

Towards autonomous characterisation of side scan sonar imagery for seabed type by unmanned underwater vehicles

L.J. Hamilton

Defence Science & Technology Group, Data 61 Building, 13 Garden St, Eveleigh NSW 2015

ABSTRACT

Surveys by autonomous underwater vehicles fitted with side scan sonar could potentially be made more efficient if they could recognize particular seabed types and modify their search strategy or software use accordingly. A limited scheme to autonomously infer actual seabed type is described. Categories are flat (even textured), sand, rippled, rough, rough periodic (seabeds of unknown type with low wavenumbers and pronounced directionality), unknown, and a class for a collection of types which presently are difficult to distinguish (including seagrass, and gravelly sands with high reflectivity speckle). The scheme has three stages: (1) even texture/rough texture is decided using the GLCM (Gray Level Co-occurrence Matrix) image processing technique, (2) periodicity and directionality are examined with 2-D Fourier and Radon Transforms, and (3) an inference of seabed type for non-rippled seabeds is made using the Gray Level Size Zone Matrix (GLSZM) scheme. The final inference is a consensus of the three stages. Some seabed types as seen in side scan sonar imagery defy categorization by any means, and problems arise from motion artefacts, saturation at higher backscatter levels, uncorrected angular effects, sea surface acoustic returns, and acoustic ambiguity.

1 INTRODUCTION

Sidescan sonars mounted on surface vessels, tow bodies, or Autonomous Underwater Vehicles (AUVs) are used to obtain acoustically derived grayscale imagery of the seafloor. For a description of sidescan sonars see Blondel (2009). Some seabeds have characteristic textures or patterns in sidescan imagery (for example, sand ripples) which allow the identification and mapping of seabed types. The imagery may also be scanned for manmade objects proud of the seafloor, particularly minelike objects (MLOs).

Use of AUVs fitted with sidescan sonar for the above purposes could be made more efficient if they were able to recognize which areas were likely suitable for survey purposes and which were not while on survey. Achieving this recognition autonomously is the topic of the present paper. The key to autonomous classification is obviously attaining an acceptable degree of error. However, machine emulation of human visual acuity and knowledge is not a simple task.

2 DATA

Side scan sonar imagery gathered by AUVs from locations around Australia, the western Pacific, and elsewhere was supplied as 256 grayscale level bitmaps of 1024x1000 pixels, with nominal swath width of 30 m a side (6 cm pixels). Sonar frequencies were 600 and 900 kHz. The AUVs nominally travelled 3 m above the seabed at 4 knots speed. Eight subimages of 256x256 pixels were extracted from each image (four from port and four from starboard) to provide 77,351 grayscale bitmaps about 15 m square. This size provided sufficient resolution and repetition of features such as ripples, stones, and smaller coral outcrops to enable reliable visual classifications of seabed characteristics. Images are not georeferenced, and are not corrected for sonar beam pattern, slant range effects, propagation losses, or variations in vessel speed, heading, attitude, or height above seabed. Nor do they contain depth information. Some images have artefacts from AUV ascents, descents, turns, adjustments to height over steeper slopes, motions in swell, interference from communications equipment, and a surface echo which appears as a higher energy line quasi-parallel to vessel motion. Vessel motions can be monitored for some of these effects when on survey, but the images are analysed without any corrections.

3 METHODS

3.1 Method 1 - Gray Level Co-occurrence Matrix (GLCM) method

The GLCM method was used to segment the sonar images into regions of different texture. This technique has been used for post-processing of sidescan sonar imagery for about 40 years (for example Pace and Dyer (1979), Reed and Hussong (1989), Keeton (1994), Blondel et al (1998)). GLCM is an image processing tech-

