Kelp Detection in Highly Dynamic Environments Using Texture Recognition

Aymeric Denuelle[†‡] **and Matthew Dunbabin**[†]
[†]Autonomous Systems Laboratory, CSIRO ICT Centre, P.O. Box 883, Kenmore, QLD 4069, Australia.
[‡]University of Burgundy, Dijon, France.
adenuelle@gmail.com, matthew.dunbabin@csiro.au

Abstract

This paper describes a texture recognition based method for segmenting kelp from images collected in highly dynamic shallow water environments by an Autonomous Underwater Vehicle (AUV). A particular challenge is image quality that is affected by uncontrolled lighting, reduced visibility, significantly varying perspective due to platform egomotion, and kelp sway from wave action. The kelp segmentation approach uses the Mahalanobis distance as a way to classify Haralick texture features from sub-regions within an image. The results illustrate the applicability of the method to classify kelp allowing construction of probability maps of kelp masses across a sequence of images.

1 Introduction

Some marine researchers would like to understand the factors that affect kelp recruitment and dieback resulting from different environmental patterns such as storms, cyclones and climate change. Collecting accurate data to classify kelp and seagrass in these areas is difficult. Regions of interest include on and around near shore reefs in 6-10 m water depth with significant swell. These conditions can result in low visibility, turbulent water at the seafloor causing the kelp to sway, and considerable natural lighting variation.

Various techniques have been proposed and demonstrated to characterise kelp and seagrass habitats. These include broader scale techniques such as satellite remote sensing [Dekker *et al.*, 2005], airborne imagery [Lathrop *et al.*, 2006], and Multi-beam sonar [Ierodiaconou *et al.*, 2010]. Recently, Roelfsema et al. [Roelfsema *et al.*, 2009] conducted a seagrass coverage study in Moreton Bay using satellite remote sensing, photo transects, and fine-scale survey. They conclude that although the field surveys provided most accurate local variability results, they are too costly for larger areas meaning a reliance on broader scale methods to understand region.

Autonomous Underwater Vehicles (AUVs) are emerging as a tool for obtaining larger habitat distribution maps with very fine detail capable of quantifying change with time. Although most studies have focused on coral reef environments, recent studies such as [Moline *et al.*, 2007] have used multi-spectral radiometers attached to an AUV to create distribution maps for eelgrass with promising results.

In March 2009, a small AUV (shown in Figure 1 [Dunbabin and Allen, 2007]) was deployed around Marmion Reef, near Perth, Western Australia to collect images of the kelp on the seafloor. The images were of relatively low resolution (640x480 pixels) captured by the downward facing navigation cameras. The test location is a highly dynamic environment resulting from swell and currents. This caused the AUV to experience significant roll and pitch during the survey. Additionally, the turbulence reduced visibility and wave action caused highly variable natural lighting which additionally affected image quality.

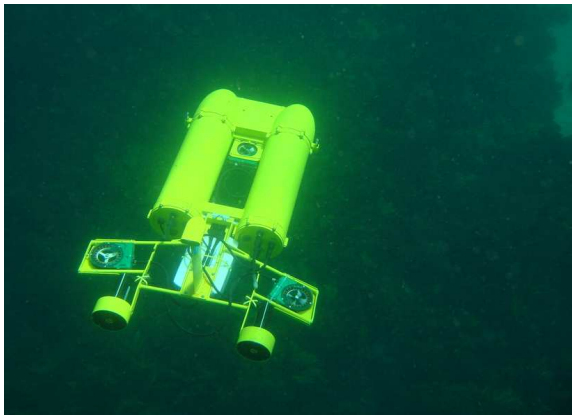


Figure 1: The Starbug AUV during image collection surveys of kelp beds at Marmion Reef.

The work described here is focused on the robust detection of kelp from these images whose quality is affected by factors resulting from operation within these dynamic environments. The objective is to obtain a kelp detection rate equivalent to manual extraction with an ultimate goal of automatically localising, monitoring and quantifying the distribution of kelp forest growth across the entire region.

