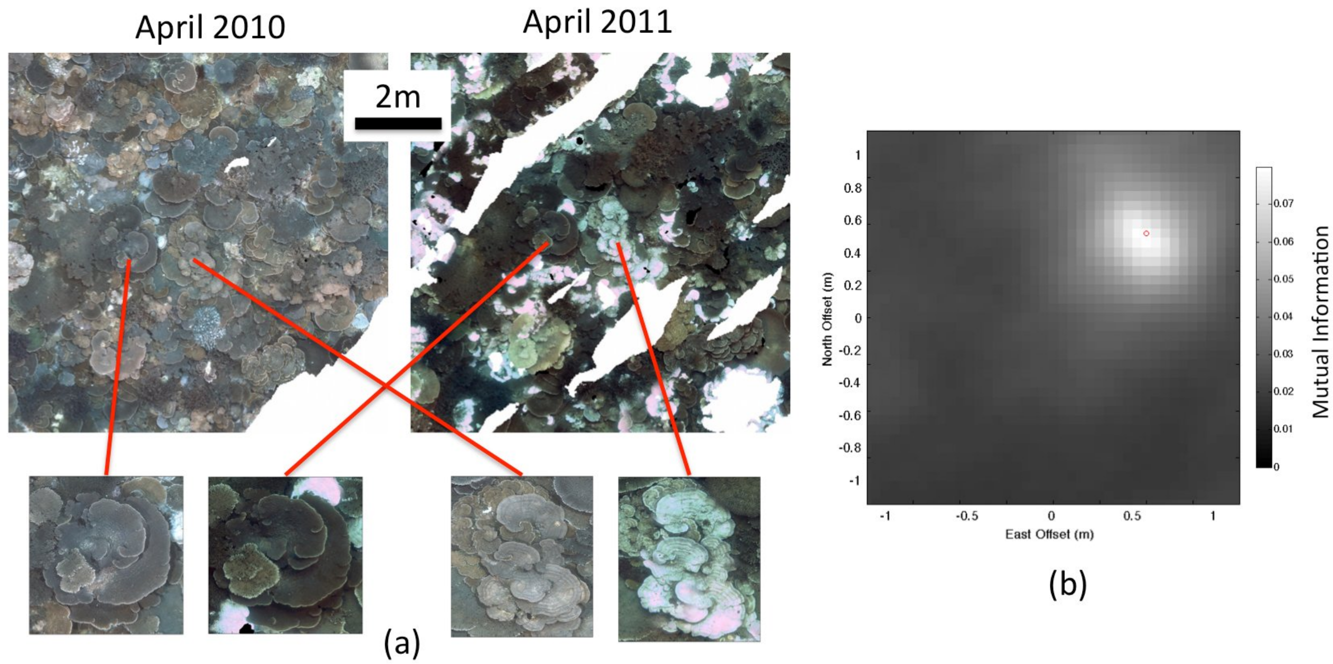

Fig. 5. Mutual information registration of coral mosaics across a 12 month period: (a) AUV mosaics of a coral reef from April 2010/April 2011 with highlighted sections of both un-affected and bleached coral, co-registered using mutual information optimisation of the mosaic greyscale images. (b) Mutual information utility surface as a function of offset between the mosaics used in computing the optimal registration of the mosaics.

information are effective registration methods over longer time periods where the SIFT feature method fails.

Future work will explore additional registration techniques and methods for performing non-rigid registration over large timescales. The inherent residual errors present in individual maps warrant adapting the internal structure of mosaics for better alignment which has been demonstrated using SIFT features (for example see [14]) but is more difficult to perform using global consistency measures such as mutual information. Future work will also focus on real-time registration and localisation methods allowing the AUV to more precisely repeat a previous survey. Real-time registration and re-planing would allow for a reduction in the accuracy requirements of external geo-referencing systems such as USBL, alleviating the logistical complexity in long-term benthic monitoring programs.

VII. Acknowledgments

This work is supported by the ARC Linkage Project LP0989998, funded by the Australian Research Council (ARC) and the New South Wales State Government. This work is also supported by the Integrated Marine Observing System (IMOS) through the DIISR National Collaborative Research Infrastructure Scheme.

References

[1] J.S. Beis and D.G. Lowe. Shape indexing using approximate nearest-neighbor search in high-dimensional spaces. In *Computer Vision and Pattern Recognition*, pages 1000–1006, 2003.

[2] P. Bunting, F. Labrosse, and R. Lucas. A multi-resolution area-based technique for automatic multi-modal image registration. *Image and Vision Computing*, 28(8):1203–1219, 2010.

[3] H. Chen, M. K. Arorab, and P. K. Varshney. Mutual information-based image registration for remote sensing data. *International Journal of Remote Sensing*, 24(18):3701–3706, 2003.

[4] O. Delaunoy, N. Gracias, and R. Garcia. Towards Detecting Changes in Underwater Image Sequences. In *MTS/IEEE Oceans Conference*, 2008.

[5] J. Van der Meer. Sampling design of monitoring programmes for marine benthos: a Comparison between the use of fixed versus randomly selected stations. *Journal of Sea Research*, 37:167–179, 1997.

[6] J. Jiang, S. Zheng, A.W. Toga, and Z. Tu. Learning based coarse-to-fine image registration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[7] M. Johnson-Roberson, O. Pizarro, S.B. Williams, and I. Mahon. Generation and Visualization of Large-scale Three-dimensional Reconstructions from Underwater Robotic Surveys. *Journal of Field Robotics*, 27(1):21–51, 2010.

[8] Y. Kita. A study of change detection from satellite images using joint intensity histogram. In *19th International Conference on Pattern Recognition*, 2008.

[9] I. Mahon, S. Williams, O. Pizarro, and M. Johnson-Roberson. Efficient View-based SLAM using Visual Loop Closures. *IEEE Transactions on Robotics*, 24(5):1002–1014, 2008.

[10] M. R. Sabuncu and P. Ramadge. Using Spanning Graphs for Efficient Image Registration. *IEEE Transactions on Image Processing*, 17(5):788–797, 2008.

[11] P. Thevenaz and M. Unser. Optimisation of Mutual Information for Multiresolution Image Registration. *IEEE Trans. on Image Processing*, 9(12):2083–2099, 2000.

[12] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms (http://www.vlfeat.org/), 2008.

[13] S. Williams, O. Pizarro, M. Jakuba, C. Johnson, N. Barrett, R. Babcock, G. Kendrick, P. Steinberg, A. Heyward, P. Doherty, I. Mahon, M. Johnson-Roberson, D. Steinberg, and A. Friedman. Monitoring of Benthic Reference Sites: Using an Autonomous Underwater Vehicle. *IEEE Robotics and Automation Magazine*, 19(1):73–84, 2012.

[14] S. Williams, O. Pizarro, M. Jakuba, I. Mahon, S. Ling, and C. Johnson. Repeated AUV surveying of urchin barrens in North Eastern Tasmania. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 293–299, 2010.