Toward Robust Image Detection of Crown-of-Thorns Starfish for Autonomous Population Monitoring

Ryan Clement†, Matthew Dunbabin‡ and Gordon Wyeth†
†School of Information Technology and Electrical Engineering, The University of Queensland.
‡Autonomous Systems Laboratory, CSIRO ICT Centre, P.O. Box 883, Kenmore, QLD 4069, Australia.
matthew.dunbabin@csiro.au, wyeth@itee.uq.edu.au

Abstract

Robust texture recognition in underwater image sequences for marine pest population control such as Crown-Of-Thorns Starfish (COTS) is a relatively unexplored area of research. Typically, humans count COTS by laboriously processing individual images taken during surveys. Being able to autonomously collect and process images of reef habitat and segment out the various marine biota holds the promise of allowing researchers to gain a greater understanding of the marine ecosystem and evaluate the impact of different environmental variables. This research applies and extends the use of Local Binary Patterns (LBP) as a method for texture-based identification of COTS from survey images. The performance and accuracy of the algorithms are evaluated on a image data set taken on the Great Barrier Reef.

1 Introduction

The Great Barrier Reef is one of Australia's most well known assets. It provides income through the hundreds of thousands of tourists that visit the reef each year bringing in an estimated $14 billion[1] per annum, and is also home to millions of different species of marine animals. It is vital for the longevity of the reef to be able to quickly and accurately monitor how well the ecosystem can maintain both the natural and human influences that affect the reef each year.

The Great Barrier Reef is also one of the world's most environmentally sensitive areas. In the last 40 years there have been numerous outbreaks of Crown-Of-Thorns Starfish observed. Crown-Of-Thorns Starfish (COTS) are a particular natural pest that feeds on coral and in the process kills it. This is shown in Figure 1 where the dead coral appears bleach white until algae begins to grow on it. If COTS reach outbreak proportions (approximately 30 per hectare [Engelhardt and Lassig, 1996]), they can rapidly cause extensive damage to large areas of coral reef.

[1] Tourism Queensland Annual Report 2001-2002

Figure 1: Crown-Of-Thorns Starfish (COTS) on coral.

Many environmental organisations, Universities, Government agencies, and members of public have provided extensive resources to understand, monitor and control COTS outbreaks. However, new outbreaks occur in cycles of 1-15 years, making it difficult to find direct causes, or even control COTS numbers.

Current monitoring of this problem is conducted entirely by humans and is very prone to error. There are two primary techniques to count COTS. In the first method, divers record video of a section of reef and humans later count COTS when reviewing the video. The second method is called a Manta Tow Survey [Bass and Miller, 1995] whereby a human is towed behind a boat and attempts to count the number of animals they see. This number is recorded and then extrapolated to approximate the distribution over many square kilometers. This technique is inaccurate as research has shown [Moran and De'ath, 1992] that it underestimates populations and can quite possibly overlook major populations of COTS when the particular transect undertaken was unrepresentative of the region.

Considering the enormous size of the reef (349,000 square kilometers) the current methods of monitoring COTS numbers is unable to generate accurate and representative large scale distribution maps. To perform this task on the scale of the entire reef is excessively expensive, with the costs of divers, ships, and the hindrance of weather and health and safety issues. Therefore, it makes sense to devise an alternate, cost effective and safe method of performing these tasks. The proposed research is to perform autonomous population monitoring involving two key activities: (1) having a vehicle autonomously collect the images, and (2) being able to perform image recognition of COTS. The first activity involves using an Autonomous Underwater Vehicle (AUV) to perform video transects of the reef as shown in Figure 2 and described in other research [Dunbabin *et al.*, 2005a; 2005b]. The second activity is the focus of this paper.

Figure 2: The Starbug AUV during reef trials.

The second activity requires the robust segmentation (recognition) of COTS from an arbitrary image. Color segmentation is not reliable as COTS vary in color considerably depending on their age, location and altitude at which the image was taken. COTS do however have distinctive thorns which leads to the potential use of texture as a reliable feature for the recognition of the starfish. A number of methods were reviewed and tested for this role being namely Local Binary Patterns (LBP), Gabor Wavelets and the Hough Transform. Initial testing showed that LBP had the greatest potential by accurately detecting the most COTS with greatest rejection rate of the rest of the image. LBP's have previously been used for coral detection [Soriano *et al.*, 2001] with accuracies of around 40% stated.