1.1 Paper Outline

The remainder of this paper is structured as follows; Section 2 provides an overview of the Haralick texture features with Section 3 describing clustering methods examined to classify kelp from images. Section 4 details the methodology for generating kelp probability maps using Haralick texture features from sliding window across the entire image. A qualitative and quantitative assessment of the methodology using images collected by the AUV are shown in Section 5, and finally Section 6 concludes the paper.

2 Haralick Texture Features

This section describes the methods used to extract some robust features from RGB images allowing characterization of texture from particular regions within the images.

2.1 Gray Level Co-occurrence Matrix (GLCM)

A co-occurrence matrix describes the patterns of neighbouring pixels in an image at a given distance. It consists of computing two-dimensional histogram matrices for different orientations of pixels pairs across the image. By varying orientation and the distance between two neighbour pixels, a rotation-invariant and multi-scale method is obtained to approximate the joint probability distribution of a pair of pixels.

Mathematically, a co-occurrence matrix of a greyscale image $I(z) = I(x,y)$ (coded on N grey levels) dealing with pixels pairs in $I(z)$ separated by a translation vector $t = (\Delta x, \Delta y)$ is defined as:

$$M_t(i,j) = \#\left\{(z, z+t) \in \Re^2 / I[z] = i,\ I[z+t] = j\right\} \quad (1)$$

The co-occurrence matrix M_t is a NxN matrix whose (i,j)th element is the number of pixels separated by the translation vector t that have the pair of grey levels (i,j).

2.2 Multi-spectral GLCM

A multi-spectral variation to the GLCM computation to process RGB images was introduced by Arvis et al. [Arvis *et al.*, 2004]. In this method, pixel pairs are found over each of the 3 colour channels. A multi-spectral co-occurrence matrix represents the total number of pixel pairs in the colour source image having a colour value i from the colour channel α, and a colour value j from the colour channel β, separated by a vector t. This was defined in the following way:

$$M_{t(\alpha,\beta)}(i,j) = \#\left\{(z, z+t) \in \Re^2 / \alpha[z] = i,\ \beta[z+t] = j\right\} \quad (2)$$

By definition, one co-occurrence matrix is obtained per pair of channels (e.g. RR, RG, RB, GG, GB and BB).

2.3 Haralick texture features

Using first order texture features such as standard deviation and variance is not sufficient to properly characterize the texture of a small area in our original image. This observation leads us to use other statistical moments.

Once the co-occurrence matrix has been computed, statistical properties can be extracted and then Haralick texture features [Haralick *et al.*, 1973] computed. From the co-occurrence matrix $M(i,j)$ which contains the texture information, Haralick defined 13 features which can be calculated directly. Only five of these are commonly used because it was shown that the 13 texture features are very correlated. The five features used are uniformity, contrast, correlation, local homogeneity and entropy defined by:

$$\begin{aligned}
f1 &= \sum_{i=1}^{N}\sum_{j=1}^{N} m(i,j)^2 \\
f2 &= \sum_{k=0}^{N-1} k^2 m_{x-y}(k) \\
f3 &= \frac{1}{\sigma_x \sigma_y}\sum_{i=1}^{N}\sum_{j=1}^{N}(ij)m(i,j) - \mu_x\mu_y \\
f5 &= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{1+(i-j)^2}m(i,j) \\
f9 &= -\sum_{i=1}^{N}\sum_{j=1}^{N} m(i,j)\log(m(i,j))
\end{aligned}$$

where

$$m(i,j) = \frac{M(i,j)}{R} \quad ; \quad R = \sum_{i=1}^{N}\sum_{j=1}^{N} M(i,j)$$

$$m_x(i) = \sum_{j=1}^{N} m(i,j) \quad ; \quad m_y(j) = \sum_{i=1}^{N} m(i,j)$$

$$\mu_x = \sum_{i=1}^{N} i m_x(i) \quad ; \quad \mu_y = \sum_{j=1}^{N} j m_y(j)$$

$$\sigma_x = \sqrt{\sum_{i=1}^{N} m_x(i)(i-\mu_x)^2} \quad ; \quad \sigma_y = \sqrt{\sum_{j=1}^{N} m_y(j)(j-\mu_y)^2}$$

$$m_{x-y}(k) = \sum_{i=1}^{N}\sum_{j=1}^{N} m(i,j)$$

such as $|i-j| = k$ where $k = 0, 1, 2, ..., N-1$.

