Proceedings of the 12th International Coral Reef Symposium, Cairns, Australia, 9-13 July 2012
5D Robots and machine learning techniques aiding coral reef science

Measuring reef complexity and rugosity from monocular video bathymetric reconstruction

Hu He[1], Renata Ferrari[2], David McKinnon[3], George A. Roff[2], Ryan Smith[1], Peter J. Mumby[2], Ben Upcroft[1]

[1]EECS, Queensland University of Technology, Brisbane, QLD 4000 Australia
[2]Marine Spatial Ecology Lab, University of Queensland, School of Biological Science, QLD, 4072 Australia
[3]IPPLEX, Suite 4 2043, Santa Monica, CA 90404, USA
Corresponding author: hu.he@student.qut.edu.au

Abstract. Topographic structural complexity of a reef is highly correlated to coral growth rates, coral cover and overall levels of biodiversity, and is therefore integral in determining ecological processes. Modeling these processes commonly includes measures of rugosity obtained from a wide range of different survey techniques that often fail to capture rugosity at different spatial scales. Here we show that accurate estimates of rugosity can be obtained from video footage captured using underwater video cameras (*i.e.,* monocular video). To demonstrate the accuracy of our method, we compared the results to *in situ* measurements of a 2*m* x 20*m* area of forereef from Glovers Reef atoll in Belize. Sequential pairs of images were used to compute fine scale bathymetric reconstructions of the reef substrate from which precise measurements of rugosity and reef topographic structural complexity can be derived across multiple spatial scales. To achieve accurate bathymetric reconstructions from uncalibrated monocular video, the position of the camera for each image in the video sequence and the intrinsic parameters (*e.g.,* focal length) must be computed simultaneously. We show that these parameters can be often determined when the data exhibits parallax-type motion, and that rugosity and reef complexity can be accurately computed from existing video sequences taken from any type of underwater camera from any reef habitat or location. This technique provides an infinite array of possibilities for future coral reef research by providing a cost-effective and automated method of determining structural complexity and rugosity in both new and historical video surveys of coral reefs.

Key words: Structural complexity, Rugosity, Monocular video, Online calibration, Reconstruction.

Introduction

Coral reefs are biologically complex ecosystems that support a variety of marine organisms: approximately 22% of marine fish species and more than 25% of all known marine species (Spalding et al. 2001). Structural complexity of coral reefs play an ecologically important role in observing coral growth or decline rates, coral coverage and levels of biodiversity, which is critical to understand fish assemblages and abundance as well as ecosystem processes (Risk 1972; McCoy and Bell 1991).

Rugosity of coral reefs is traditionally measured *in situ* by SCUBA divers along a single, linear profile using chain-tape methods or profile gauges (McCormick 1994). Specifically, divers employ quadrats and transects to measure the size and distribution of coral colonies *in situ*. The chain-tape method is the most commonly used and its accuracy is highly limited by transect length and chain links size. These human-based survey methods are labor intensive and expensive due to high operational cost and human physiological limits on dive time and depth. Hence, surveys can only monitor a small number of individual corals and limited reef area. As a result, these surveys are spatially and temporally sparse and difficult to replicate. These limitations restrict our understanding of the mechanistic processes occurring at different scales.

Alternative methods for surveying large reef areas (tens to thousand of meters) are proposed based on remote sensing methods, such as aerial LIDAR and satellite-based imagery. However, these remote sensing methods cannot provide fine details of underwater structures due to the limited spatial resolution (Brock 2004) and the limited depth range caused by severe attenuation of light in water.