Pilot studies on the use of Gabor Wavelets and Hough Transforms on small data sets were not promising. Therefore, the focus of this paper is on LBP classification with a discussion on these techniques given in Section 4.

This paper describes the preliminary investigation into the use of LBP's as a method for texture mapping and correlation of COTS from survey images. A large image data set was obtained from the Australian Institute of Marine Research (AIMS) which contained COTS in various poses. The performance and accuracy of the algorithms are evaluated on the image data set. This paper focuses on the detection of 'blobs' of starfish texture and not the process of how to differentiate between two or more touching animals.

1.1 Paper Outline

The remainder of this paper is structured as follows: Section 2 describes the texture-based classification procedure developed using LBP with Section 3 presenting results on the classification performance for identifying COTS from the image set obtained from AIMS. Finally, Section 4 will discuss the issues encountered during this research and focus of future research.

2 Texture-based Marine Biota Classification

Texture recognition is the most suitable form of segmentation when performing underwater biota identification. The primary reason for this comes down to the environment. In underwater environments red wavelengths of light get absorbed much faster then the shorter (i.e. blue) wavelengths. This means that depending on the depth, the environment will generally appear a blue/green colour. Di Gesu et.al [Di Gesu, 2003] present a method of starfish identification and tracking using color segmentation and background subtraction. These starfish are generally brightly colored with respect to their surroundings and have a regular shape. COTS on the other hand are reasonably camouflaged and highly flexible creatures that are found in varying poses (e.g. flat, curled, partially hidden) making shape or template matching inadequate in this particular instance. The thorns however, are the most distinguishing and recognisable feature (texture) of COTS and will be utilised for this automated detection scheme.

2.1 Local Binary Patterns

This is a greyscale invariant method of detecting textures. This method has been used previously with limited success in the detection of coral in underwater video sequences [Soriano *et al.*, 2001]. To achieve greyscale invariance, a pixel "g_c" is selected and its greyscale value subtracted from each of the neighbouring pixels (g_p). Normalisation is then achieved by the addition of these differenced neighbouring pixels. However, it is still possible to observe large greyscale shifts. To overcome this scenario, only the signs of the neighbouring pixels are chosen [Ojala *et al.*, 2002]. If a pixel is greater or equal to zero it is assigned '1', otherwise '0'. Therefore the basic LBP can be define as Equation 1.

$$LBP = \sum_{p=0}^{P-1} s\left(g_p - g_c\right) = \sum_{p=0}^{P-1} \left(g_p \geq g_c\right) \qquad (1)$$

where $s(x)$ is defined as

$$s(x) = \begin{cases} 1, \; x \geq 0 \\ 0, \; x < 0 \end{cases} \qquad (2)$$

While this equation is greyscale invariant, it does not provide enough extra information for adequate correlation when comparing with images. Studies have shown that there are a number of ways to get rotationally invariant information [Ojala *et al.*, 2002; Pietikainen *et al.*, 2000]. Equation 3 shows an alternate LBP with improved rotational invariance. Ojala [Ojala *et al.*, 2002] concluded this as the best method for LBP as it improved discrimination by only using uniform patterns rather than all patterns. Uniform patterns are defined as being able to switch from 1 to 0 once and 0 to 1 once. This eliminates problems arising from high frequency patterns such as 10101010.

$$LBP_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s\left(g_p - g_c\right), U\left(LBP_{P,R}\right) \leq 2 \\ else, \;\; P+1 \end{cases} \tag{3}$$

where

$$\begin{aligned} U\left(LBP_{P,R}\right) &= \left|s\left(g_{P-1} - g_c\right) - s\left(g_0 - g_c\right)\right| \\ &+ \sum_{p=1}^{P-1} \left|s\left(g_p - g_c\right) - s\left(g_{p-1} - g_c\right)\right| \end{aligned}$$

In Equation 3, R specifies the radius from the center pixel and P specifies how many positions are available in the pattern. This means when considering a pixels nearest neighbours $P = 8$. If the pattern is not uniform, then it is assigned an invalid number $(P+1)$.