2.4 Co-occurrence matrices and Haralick features computation

By definition, co-occurrence matrices are symmetric so it is sufficient to describe half of the space around one pixel to get the whole neighbourhood. Consequently, to achieve a degree

of rotation invariance, the co-occurrence matrix is computed matching the angles 0°, 45°, 90° and 135°, and by symmetry a total of 8 co-occurrence matrices are summed to obtain one final rotation-invariant matrix.

As the data set consists of multi-scale images, it is necessary to use several radii to build the co-occurrence matrix (four were chosen in this case). As a result, 4 rotation-invariant co-occurrence matrices are computed for each pair of colour channels. As the red channel is highly absorbed underwater and dependant on the altitude and depth at which the image was taken, we ignore the contribution of this channel to compute the multi-spectral co-occurrence matrix and only consider the co-occurrence matrices within each of the green and blue channels.

This method leads to the computation of 60 Haralick texture features per image (5 features x 4 rotation-invariant co-occurrence matrices (different radii) x 3 pairs of colour channels).

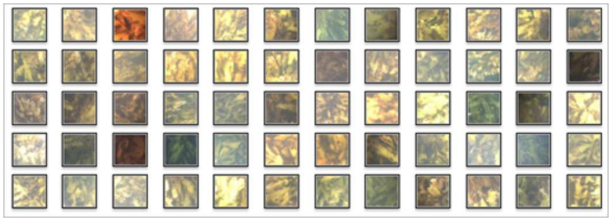
Several methods can be applied to considerably reduce the feature computation time using the properties of the rotation-invariant co-occurrence matrices jointly with algorithmic methods [Miyamoto and Merryman, 2005]. As the number of grey levels used increases, the co-occurrence matrix calculation becomes exponentially time consuming. Therefore, feature computation is only calculated on 32 gray-level co-occurrence matrices. Furthermore, using co-occurrence matrix symmetry, it is easy to obtain the following equalities: $\mu_x = \mu_y$ and $\sigma_x = \sigma_y$.

A second approach which achieves a significant reduction in Haralick feature computation time is via an efficient manner in which the calculations are performed. That is by gathering the features which loop across the data in similar ways such as combining the calculation of $f1$, $f3$, $f5$ and $f9$.

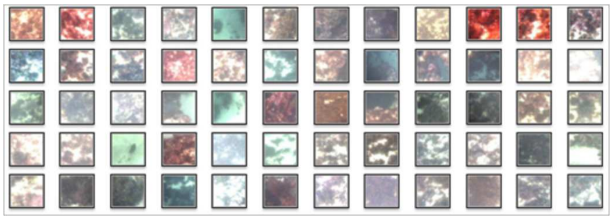
3 Feature Clustering

The classification parameters used are the Haralick texture features (described in Section 2) extracted from a series of representative kelp and non-kelp 100x100 pixel sub-images obtained from the entire data set. It was important to use a large training set with a mix of bright and dark samples taken in poor, clear conditions with different degrees of red channel absorption as shown in Figure 2.

All images taken by the AUV are georeferenced by the variables roll, pitch, raw, altitude and depth. This information can be used to bootstrap the clustering procedure for the size and perspective of texture and red light attenuation. Therefore, to provide a "scaling" factor for texture features, the previous 60 Haralick texture features are augmented with altitude information from the AUV creating a 61-dimensional vector for each training image. Combining the 61-dimensional vectors from all training images, two centroids (61-dimensional vectors) representing the kelp and non-kelp classes can be created.



(a) Kelp texture samples.



(b) Non-kelp texture samples

Figure 2: Example clustering training set of kelp and non-kelp 100x100 pixel sub-images extracted from the entire data set.