Underwater video cameras are becoming more commonly used in research because they can provide high resolution images of coral colonies and cover large areas quickly. Additionally, they can also produce a permanent visual record of reef condition and prevent the potential damage due to untouched measurement. One work (Lirman et al. 2007) applied video mosaic to construct 2D spatially accurate high resolution mosaic for underwater surveying. However, this method is not geometrically accurate due to ignorance of 3D structure of the scene. The Australian Centre for Field Robotics (ACFR) employed Autonomous Underwater Vehicles (AUVs) equipped with stereo cameras to collect reef data and calculate structural complexity based on bathymetric stereo image reconstruction (Friedman et al. 2010). However, the operation of such vehicles and their support workstation is expensive, and their large size is not suitable for survey in shallow water. Moreover, stereo vision system requires additional design for use by AUVs or divers comparing to monocular cameras. Additionally, stereo cameras have only been introduced very recently to a few coral reef surveys (Bridge 2011), hence most if not all existing video surveys have used monocular cameras. The method to estimate coral reef structural complexity and bathymetric reconstruction from monocular video will provide significant benefit. Note that the drawback of monocular video reconstruction is that there is no intrinsic method for computing the scale of the scene. However, most of the targeted experiment spots have some artificial marks, which can help resolve true scale.

The method proposed here uses the uncalibrated monocular video obtained by a SCUBA diver to generate dense bathymetric reconstruction of a selected transect in a fore reef. Then seafloor can be detected via plane fitting approach. Maximum height and structural complexity of each quadrat in the transect are computed, based on a reconstructed 3D model. Finally, the accuracy of computed maximum height profile and structural complexity are evaluated by comparing *in situ* measurements of the same area obtained by a team of scientific divers.

Material and Methods

Video data acquisition

The field activities for this study were conducted on Glovers Reef Atoll (87°48'W, 16°50'N), located 52 km offshore and 15 km east from the Mesoamerican Barrier Reef off Belize, Central America. Particularly, the experimental videos and measurements were collected from the Montastraea annularis dominated fore reef (depth of 10-12 m) on the eastern side of the atoll during 2009. The system has high wave energy and water flow due to its windward orientation (McClanahan 1998; Renken 2009), which has a direct effect on the structural complexity of the reef.

An area of approximately 16 x 20 m was selected on the fore reef slope of the atoll, approximately 2 m away from the drop off in order to avoid a steep slope or wall. The perimeter of the area was marked using a thin rope (5 mm in diameter) and string was used to subdivide the area in eight equal transects of 2 x 20 m. Subsequently each transect was subdivided into 10 quadrats (2 x 2 m), flagging tape was used to mark the corners of each quadrat. Video transects following a lawnmower pattern were taken using a high definition (1280 x 720) Sanyo Xacti video camera at 30*Hz* in an Epoque housing. The camera was held perpendicular to the substratum at a height of approximately 1 m (Fig. 1).

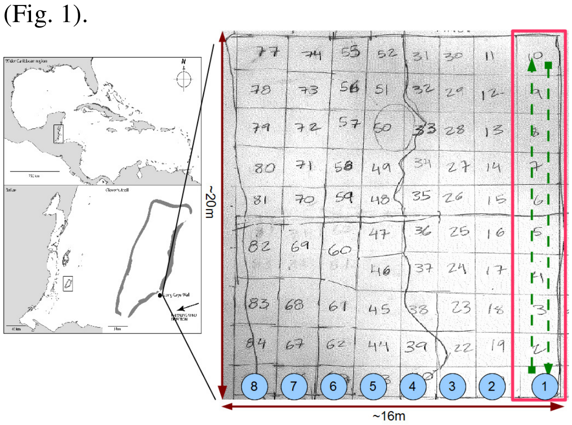

Figure 1: Area targeted for diver surveys (about 16 x 20 m). The surveyed spot consists of 8 transects (about 2 x 10 m) indexed by the blue numbered circles. Each transect contains 10 quadrats of 2 m by 2 m. This study is conducted on the first transect including quadrat 1 to 10 highlighted by the red rectangle on the right. The green dash line indicates the lawnmower pattern swam by the diver while filming, therefore each video swim covered half a transect.

Ground truth from in situ measurements

A team of 4 divers mapped and characterized the benthos of the area following a modified version of the Atlantic and Gulf Rapid Reef Assessment v4 (AGRRA) methodology. This methodology was chosen because: (1) it is the most effective methodology at quantifying the condition of coral reef communities, (2) it is widely used and it has a regional database of Caribbean coral reef condition (ww.agrra.org), (3) it is the most detailed methodology available in the region (Kramer 2003) and, (4) it measures rugosity and coral density (www.agrra.org).