nique developed by Haralick et al (1973) which analyses texture and tone. *Tone* refers to the backscatter amplitude (the gray scale) of image pixels, with dark tone indicating lesser backscatter return than light tone. *Texture* refers to a repeating pattern, such as sand ripples. Each pixel is classified with up to 32 indices. “GLCMs address the average spatial relationships between pixels of a small region, by quantifying the relative frequency of occurrence of two gray levels at a specified distance and angle from each other. ... Because they are difficult to manipulate and interpret, GLCMs are described by statistical measures, called indices.” (Blondel et al 1998). Possible look angles are 0, 45, 90, 135°. “Effectively, a two-dimensional histogram is created (dimensioned to $G \times G$ [where G is the number of gray scale levels in the image]) that represents the total number of occurrences for each (i, j) gray level pair [i and j are integer gray scale levels] that occurs in a fixed sized window given a certain spatial (and angular) offset. ... The co-occurring probabilities are determined by dividing $C(i,j)$ [the number of occurrences for a particular (i,j) pairing] by the total number of counts across all i and j” (Clausi & Zhao, 2003).

3.1.1 Choice of GLCM parameters

The simplest GLCM usage occurs if a single parameter can be used to characterize imagery. Since both texture and tone are important for the present work, this is unlikely to be the case. According to Blondel et al (1998) only the two GLCM parameters Homogeneity and Entropy are needed to segment sidescan imagery. However, Baraldi and Parmiggiani (1995) consider Homogeneity to be of little use in texture analysis. Other papers have similar contradictions. Clausi and Jernigan (1985) place GLCM indices into three groups (smoothness, uniformity, correlation), stating that one index from each group should be used. Clausi (2002) recommends Contrast, Entropy, and Correlation. Contrast is not used for the present work because many of the sonar images have saturated grayscale values of 255, which would likely produce unreliable statistics.

The eleven GLCM parameters provided by modifications of the efficient linked list C++ coding of Clausi & Zhao (2003) were evaluated. Pairwise comparisons showed Correlation to be independent of the other ten, all of which were strongly related to each other, and only Entropy and Correlation were selected. The Clausi and Zhao (2003) Entropy and Correlation parameters were replaced by normalized versions from Bianconi & Fernandez (2014) and rescaled to range 0-255. A window size of 19x19 pixels was selected as visually optimal for the data set with respect to likelihood of finding MLOs. A smaller window produces too much clutter. A larger window obscures features of interest.

The classification space formed by the two GLCM parameters was manually segmented to match visual estimates of seabed smoothness/roughness and clutter at the scale of MLOs. Four roughness classes were used (green, yellow, maroon, red), with a fifth for shadowed / unenisonified / saturated pixels (gray). Each pixel is given a classification, except for strips 9 pixels wide (half the window size) along image borders.

Classified image pixels were defined by polygonal regions using crack code contour methods (Kovalevsky, 2006). The crack code also calculates horizontal and vertical extents and areas (numbers of pixels) of the polygonal objects.

3.1.2 GLCM examples

Examples of sonar images and their GLCM classifications are shown in Figures 1 to 4.

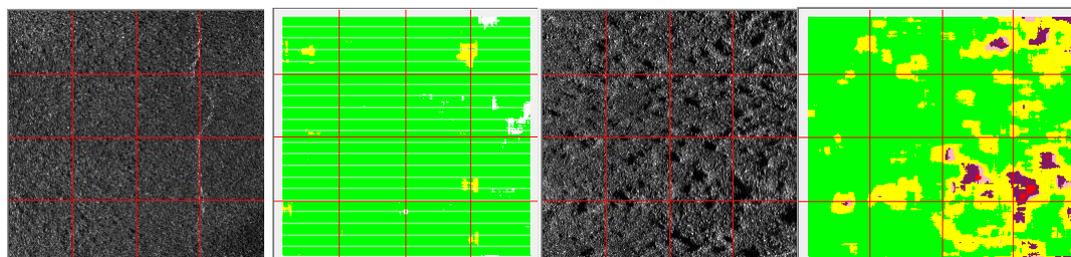


Figure 1. Image / GLCM pairs. Left. Relatively featureless seabed with a weak surface return. Right. Rougher seabed. The GLCM classes (1) to (5) are color coded as green, yellow, maroon, red, gray/black respectively.