An important consideration before processing any feature clustering operation is that of normalization. Several normalization methods exist although they are generally based on properties extracted from the entire data set. Therefore, as during the classification phase it is impractical to normalize the sample against all the other samples, a methods not requiring any normalization of the data is desired. The following sections describe two potential methods of clustering depending on the application requirements.

3.1 Unsupervised Clustering

The *k*-means algorithm was considered for an unsupervised clustering method of the available training samples. In order to improve the results given by the *k*-means algorithm, Principal Component Analysis (PCA) is first applied on the raw training data vectors. Using PCA before building the reference kelp cluster reduces the possible number of errors during the clustering phase.

However, the Euclidian distance cannot be used with non-normalized texture features. Indeed, due to the large numerical values of certain features, the computed distance would be dominated by these. As a proper normalization of each new feature vector is not possible (accordingly to the normalization done during the training phase), the largest features are divided by an appropriate factor in order to scale them within the range [0:1].

Finally, the two centroids are determined that are used to measure the distance between the current features vector (matching with the texture of the current observed region of the image) and each class.

3.2 Supervised Clustering

To achieve a supervised clustering, it is proposed to use the Mahalanobis statistical distance measure. Contrary to the Euclidian distance, it takes into account the correlations of the data set and is scale-invariant. However, it first requires the data set to be clustered (kelp and non-kelp clusters) with this distance used to estimate the closest cluster from the data sample to be classified. Formally, the Mahalanobis distance from a certain cluster for a texture features vector X is defined as:

$$D_M(X) = \sqrt{(X-\mu)^T S^{-1} (X-\mu)} \quad (3)$$

where μ and S are respectively the mean vector and the covariance matrix of the cluster in question.

A supervised method is considered the most appropriate approach in this case as the clusters do not require a lot of samples to make the processing working well. In addition, an unsupervised technique adds errors from the clusters construction to the possible errors encountered during the latter classification phase. The main advantage of using the Mahalanobis distance is that it does not require any normalization or scaling of the data set, but this method completely relies on the strength of the Haralick texture features from sub-images to those of the training feature sets.

4 Kelp Probability Mapping

Using the clustering methods described in Section 3 it is desired to build a probability map indicating the likelihood of kelp occupying a particular region of an image. In order to build such a map, it is proposed to use a sliding window over the original image and for each sub-image classify it from the texture features to determine if this neighbourhood has been recognised as a kelp textural region.

To build the probability map, the sliding windows are overlapped and the mean value determined over each pixel of the image. As a result, a likelihood of how close each region of the image is from the kelp cluster can be determined. The probability map is the result of averaging the binary output from the texture classification for overlapping sub-images. To improve accuracy, a sliding window of the same size as the training samples was used. Also in order to ensure the same number of overlaps per pixel (especially for the edges), the original image was resized by mirroring out the edge pixels into a depth equal to the window size around the entire boarder. Finally, the number of overlaps per sub-image was chosen as a compromise between computation time and desired accuracy.

A final normalisation step can be performed allowing the use of the entire available grey-level range to build the probability map. This normalization will permit the removal of the lowest probabilities (so only the centre of low probability areas will be kept for better accuracy) while enhancing the kelp masses. However, such a normalisation should be performed after a thresholding process over the final image to reduce enhancement of low probabilities due to false positives.

The final procedure developed for image-based kelp detection and probability map construction is shown in Figure 3.

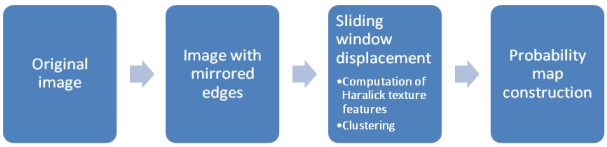


Figure 3: Final method adopted for kelp detection and probability map construction.

5 Results

5.1 Unsupervised versus supervised clustering

A comparison between the unsupervised and supervised clustering methods of Sections 3 and 4 using the training samples from Figure 2 on a random image containing kelp is shown in Figure 4 for a 100x100 pixel sliding window.