An optimised method of calculating the LBP based on the work of Pietikainen [Pietikainen *et al.*, 2000] was implemented in this investigation. This optimisation is especially beneficial when using Matlab which performs matrix operations much faster than looped operations, however, it requires more memory. The improvement in processing speed significantly lowered the LBP generation time from 15 minutes to approximately 20 seconds. The algorithm works by creating P copies of the image. The size of the image is increased by padding an extra R zeros around the outside of the first image. In the case of the nearest neighbour solution ($R = 1$), one layer of zeros are placed around the outside of the image as shown in Figure 3. The subsequent copies are positioned such that each image is shifted R pixels around the central image to create the circle. Again the nearest neighbour solution means that the image is moved to occupy eight positions. Once this has occurred, the output image is simply an addition of matrix logical comparisons to see if the pixels are greater than the centred image pixels.

2.2 Texture Database

A texture database of 384x384 pixel textures was created from representative images. A total of 27 base textures were created of which 12 were COTS and 15 were desired reject textures of various corals and other marine surfaces. Each texture has its Local Binary Pattern created and stored for use in image classification.

It was observed that instead of looking for changes of the circular radius for multi-scale analysis, improved results could be obtained by scaling the sample textures. Therefore, in this analysis 3 different texture scales were used; one that was nearest neighbour, one that was down sampled by 25% before applying nearest neighbour, and the final up sampled by 25%. This has obvious limitations when images are at high altitude which is the focus of future research.

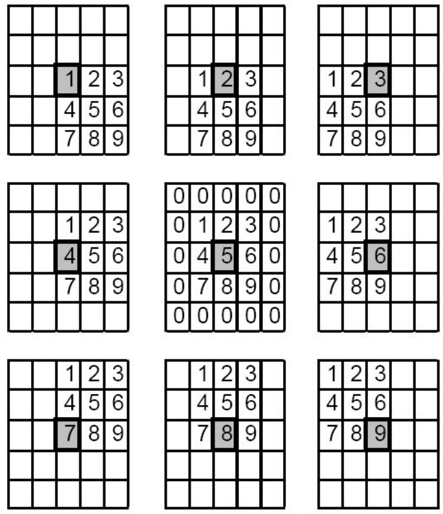


Figure 3: Optimised matrix method for LBP calculation.

It was also found that some sample coral textures had similar histograms to representative COTS textures. Figure 4 shows a series of sample textures and their respective histograms. Although some histograms appear very similar, the combinations from just nearest neighbour and the multi resolution approaches assist to differentiate the different histograms and provide a ‘good’ solution. It is also interesting to not that absence of data in the ninth bin indicates that the textures contain only uniform patterns.

2.3 Texture Matching

A log-likelihood measure is used for comparing the similarity between textures as described by

$$L\left(S,M\right) = \sum_{b=1}^{B} S_b \log M_b \tag{4}$$

where B is the number of bins in the histogram, S_b and M_b correspond to the sample and model probabilities at bin b respectively. In this implementation 9 bins were chosen to differentiate between various textures. However, this statistic is not quite adequate for scale invariant solutions so it has been extended [Topi *et al.*, 2000] to accommodate for the extra radii calculated for the LBP as described by Equation 5.

$$L\left(S,M\right) = -\sum_{h=1}^{H} \sum_{n=1}^{N} \frac{T_{hs} S_{hn}}{\sum_h T_{hs}} \ln \frac{T_{hm} M_{hn}}{\sum_h T_{hm}} \tag{5}$$

where S_{hn} and M_{hn} correspond to the probabilities of bin n in the h^{th} sample and model histogram, respectively. T_{hs} and T_{hm} denote the total number of entires used in producing sample and model histogram h, respectively. Using this classification method it is possible to determine the ‘best’ matched texture for a particular region on the screen.

2.4 Process

Once an appropriate texture database has been generated as described in Section 2.2, the complete image processing technique consists of the 6 following steps:

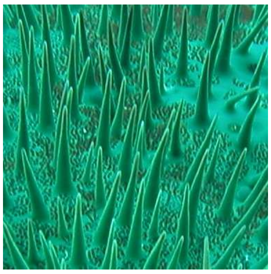
1. Top Hat filtering with disc structuring element
2. Grey scale conversion (MATLAB)
3. Local binary pattern created
4. Histograms created
5. Log-likelihood measure performed on image blocks
6. Count number of detected COTS ‘blobs’.