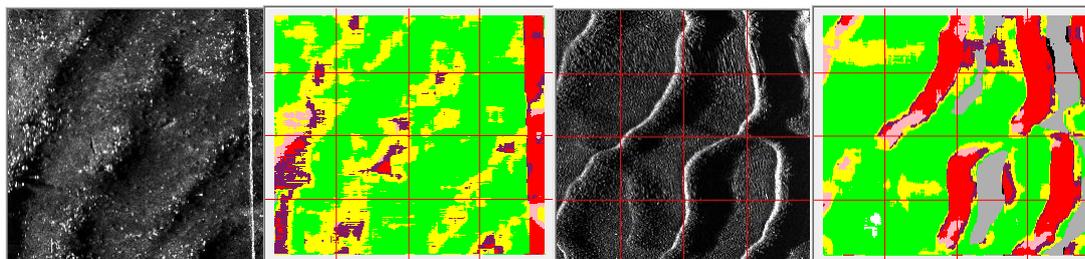


Figure 2. Image / GLCM pairs. Left. Parallel linear trending undulations with reflective crests, and a surface return. Right. Sharply defined sand waves with shadow regions.

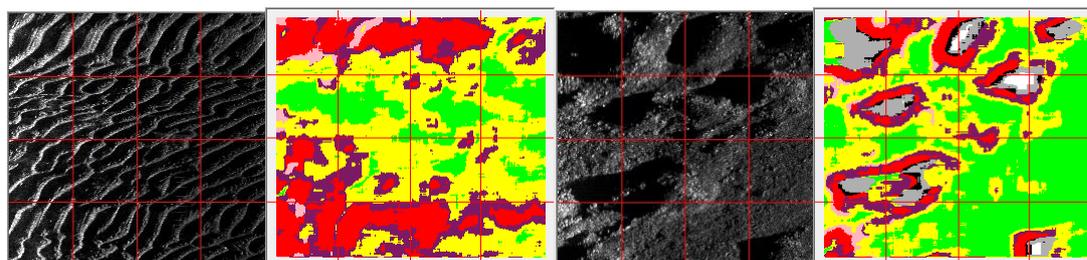


Figure 3. Image / GLCM pairs. Left. Larger ripples. Note the range of classes depending on acoustic contrast, ripple wavelength, and ripple orientation with respect to the sonar. Right. Rough rock seabed with depressions and unsonified shadow regions.

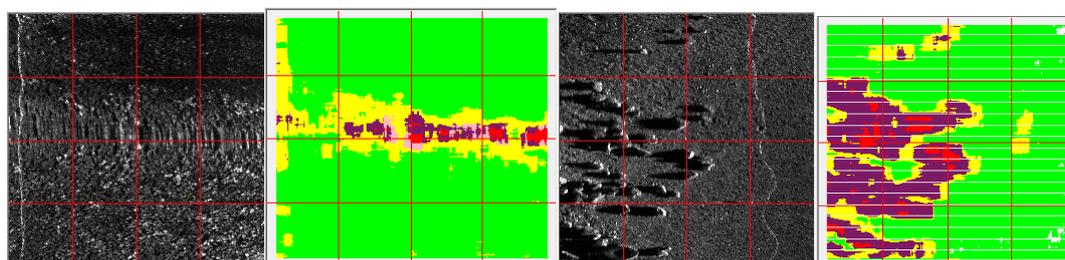


Figure 4. Image / GLCM pairs. Left. A horizontal distortion band with parallel vertical structures due to AUV motion. Right. A boulder field.

3.1.3 Whole of image indicator for feasibility of hunting

A simple quantitative indicator of hunting difficulty for each 256x256 sonar bitmap image was obtained using the percentage of pixels in the roughest class (4) and shadow class (5). This simple index exhibited consistency of geospatial classification and usefulness for whole surveys (Figures 5, 6). Indices with index less than 3% corresponded to evenly textured seabeds. Indices above 3% identified images having rougher areas, longer wavelength sandwaves, and/or higher contrast, and artefacts produced by vessel turns or departures from the horizontal.