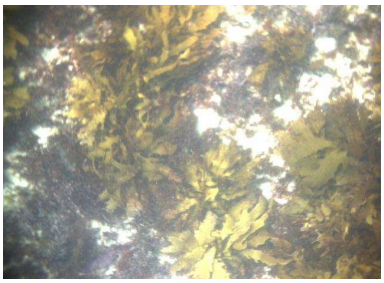
As shown in Figure 4 both clustering methods provide similar results in terms of segmentation and probability maps. Even if Euclidian distance method (unsupervised) returns very few false positives, the supervised process seems to be able to detect more kelp areas. Moreover, for more convenience (no scaling or base shift required on the raw data), it is preferred to use the Mahalanobis distance as the classification method.

Despite the observed classification performance, it was found that a considerable number of false positives can be encountered with both methods when processing images with significant colour unbalances as shown in Figure 5. Consequently, the sample images used as a training set has to be as diversified as possible in order to counter the different random lighting conditions, underwater visibility or red channel absorption in these highly dynamic environments.

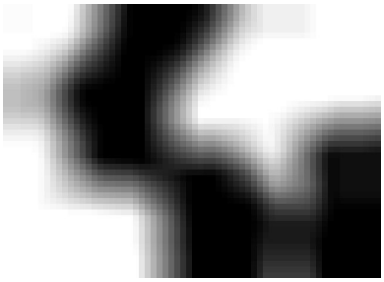
5.2 Sliding window size

The size of the sliding window impacts the computation time required to collect texture features. However, the texture "aspect" of an object is not the same when it is seen from differing view heights. Figure 6 shows an example probability map generated using the Mahalanobis distance with two different sliding window sizes.

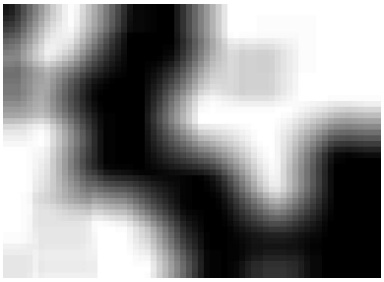
As seen in Figure 6, using a 100x100 pixels sliding window allows a quite good general definition of kelp texture but it gives less accurate spatial detection of the kelp as the probability of a window mixing kelp and non-kelp texture becomes higher with window size. However, with a small sliding window (50x50 pixels) it becomes more difficult to define a kelp texture, i.e. the kelp texture becomes non regular with important texture differences occurring between the base and the



(a) Original image.



(b) Probability map from unsupervised clustering.



(c) Probability map from supervised clustering.

Figure 4: Comparison of kelp probability maps from an image generated by the Euclidean (unsupervised) and Mahalanobis (supervised) clustering methods using a 100x100 pixel sliding window.

leaves of the kelp. A larger sliding window tended to average such differences.

Indeed, as shown in Figure 7 the higher textural region averaging resulting from larger sliding windows appears to be more robust to the detection of kelp masses in the darkest areas than from a smaller window.

On another hand, working with a smaller window allows us to increase the number of overlaps over each pixel (as well as improve the computational time) which consequently allows better spatial definition of the kelp areas. As shown in Figure 8 a smaller sliding window permits detection of very small scattered kelp areas that have not been recognised by a larger averaging window.

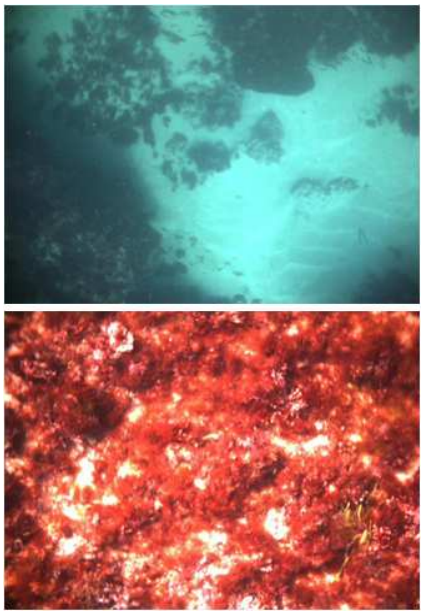


Figure 5: Examples of two significantly color unbalanced images which can result in large numbers of false positives.