Steps 1-3 are performed once on the whole image, whilst steps 4-6 are performed in loop-fashion on sub-sections of the image. The first step, Top Hat filtering, is an operation that subtracts the original image from an “opened” version of the image. An “opened” image is defined as the erosion of the image using a structuring element and then dilation of resulting image with the same element. This operation can usually highlight shadowed regions in images and is performed to enhance the LBP and further reduce major greyscale shifts. Figure 5 shows a representative image before and after Top Hat filtering.

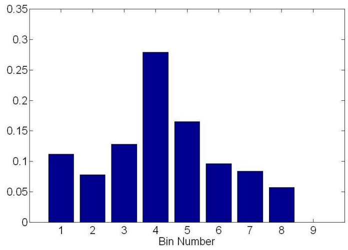
The LBP was computed once for the entire image, however, for scale invariance, multiple radii are required. In this investigation, it was decided to have multiple radii on the texture database rather than on the input image as these can be precomputed, significantly speeding up the entire image process time. Figure 6 shows the LBP of the Top Hat filtered image of Figure 5.

The first of the looped operations on the image is the creation of histograms for texture comparison. These histograms were first computed in 2D and then summed together to create the nine bins. When computing a histogram, a 50x50 pixel block is selected with an overlap region of 200 pixels applied to create a 450x450 pixels region. This large block size enables the algorithm to classify greater sections of COTS especially around the COTS leg area where it can change from COTS to coral then back to COTS again. The log-likelihood measure is then used to find the best matched texture for this region. When matching the histograms against the textures, the resulting output region is only the central 50x50 pixel block. The entire process is computed again with the central block shifted 50 pixels and repeated across the entire image.

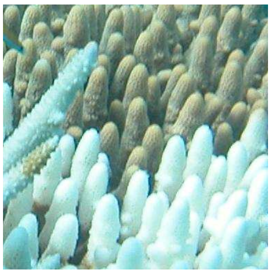
The resulting binary output from texture comparison is checked to see if any holes exist within an identified object. The likelihood that multiple COTS touching form a ‘donut’ shape is minimal, so it is reasonably valid to ‘fill in’ any closed holes that arise. Figure 7 shows the results of a COTS classification before and after the central region has been filled by the algorithm.



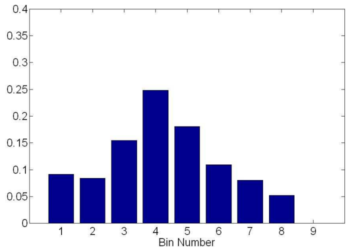
(a) COTS



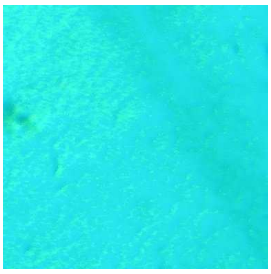
(b) COTS histogram



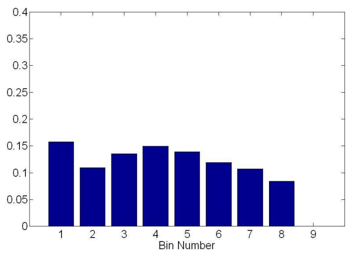
(c) Coral



(d) Coral histogram



(e) Sand



(f) Sand histogram

Figure 4: Sample textures and histograms for COTS, coral and sand.

(a) Original image

(b) Image after Top Hat filtering

Figure 5: Effect of top Top Hat filtering on reef image.

Additionally, a minimum ‘blob’ size constraint was added to minimise detection of random blocks. If a region is connected with 5 or less blocks it is deemed to be an unlikely texture. This problem of misclassification often arises with changes of depth in the water. The proposed AUV that will collect images for this processing is capable of altitude control, therefore, the size of COTS in the image would be more consistent than those from this investigation.

3 Results

The data set used for this analysis was sourced from the Australian Institute of Marine Science (AIMS). A total of 80 images were obtained both with and without starfish in the image. The images used were approximately 3 Mega pixels in size and taken from various angles, distances and depths from the COTS and at different image and quality (focus).

The algorithm described in Section 2 was applied to all images with the number of COTS detected by the algorithm and the area of classified image (in pixels) recorded along with the actual number of animals obtained from individual analysis of each image. Figure 8 shows the results of the detected number of starfish compared with the corresponding actual number of starfish in the scene for all 80 images.

Figure 6: Local Binary Pattern Output.