The main modification made to the AGRRA protocol was the use of quadrats instead of transects, and 100% of the benthos, not only a proportion of it, was mapped. Within each quadrat every structure larger than 10 *cm* in diameter was mapped measured and identified to species (hermatypic corals) or family (sponges, calcareous algae, etc.) level. In order to assess size three measurements were taken from each colony (dead or alive) and rigid sponge: (1) the maximum diameter a, (2) perpendicular diameter b (both perpendicular to the axis of growth) and (3) maximum height h_{colony} (parallel to the axis of growth); these were measured to the nearest centimeter using a 1 m pvc pole marked every 5 cm. The structural complexity index (*0-1*) for each quadrat was calculated by combining the spatial distribution and size data using the following equation:

$$SC_{quadrat} = \frac{\pi abh_{colony}}{4l_{quadrat} w_{quadrat} h_{transect}} \quad (1)$$

where $SC_{quadrat}$ represents the structural complexity of the quadrat, $l_{quadrat}$ and $w_{quadrat}$ denote the length and width of each quadrat, respectively, which are 2 m equally, and $h_{transect}$ denotes the maximum height of the surveyed transect across ten quadrats.

Monocular bathymetric reconstruction
Multiple-view stereo based reconstruction method (Snavely et al. 2006; Furukawa et al. 2010) in computer vision has been employed in this paper. Firstly, camera positions for each frame are obtained by a structure-from-motion algorithm which consists of feature extraction and tracking, followed by camera poses estimation routines and bundle-adjustment to refine the solutions.

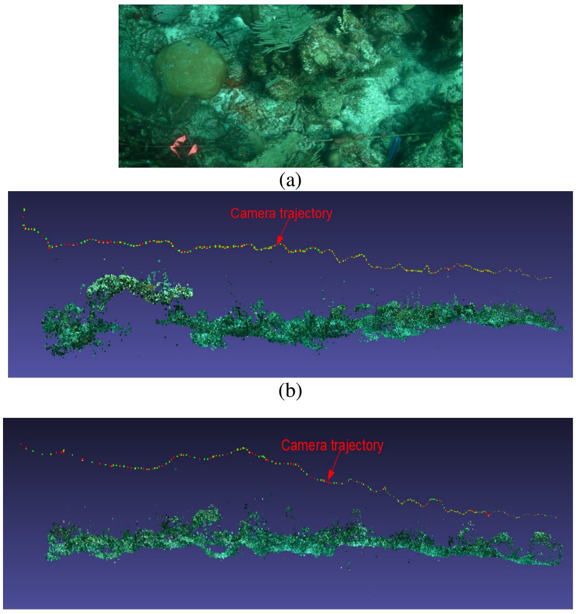

Figure 2: Sample frame and dense reconstructions. (a) A sample frame from video data for the first transect, (b) dense reconstruction of quadrats in the left half side of the first transect, and top color line denotes the estimated camera path, (c) dense reconstruction of quadrats in the right half side of the first transect, and top color line denotes the estimated camera path.

As camera parameters (focal length, principle points, distortion etc.) are unknown, calibration results were determined on the fly (Pollefeys et al. 1998). Then, a dense depth map for each frame was obtained based on triangulation of matched features and precomputed camera poses. Finally, a dense 3D point cloud model were obtained by projecting features with depth to 3D space and discarding the outlines through depth consistency checking (McKinnon et al. 2011).

Two video clips were taken for each transect separately following a lawnmower pattern. By watching the video empirically, each video clip covered 50 - 60% area of the transect. The monocular reconstructions of each of these video clips covering half transect may have different scales. To overcome this challenge, we reconstructed the two half-transects individually and then scaled them based on the metric information from real data. In other words we used the known size of an object in the video to obtain the scale, for example, Fig. 2 (a) is a sample frame of the video clip, and the line and flag at the bottom of the frame can be used to infer the scale information. Fig. 2 (b) and (c) illustrate the dense reconstructions of two half sides of the first transect, respectively. The red and green dots at the top of these figures represent estimated camera trajectory, which does not need to be parallel to sea floor. Thus, our method relaxed the restriction on video capture.