5.3 Classification Results

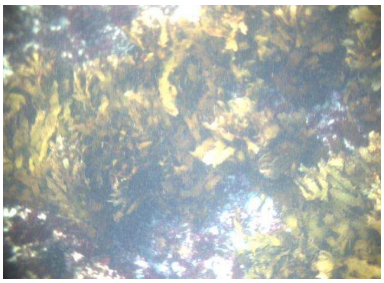
An evaluation of the classification performance was conducted using 45 images randomly selected from the entire image set containing kelp, no kelp, high altitude, over exposed images, and poor visibility. The total area of kelp in each image was segmented out by a marine scientist and this was compared against the proposed classification technique described above. Figure 9 shows two example images of manually segmented kelp along with the classified kelp probability map for the images.

Figure 10 shows the ROC curves for sliding window sizes of 50x50 and 100x100 pixels using the Mahalanobis distance. Each curve was obtained by comparing pixel by pixel the manually segmented image and calculated probability map for all 45 images and varying the probability threshold for classification. The ROC curves show that the larger sliding window has slightly better overall classification performance. Additionally, the minimum Euclidean distance to the optimum classification corresponds to a probability threshold of 0.55.

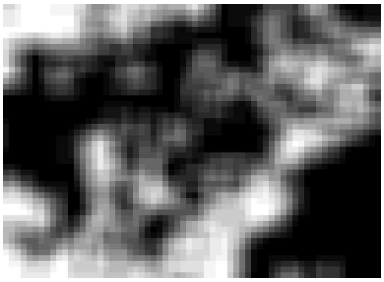
Figure 11 shows the percentage of classified kelp area (both true and false positives) using a probability threshold of 0.55 against the percentage of manually segmented kelp for the 45 images. These results show generally good agreement between the proposed technique and manual segmentation with the larger outliers due to very poor visibility and/or high altitude images.

5.4 Kelp distribution mosaicing

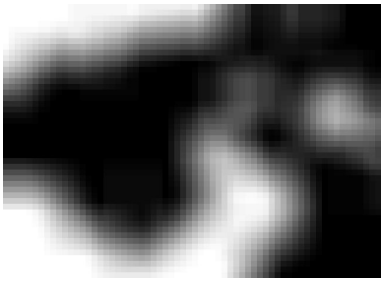
One main difficulty encountered when building a mosaic from underwater images containing large amounts of kelp is



(a) Original image.

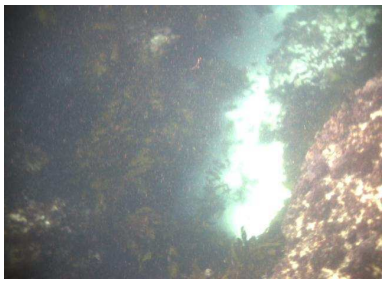


(b) Probability map from 50x50 pixel sliding window.

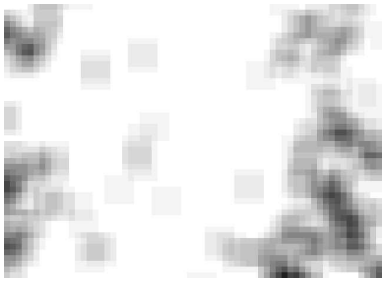


(c) Probability map from 100x100 pixel sliding window.

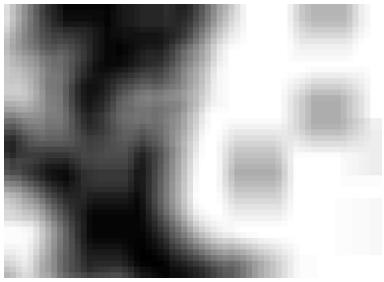
Figure 6: Comparison of probability maps generation using the Mahalanobis distance with varying sliding window size.



(a) Original image.



(b) Probability map from 50x50 pixel sliding window.



(c) Probability map from 100x100 pixel sliding window.