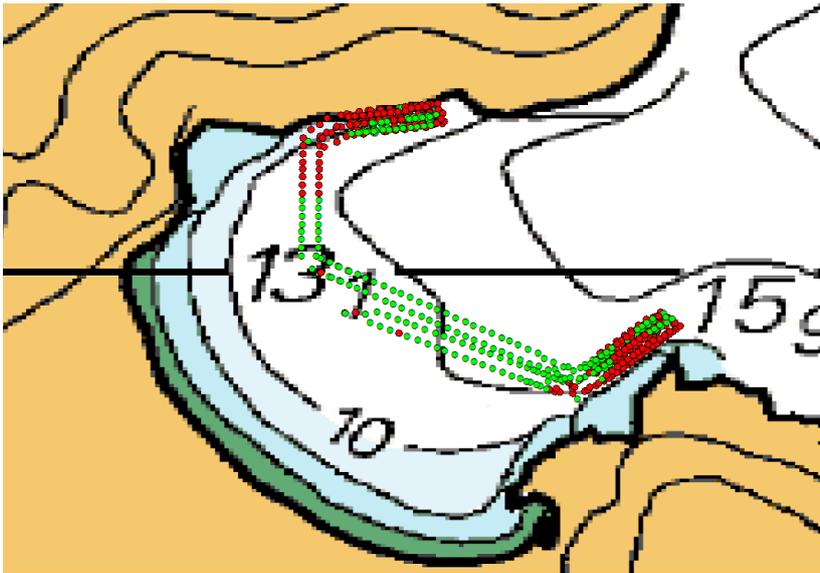


Figure 5. Classification of seabed as even textured or as rough (including artefacts). Each colored symbol represents a 256x256 pixel image. Red symbols show images with more than 3% rough pixels (corresponding to rocky seabeds in this case), green symbols for images with less than 3% rough pixels.



Figure 6. Classification of seabed as even textured or as rough (including artefacts). Red areas have more than 3% rough pixels, green areas have less than 3%. Red images at the ends of north-south sections correspond to image distortions caused by turns. In the west the quasi-zonal band along the centre of the surveyed area is caused by image distortions from terrain following height adjustments over a steep slope. The red eastern area corresponds to dark imagery with little information.

3.1.4 Selection of Template Images

Characterization of images for actual seabed type was sought as a whole using template matching methods, rather than individually labelling portions of imagery. Templates were identified using statistical clustering methods described in Hamilton and Cleary (2015). Three clusterings were made using the fast CLARA statistical clustering software of Kaufman and Rousseeuw (1990). First, the GLCM 2-parameter classification space (Entropy vs Correlation) for each image was segmented into 50x50 squares, the number of image pixels falling in each of the 2500 squares was counted, and the 77,351 vectors of 2500 counts were clustered. Secondly, probability contours were calculated for the GLCM 2-parameter classification space using Gaussian Mixture Model (GMM) techniques, and particular contours were clustered (Figure 7) using the techniques of Hamilton (2013). Thirdly, the 77,351 gray scale brightness histograms (values 0-255) were clustered. Results were mixed but templates were selected from clusters with more consistent seabed types.

Thirteen groups of 10 images each were compiled (bad data, drag marks, holes and pock marks, even texture, lightly rippled, rippled, ripple bands or sand ribbons, mega-ripples, fields of stones or objects on an otherwise uniform background, coral outcrops on a uniform background, rock, very rough (usually rocky very uneven surfaces), and seagrass. Ten images per group were used in an attempt to account for the natural variation in seabed type. Seabeds will not all have the well defined property ranges of the template sets, but identification of more uniform seabeds is likely to be more tractable than identifying mixed types.