Maximum height profile and structural complexity
As the initial reconstructions of two half-transects do not lie in the same coordinate system as *in situ* measurements, we align the reconstructed sea floor with the *x-y* plane to make straightforward comparison with ground truth. Firstly, a robust plane fitting algorithm was applied on reconstructed sea floor to find the normal direction of sea floor, then we transform the reconstructions to the same coordinates system as the *in situ* measurements. Both reconstructions are thus scaled with actual metric dimension. In Fig. 3, two reconstructed half transects are merged via recovered scale and displayed in 3D space with different colors representing different quadrats. Moving objects (fish, gorgonians, etc.) are not reconstructed, they resulted in sparse regions and are shown in white in the reconstructions. Because our reconstruction method assumes the reconstructed scene is static, which means features from scene should exist in at least three frames, if otherwise the reconstructed points will be discarded.

Given the reconstructions have been registered in the same coordinates, the maximum height for each quadrat in the first transect is computed as below:

$$h_{max}(i) = \begin{cases} h^1_{max}(i) & h^1_{max}(i) > h^2_{max}(i) \\ h^2_{max}(i) & h^1_{max}(i) < h^2_{max}(i) \end{cases} \quad (1)$$

where $h_{max}(i)$ is the maximum height in the i^{th} quadrat, while $h^1_{max}(i)$ and $h^2_{max}(i)$ denote the maximum height for the i^{th} quadrat on the left half side and the right half side of the first transect, respectively.

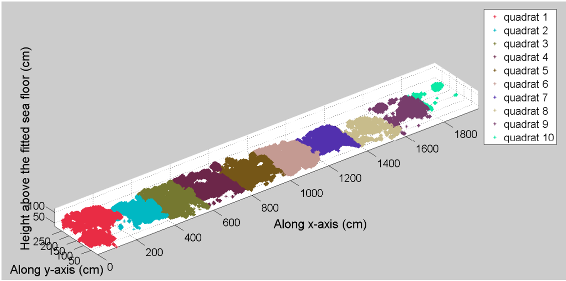

Figure 3: Alignment of each quadrat from the first transect in 3D space. Different color denotes different quadrat.

Structural complexity of the first transect are defined as the volume integral over each quadrat.

$$SC_{quadrat}(i) = \int\int h_i(x,y)\,dxdy \qquad (2)$$

where i is the index of quadrat in the first transect, and $h(x,y)$ is the height distribution of each quadrat.

Finally, we compared our structural complexity values using a two tailed pairwise t-test. The accuracy of our method is also evaluated by the relative absolute error (RAE) as in equation 4 (Harvey et al. 2000) in the estimations of maximum height and structural complexity of each quadrat from the first transect between the values obtained from 3D reconstruction X_E and the values obtained from the *in situ* measurements (ground truth) X_G:

$$RAE = \frac{|X_E - X_G|}{X_E} \qquad (4)$$

Results

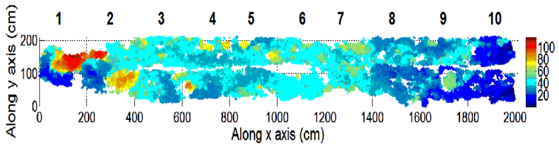

Figure 4: Height map for ten quadrats indexed by number (1-10) in the first transect. Note that the dimension has been scaled with metric measurements, where red indicates the highest colonies and dark blue the shorter ones (in cm). The sparse white regions indicate no reconstructions.

Fig. 4 shows the height map for ten quadrats in the first transect. There are some sparse white regions indicating no reconstruction for that area due to moving objects. The highest coral colony (140 cm) lies in the first quadrat, which is qualitatively correct by comparing with diver's measurements of the same quadrat. We can also find that the first quadrat of the transect is reconstructed partially. Because the diver filmed the scene at roughly constant height above the sea floor, the higher the coral colony is, the closer the substrate was to camera, resulting in a small part of the quadrat not being captured in the video. Therefore, the actual area covered by field of view is smaller provided the focal length is fixed. Moreover, it implicitly indicates large error of structural complexity will be introduced for the first quadrat (Fig. 5b, Table 1).