As seen from Figure 8, the algorithm resulted in accurately detecting the presence, or non-presence, of a starfish in 65% of the images and correctly counted the number of starfish in 48% of the images. The output from a correctly classified COTS is shown in Figure 9(a). Later analysis of the images for which the algorithm failed to detect any COTS (17% of the images), showed that they were slightly out of focus or taken at high altitude compared to the other images, potentially contributing to the poor detection in these images.

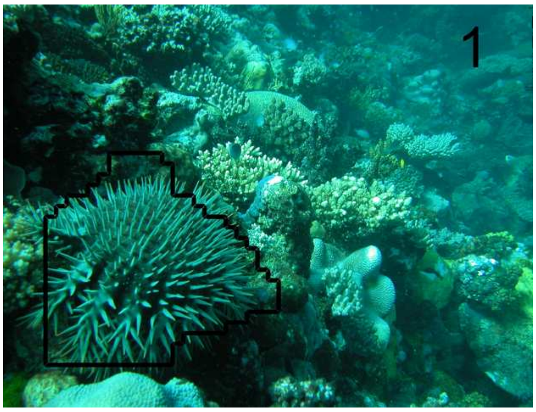
The algorithm did however, falsely classify some non-starfish textures as starfish. The percentage of a false positive classification was 31%. An example of false positive classification is shown in Figure 9(b). In general, false positive classification resulted from images taken at relatively high altitudes. The deeper the image the more likely chance of misclassification.

Figure 10 shows the area of image that is classified at COTS compared to the actual area for all the images. The algorithm on average found 49% of the actual area identified as COTS with a median of 65%. This average is biased by the results where the algorithm did not detect a COTS. Discarding, those cases of non-detection, the algorithm on average correctly classifies 77% of the actual COTS in the image.

The current algorithm implementation is written in MATLAB and takes just over 2 minutes to run per image (Pentium 4 3GHz, 1024MB RAM). The bulk of the processing is actually contained in computing and comparing histograms and not in the creation of the Local Binary Pattern. The final stage that removes the small blocks, counts the number of correct pixels and overlays the number on the image takes the largest proportion of time approximately one third of total time. The following list describes the key steps and the percentage of total processing time required to perform the task:

1. Top hat filtering, resizing and rgb2gray (23.6%)
2. LBP calculation (5.2%)
3. Calculating histogram (22.5%)
4. Compare textures (7.7%)

(a) Before



(b) After

Figure 7: Classification results of COTS before and after the central unclassified region has been filled.

5. Count and render to screen (41.0%)

From this it can be seen that rendering to screen takes a significant proportion of the processing time whilst also taking a large amount of memory thus slowing the system as it has to page to get the required memory.

4 Discussion

Whilst the above analysis is far from exhaustive, it has highlighted a number of conditions that are required to obtain best COTS identification accuracy. The classes of images for which the algorithms were observed to performed less accurately were:

1. Images taken at altitude where the COTS and its texture are a small proportional of the overall image (e.g. see Figure 11(a)).
2. Textures whose LBP's look similar to COTS (e.g. see Figure 11(b)).
3. Images that are slightly out of focus.

The first condition can be moderated by performing controlled survey transects such as those obtained using an AUV, whereby reducing the altitude from the camera to the seafloor and increasing the size of objects in the image. Addressing the second problem is more difficult and the area of current research. Image quality is important, and must be addressed at the hardware level to ensure the thorns are clear and sharp.

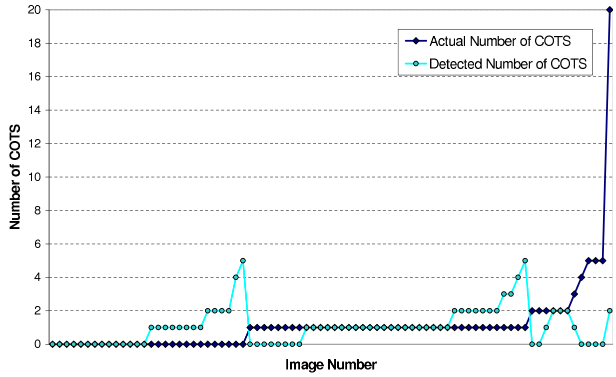


Figure 8: Number of COTS detected and counted for all images.