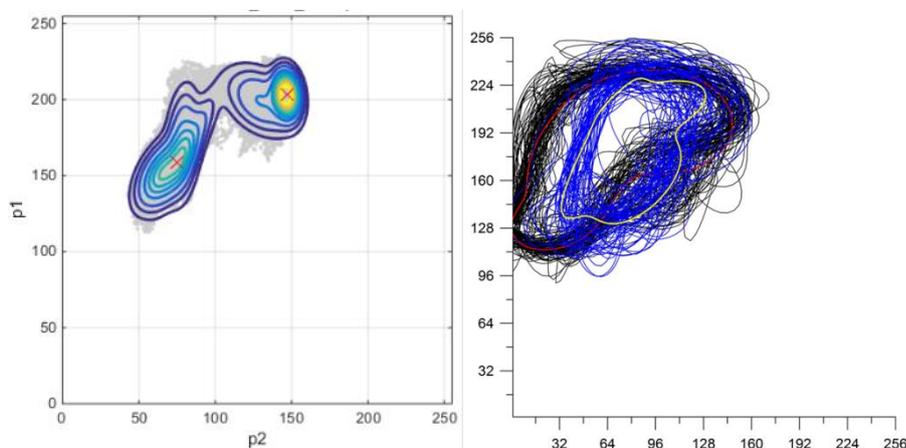


Figure 7. Gaussian Mixture Model (GMM) probability contours of GLCM data. (a) GMM contours for an image with two different textures. (b) Examples of clustering of GMM contours (black and blue curves), and their central tendencies (the red and yellow curves).

3.1.5 Template Matching

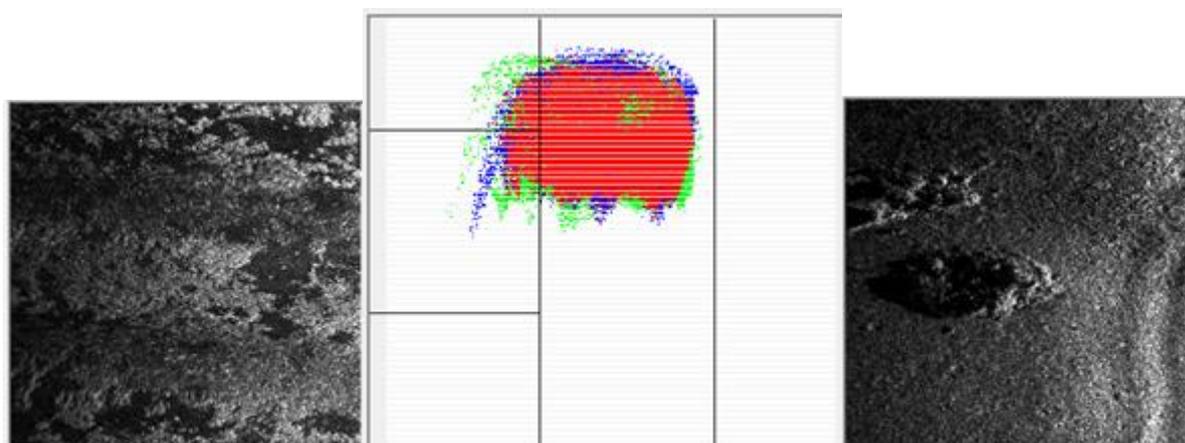


Figure 8. Ambiguity in a 2-parameter GLCM signature. Seagrass (left) and outcrops on sand (right) have near the same GLCM template. Red pixels show matches, blue and green show differences.

None of the GLCM, histogram, or contour methods sufficed as a standalone method for template matching. Some very different seabeds had near the same GLCM signatures (Figure). The GLCM method was more useful in indicating templates that did *not* match. The GLCM and histogram methods were then combined. The best five matches of an incoming image with template GLCM signatures and the best five matches with histograms were found. If the same template appeared in both sets of five then it was automatically the pick. If a matching pair did not exist then the GLCM pick with histogram closest to that of the incoming image was used. This simple scheme had more success, and outperformed several Context Based Image Retrieval (CBIR) packages, but did not provide sufficient overall reliability for classification. More directed methods were then sought.

3.2 Method 2 - Directionality And Periodicity

Sams et al (2004) used Scharr convolution matrices to estimate directionality in sidescan sonar images, because they found that 2-D Fourier Transforms did not give reliable results. However, for the present work Scharr counts were not useful. They were a function of the wavelength of sandwaves, and secondary directionalities were often swamped by the dominant direction of features. A coupled 2D-FFT/Radon transform/auto-correlation function method was developed instead.