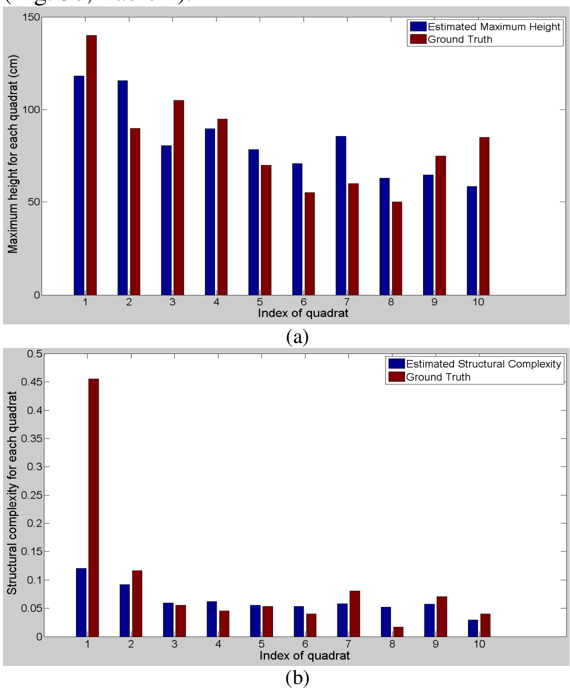

Figure 5: Comparison between estimated maximum height and structural complexity of each quadrat in the first transect and ground truth data. (a) shows maximum heights plot of each quadrat from estimation (blue) and from ground truth (red), (b) indicates the estimated structural complexity (blue) of each quadrat with respect to ground truth (red).

Table 1: Relative absolute error of maximum height and structural complexity for each quadrat in the first transect, as computed by Eq. 4.

Index of Quadrat	Maximum Height (RAE)	Structural Complexity (RAE)
Quadrat 1	0.18	2.78
Quadrat 2	0.22	0.27
Quadrat 3	0.3	0.07
Quadrat 4	0.06	0.26
Quadrat 5	0.11	0.04
Quadrat 6	0.22	0.26
Quadrat 7	0.3	0.4
Quadrat 8	0.2	0.67
Quadrat 9	0.16	0.23
Quadrat 10	0.45	0.36

Fig. 5 illustrates the quantitative comparison between the maximum height as well as structural complexity of each quadrat in the first transect and ground truth. In Fig. 5a, the trend of maximum heights are quite similar to *in situ* measurements (with an average error of 20.2 cm). Table 1 shows that there is a large error in quadrats 3, 7 and 10 (around 30 – 45% RAE of the estimated maximum height).

This is because the scenes occluded by moving objects cannot be reconstructed. A high density of large moving gorgonians was presented in quadrats 3 and 7, resulting in partial occlusion of coral colonies. Thus some information on said coral colonies could not be reconstructed. This indicates that small size coral colonies would be easier affected by occlusions.

The error of maximum height for quadrat 10, however, is introduced by drift of the proposed algorithm. The estimated camera path over time drifted because the present video data was captured by an uncalibrated camera. Using precalibrated cameras or robust online calibration, or fusing with other sensors (IMU, GPS, DVL, etc.) can improve camera poses estimation (Roberson et al. 2010).

Fig. 5b shows the quantitative comparison between the structural complexity indexes of each quadrat in the first transect obtained from the 3D reconstruction and from the ground truth data. As discussed above, quadrat 1 is only partially reconstructed, thus large error appears comparing with *in situ* measurement, hence we compared our estimations with the ground truth data twice, one with all the quadrats and the second time excluded quadrat 1. Estimated structural complexity from our 3D reconstruction method matched the ground truth data accurately (S.D. ±0.11, Table 1). Structural complexity is not significantly different among our 3D reconstructions and *in situ* measurements ($p = 0.359$ including all quadrats, and $p = 0.937$ excluding quadrat 1).