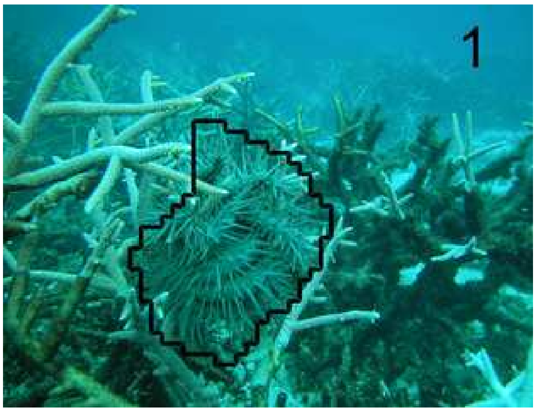
Another observed deficiency of the current processing method is identifying (counting) multiple COTS in an image that are touching each other. Figure 12 shows a processed image where the interconnected COTS were correctly identified, however, the counting routine only identified two animals when there is actually 5 in the scene. Improving the counting routine is a current area of research and will be required if accurate population monitoring is to be obtained.

Preliminary investigations have also been undertaken into alternate image processing techniques to improve COTS classification. These were:

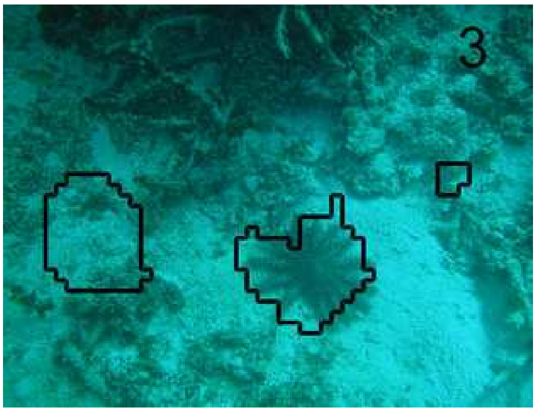
1. Gabor wavelets
2. Hough transforms.

The initial failure of Gabor Wavelets is believed to be credited to the environment in which the photos were taken. Since lighting is different in all images it is hard to match textures accurately all the time. However, there has been some success in other areas such as Arial photography [Manjunath and Ma, 1996] which warrants continued research in this area.

The unique characteristic of the starfish are its straight thorns, which has lead to the development of a technique to find these straight lines. Taking a Hough transform of the image and finding intersecting lines of the size and density of the thorns is highly dependant on the image processing being able to resolve sharply the thorns of the COTS. If the depth of COTS is too deep, the thorns are no longer clearly defined to be detected. An initial investigation of this technique was promising but also falsely classified shelf or stag coral. The straight lines from this type of coral would be detected as thorns and cause the algorithm to have many false classifications.



(a) Correctly detected COTS



(b) False positives

Figure 9: Classification results of COTS in scene.

5 Conclusions

This paper presented and evaluated the application of Local Binary Pattern's as simple but powerful feature descriptor for classifying Crown-Of-Thorns Starfish from an image. The algorithm presented was applied to a set of 80 images obtained from the Great Barrier Reef. The results show that it could correctly detect the presence of a COTS in 65% of the images which had COTS in them. The algorithm was on average also able to find 77% of the total COTS in the image when a starfish was detected.

The technique was found to perform less accurately on images that were taken at altitude where the texture and size of the starfish was small compared to the image resolution, and when the image was slightly out of focus.

Future research is expanding the classification scheme to accommodate greater image altitude, poorer image quality and greater rejection of similar looking corals. Additionally, the system will be integrated into the AUV for autonomous population monitoring.

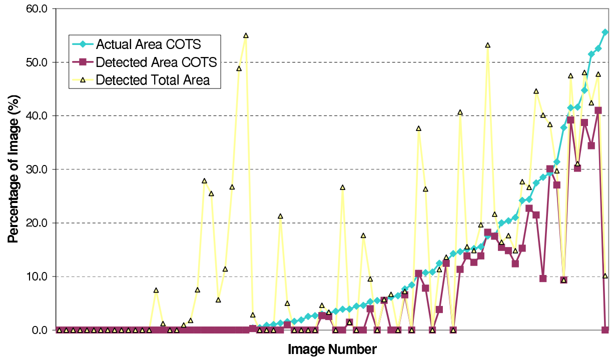


Figure 10: Percentage of total image area detected as COTS.

Acknowledgments

The authors would like to thank Kate Osborne from the Australian Institute of Marine Science (AIMS) for supplying many of the original images used in this study.