Figure 7: Comparison of probability maps generation using the Mahalanobis distance with varying sliding window size on darker images.

the swaying resulting from wave action. This complicates the features matching between two overlapping images. By building a probability map of the seagrass areas, an estimate of the location of the kelp mass centres can be calculated giving more matching features in case of mosaicing processing. Figure 12 shows a sequence of images from the AUV and the resulting kelp probability maps before mosaicing.

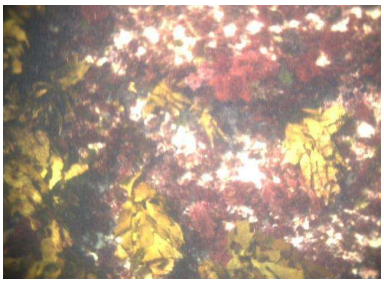
Current work is developing methods to use these probability maps from image sequences with kelp sway to reliably quantify kelp abundance.

6 Conclusions

This paper has presented a qualitative and quantitative analysis of a texture recognition based method for classifying kelp from images collected in highly dynamic shallow water environments with uncontrolled lighting and significant perspective and visibility variation. This approach uses the Mahalanobis distance as a way to classify Haralick texture features of a sub-region of the original image. These features are determined for multiple rotations and scales as well as color channels to improve classification performance.

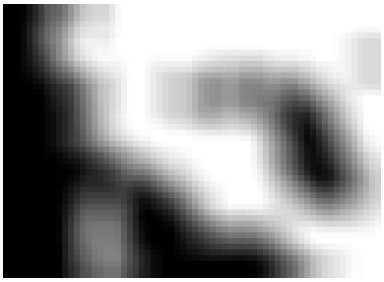
Kelp probability maps are generated by averaging the texture classifications of a sliding window sub-image across the entire image. A compromise between kelp detection performance, spatial definition and computation time can be obtained by varying the sliding window size of the sub-image used for classification.

Preliminary results from a data set collected in relatively shallow water by an AUV have shown the method to reliably segment kelp from a range of image sequences with significant variability in lighting, visibility and seafloor characteristics. Additionally, kelp probability maps have been successfully formed to allow approximation of the spatial density of swaying kelp. Future work will consider an adaptive sliding



(a) Original image.

(b) Probability map from 50x50 pixel sliding window.



(c) Probability map from 100x100 pixel sliding window.

Figure 8: Comparison of probability maps generation using the Mahalanobis distance showing kelp spatial definition with varying sliding window size.

window size during the probability map construction phase to account for significant variations in altitude across an image typical of reef like environments.

Acknowledgements

This internship was funded by in part by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) through its National Research Program in Wealth from Oceans. The authors would like to thank Dr Russell Babcock from CSIRO Marine and Atmospheric Research for his assistance in collection of data and processing of results.

References

[Arvis *et al.*, 2004] V. Arvis, C. Debain, M. Berducat, and A. Benassi. Generalisation of the cooccurrence matrix for colour images: Application to colour texture classification. *Image Analysis and Stereology*, 23(1):63–72, February 2004.

[Dekker *et al.*, 2005] A.G. Dekker, V. Ernesto Brando, and J.M.Anstee. Retrospective seagrass change detection in a shallow tidal australian lake. *Remote Sensing of Environment*, 97:415–433, 2005.

[Dunbabin and Allen, 2007] M. Dunbabin and S. Allen. Large-scale habitat mapping using vision-based AUVs: Experiences, challenges & vehicle design. In *Proc. OCEANS 2007 Europe*, pages 1–6, Aberdeen, UK, 2007.

[Haralick *et al.*, 1973] R.M. Haralick, K. Shanmugam, and I. Dinstein. Generalisation of the cooccurrence matrix for colour images: Application to colour texture classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6):610–621, November 1973.

[Ierodiaconou *et al.*, 2010] D. Ierodiaconou, J. Monk, A. Rattray, L. Laurenson, and V.L. Versace. Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. *Continental Shelf Research*, 2010.