Discussion

In this paper we show that the proposed monocular bathymetric reconstruction method can obtain height profile and structural complexity of the surveyed region automatically and quickly with an accuracy of 20.2 cm and ±0.11 S.D., respectively. As the use of digital imagery and video surveys in benthic monitoring has increased dramatically, accurate and efficiently image based method become more and more important in coral reef research.

There are different ways in which the accuracy of the estimation from our algorithm could be improved: (1) currently, cross section of a coral colony is assumed to be ellipse, and volume of coral colony is computed by the area of approximate ellipse multiplied by maximum height of that coral colony, potentially introducing a large error when a coral colony is not of an elliptic shape. (2) Improved online calibration might prevent camera poses drift. Better reconstruction could be achieved by using color correction on current video data.

In conclusion, the proposed method is promising to provide ecological information of coral reef cheaply and efficiently. It can be used for old videos conducting temporal comparison for the same coral colony which is critical for accurately measuring coral growth and bioerosion (McKinnon et al. 2011). Furthermore, it might be incorporated into existing monitoring protocols to obtain accurate measurements of rugosity and structural complexity.

References

Brock JC, Wright CW, Clayton TD, Nayegandhi A (2004) Lidar optical rugosity of coral reefs in biscayne national park, florida. Coral Reefs, 23: 48-59

Bridge TC, Done TJ, Beaman RJ, Friedman A, Williams SB, Pizarro O, Webster JM (2011) Topography, substratum and benthic macrofaunal relationships on a tropical mesophotic shelf margin, central Great Barrier Reef, Australia. Coral Reefs, 30: 143-153

Friedman A, Pizarro O, Williams SB (2010) Rugosity, slope and aspect from bathymetric stereo image reconstructions. OCEANS, Sydney, pp 1-9

Furukawa Y, Ponce J (2010) Accurate, dense, and robust multi-view stereopsis. PAMI, 32: 1362-1376

Harvey E, Fletcher D, Shortis M (2001) A comparison of the precision and accuracy of estimates of reef fish lengths determined visually by divers with estimations produced by a stereo video system. Fisheries Bulletin, 99, 63-71

Kramer PA (2003) Synthesis of coral reef health indicators for the western Atlantic: Results of the AGRRA program(1997-2000). Atoll Research Bulletin, 496: 1-57

Lirman D, et al. (2007) Development and application of a video-mosaic survey technology to document the status of coral reef communities. EMS, 125: 59-73

McKinnon D, He H, Upcroft B, Smith R (2011) Towards automated and in-situ, near realtime 3D reconstruction of coral reef environment. Oceans, Hawaii

McCoy ED, Bell SS (1991) Habitat structure:the evolution and diversification of a complex topic. Habitat Structure:the Physical Arrangement of Objects in Space. Chapman&Hall, London, pp 3-27

McCormick MI (1994) Comparison of field methods for measuring surface topography and their associations with a tropical reef fish assemblage. MEPS, 112: 87-96

McClanahan T, Muthiga N, Mangi S (2001) Coral and algal changes after the 1998 coral bleaching: interaction with reef management and herbivores on Kenyan reefs. Coral reefs, 19: 380-391

Pollefeys M, Koch R, Gool LV (1998) Self-calibration and metric reconstruction in spite of varying and unkown internal cameras parameters. ICCV, India, pp 90-95

Renken H, Mumby PJ, Matsikis I, Edwards HJ (2010) Modelling the dynamics of coral reef macroalgae using a Bayesian Belief Network approach. Ecological Modelling, 220: 1350-1314

Risk MJ (1972) Fish diversity on a coral reef in the Virgin Islands. Atoll Res. Bull. 153: 1-6

Roberson MJ, Pizarro O, Williams SB, Mahon I (2010) Generation and visualization of large scale three dimentional reconstructions from underwater robotic surveys, JFR, 27: 21-51

Spalding M, Ravilious C, Green E (2001) World Atlas of Coral Reefs. University of California Press, Berkely, USA

Snavely N, Seitz SM, Szeliski R (2006) Photo tourism: exploring photo collections in 3D. SIGGRAPH, Boston, pp 835-846