3.2.1 2-D Fourier Transforms and Radon Transforms

A geometrically perfect 2-D sine wave of higher frequency appears as two points equidistant from the origin of the optical transform expression of a 2-D FFT, corresponding to positive and negative frequency. Real world ripples appear as two point spread functions rather than two points (Figure 9). The two points (or point spreads) and the origin lie on a straight line (Figure 9). The average distance of the point spread to the transform centre is proportional to spatial frequency. The orientation of the line joining the transform centre to the point spread is the direction of the sine wave. Short wavelength ripples and longer wavelength sandwaves often coexist, and may have the same or different average orientation.

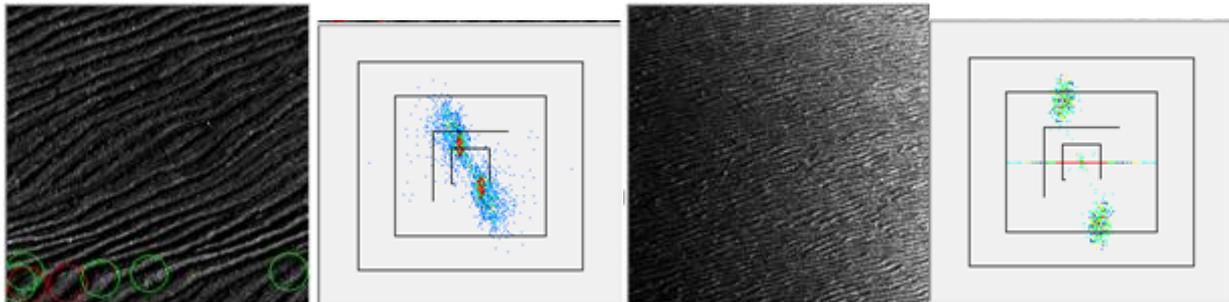


Figure 9. Ripples and sandwaves have point spreads in a 2D-FFT.

Amplitudes for 2-D FFTs for each 256x256 image were squared and summed to form energy estimates for 64 annuli of equal width (2 FFT pixels) centred on the optical transform view. The 64 energy values were used to estimate periodicity through the presence of spectral peaks. The major problem was inability to separate longer wavelength features from near flat areas. Some undulations only have two or three cycles in a 256x256 image, and they are not usually regular enough in wavelength or direction to be separated from DC. They can appear as a broadly linear distribution of points passing through the origin, indicating a dominant direction, a spread in wavelengths, and the possible presence of larger unresolved wavelengths (Figure 10).

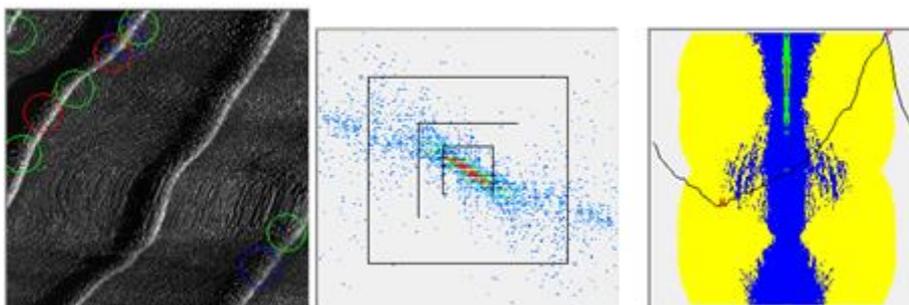


Figure 10. 2D-Fourier and Radon transforms for long wavelength features.

Useful directional information was extracted by applying the Radon Transform code of Hoiland (2007) to the 2D-FFT. Lines in images transform to points in the Radon Transform according to their distance and direction from the centre of the images (Hoiland (2007) shows examples). The spread of real world data in the 2D-FFT will cause points that fall on lines passing through the origin of the 2D-FFT to transform to a spread of points appearing on and about the zero displacement line of the Radon Transform. Their position along the zero displacement line indicates their orientation in the sonar image. The two point clouds for higher frequencies will also transform to one point spread in the Radon transform. Radon Transform energies (amplitudes squared) for displacements close to zero were summed for 1° direction increments and peaks in the energies were used to infer directionalities in the sonar image.