References

[Bass and Miller, 1995] D.K. Bass and I.R. Miller. Crown-of-thorn starfish and coral surveys using the manta tow and SCUBA search techniques. Technical Report Standard Operational Procedure Number 1, Australian Institute of Marine Science, 1995.

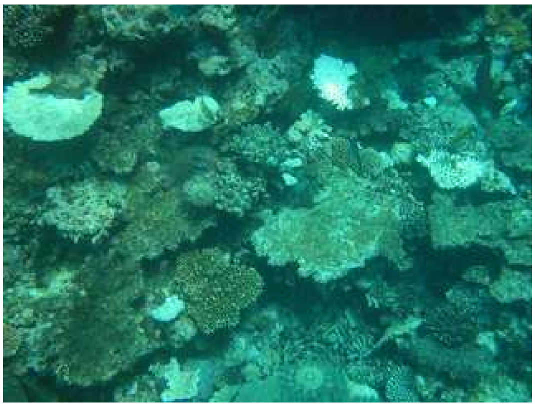
[Di Gesu, 2003] V. et.al. Di Gesu. Finding essential features for tracking starfish in a video sequence. In *Proceedings of the 12th International Conference on Image Analysis and Processing*, 2003.

[Dunbabin *et al.*, 2005a] M. Dunbabin, J. Roberts, K. Usher, G. Winstanley, and P. Corke. A hybrid AUV design for shallow water reef navigation. In *Proc. International Conference on Robotics and Automation*, pages 2117–2122, Barcelona, April 2005.

[Dunbabin *et al.*, 2005b] M. Dunbabin, K. Usher, and P. Corke. Visual motion estimation for an autonomous underwater reef monitoring robot. In *Proc. International Conference on Field & Service Robotics*, pages 57–68, Port Douglas, 2005.

[Engelhardt and Lassig, 1996] U. Engelhardt and B.R. Lassig. A review of the possible causes and consequences of outbreaks of the crown-of-thorns starfish (acanthaster planci) on the great barrier reef-an australian perspective. In *Proceedings of The Great Barrier Reef - Science, use and management (Volume 1: Invited Papers)*, pages 243–259, 1996.

[Manjunath and Ma, 1996] B.S. Manjunath and W.Y Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:837–842, 1996.



(a) High altitude image

(b) Coral with similar texture to COTS

Figure 11: Typical image classes causing inaccurate COTS classification.

[Moran and De'ath, 1992] P.J. Moran and G. De'ath. Suitability of the manta tow technique for estimating relative and absolute abundance of crown-of-thorn starfish (acanthaster planci). *Aust. J. Mar. Freshwater Res.*, 43:357–378, 1992.

[Ojala *et al.*, 2002] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:971–978, 2002.

[Pietikainen *et al.*, 2000] M. Pietikainen, T. Ojala, and Z. Xu. Rotation-invariant texture classification using feature distributions. *Pattern Recognition*, 33:43–52, 2000.

[Soriano *et al.*, 2001] M. Soriano, S. Marcos, C. Saloma, M. Quibilan, and P. Alino. Image classification of coral reef components from underwater color video. In *Proceedings of the Oceans 2001 MTS/IEEE - An Ocean Odyssey*, Honolulu, HI., November 5-8 2001.

[Topi *et al.*, 2000] M. Topi, P. Matti, and O. Timo. Texture classification by multi-predicate local binary pattern operators. In *Proceedings of the 15th International Conference on Pattern Recognition*, 2000.

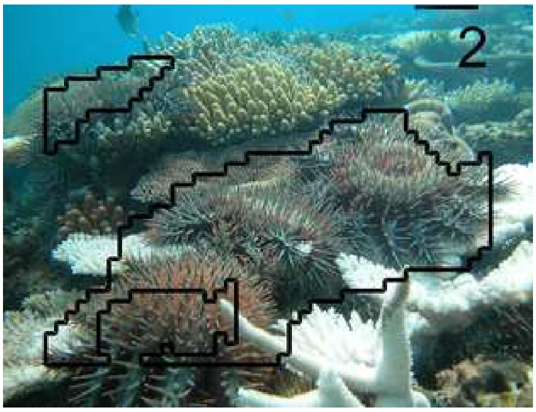


Figure 12: Typical image where number COTS texture has been correctly identified, but number of animals incorrectly counted.