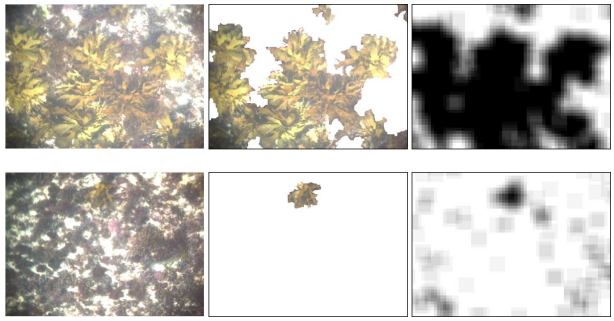


Figure 9: Example segmentation results for classification performance assessment. (Left) original image, (middle) manually segmented image, and (right) probability map using 100x100 pixel sliding window with Mahalanobis distance.

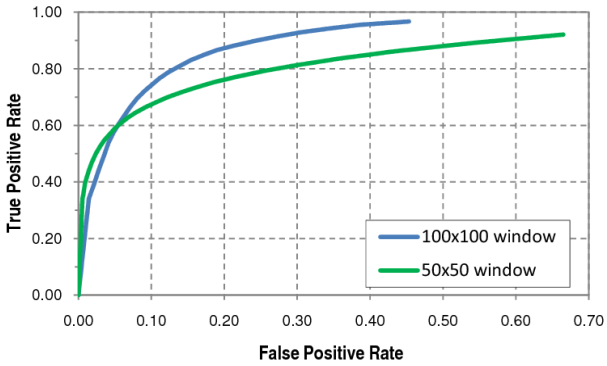


Figure 10: Receiver Operating Characteristic (ROC) curves for classification results with varying sliding window size using the Mahalanobis distance.

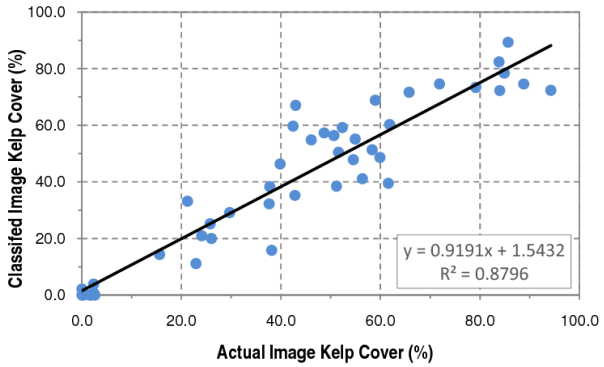


Figure 11: Percentage of total area of kelp classified using 100x100 pixel windows and Mahalanobis distance against percentage of manually segmented kelp area for 45 images.

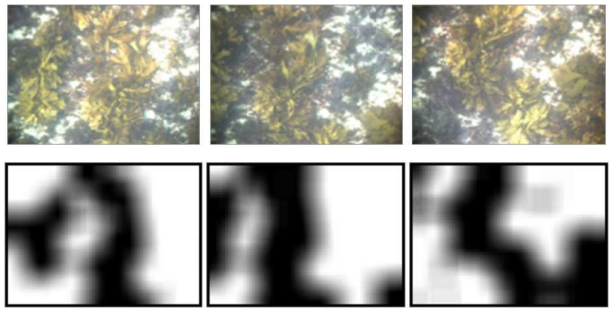


Figure 12: Sequence of overlapping images and the corresponding probability maps.

[Lathrop *et al.*, 2006] R.G. Lathrop, P. Montesano, and S. Haag. A multi-scale segmentation approach to mapping seagrass habitats using airbourne digital camera imagery. *Photogrammetric Engineering & Remote Sensing*, 72:665–675, June 2006.

[Miyamoto and Merryman, 2005] Eizan Miyamoto and Thomas Merryman. Fast calculation of haralick texture features, 2005.

[Moline *et al.*, 2007] M. Moline, D.L. Woodruff, and N.R. Evans. Optical delineation of benthic habitat using an autonomous underwater vehicle. *Journal of Field Robotics*, 24:461–471, 2007.

[Roelfsema *et al.*, 2009] C.M. Roelfsema, S.R. Phinn, N. Udy, and P. Maxwell. An integrated field and remote sensing approach for mapping seagrass cover, Moreton Bay Australia. *Spatial Science*, 54:45–62, June 2009.