Cross-sections of backscatter (and of 2D-FFT) values were made along the Radon inferred directions for lines passing through the centre points of the original image (and the 2D transform). 1-D Fourier Transforms were made for the extracted backscatter cross-section, and autocorrelation functions formed the backscatter section. The 2D-FFT section and the 1-D FFT did not provide reliable indications of low spatial frequencies, but the autocorrelation function was more successful in this aim.

Energy peaks above a threshold for annuli 5 to 32 were assumed to indicate rippled sandy seabeds. Seabeds with spectral peaks in the first four 2-D FFT annuli indicated by the Radon methods to have directionality were labelled RoughPeriodic, otherwise they were labelled Rough. This is suitable to assist the mine hunting aim, but not for inference of actual seabed type. Some rock seabeds in the RoughPeriodic class were differentiated from undulations and sand waves by comparing curves of counts of pixel energies calculated for five broader areas about the centre of the optical transform. The first two line segments of rock energy curves are typically concave, whereas curves for other seabeds are convex. Rock typically has higher energy at any frequency than other seabeds, and the initial dropoff of energy is not as rapid. These criteria do not separate all cases, but are effective for brighter or more well ensounded surfaces.

3.3 Method 3 - Gray Level Size Zone Matrix (GLSZM)

The GLCM / histogram / FFT and Radon Transforms methods largely use indirect image characteristics, not actual properties of numbers and sizes of seabed features such as stones or outcrops. The Gray Level Size Zone Matrix (GLSZM) scheme (Thibault et al 2009) employs this direct information. For each grayscale bitmap the GLSZM technique processes the number of zones of size S (as number of pixels) and gray level G (G is quantised to 16 levels or some suitable number). The size zones were obtained using pixel crack code polygons (Kovalevsky 2006). As many as 16 GLSZM indices may be calculated to describe image texture (<http://www.thibault.biz/Research/ThibaultMatrices/MGLSZM/MGLSZM.html>). Unlike the GLCM method, calculations are not directional. Pairwise plots showed that only two indices (Small Zone Emphasis and Gray Level Non-uniformity) were independent (with a third partially independent). The GLSZM technique was evaluated by application of these two indices to the 13 template groups.

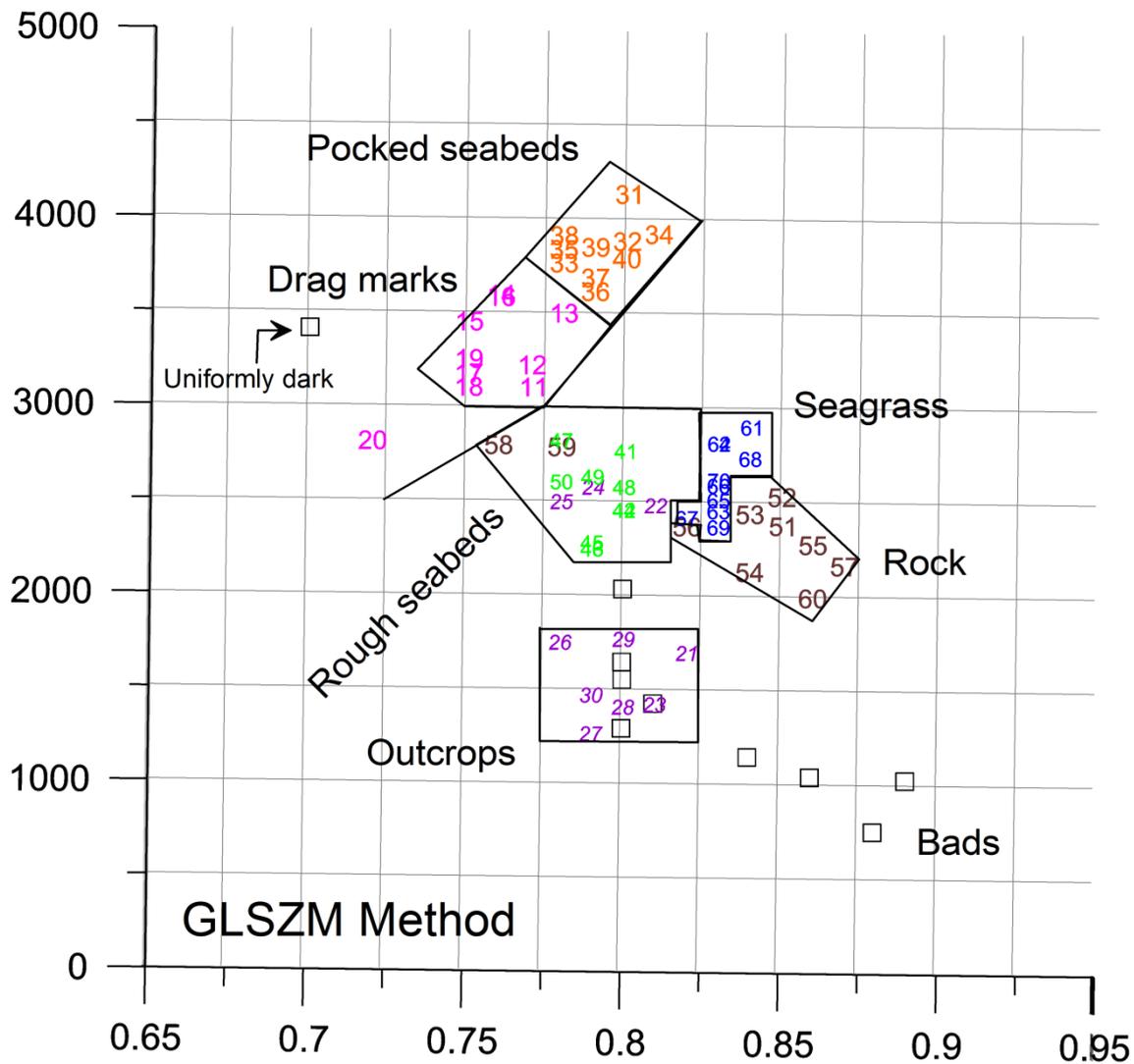


Figure 11. A 2-parameter GLSZM classification of selected template sets. X-axis is Small Zone Emphasis, Y-axis is Gray Level Nonuniformity.

Several template sets occupied largely separate areas of the GLSZM 2-parameter space (flatter areas, outcrops, rough ground, rock, seagrass) (Figure 11). *GLSZM was the only method to achieve separation of seagrass and rock templates.* The GLSZM method was tuned to the templates. *A problem with using templates for evaluation is that they may not be classed by the properties the user requires.* GLSZM did not see the drag lines or pocks templates as different from templates for flatter areas, as these features usually had gray level differing little from their surrounds.

4 Final method of seabed classification

Six broad classes were chosen as being achievable. These were flat (evenly textured), ripples, sand, rough (outcrops, rock, rough rock, stone fields), RoughPeriodic (rough seabeds exhibiting directionality, or undulations and sandwaves with unresolved larger wavelengths), and dark (near black, unenisonified, or shadowed images). Two unknown classes were also formed. The first unknown class is for irreconcilable differences in classification for different methods, for example, even texture from one method and rough from another. The second unknown class is for images with multiple bright pixels or highlights (speckle) which can range from some types of highlighted sands and gravels to seagrasses. No methods were able to consistently separate images in this class, although visibly they are very different. This remains to be resolved.

Images are first classified by the GLCM method as being flat (actually evenly textured), or non-flat seabeds. Non-flat seabeds are classed as rippled (using the 2D-FFT and Radon Transforms), or rough. Counts of stone

or boulder size objects are made for non-rippled surfaces using crack code results. Flat and rough seabeds are then further classified with the GLSZM method. If the GLCM and GLSZM classifications both agree that a seabed is flat or is rough the classification is accepted, and the GLSZM class is used as a secondary label. If not, then the most likely is chosen where possible, else the seabed is classed unknown. A surety index is provided by the number of methods in agreement.

4.1.1 Statistics And Reliability

A survey with 1456 images returned a classification as follows: 5 dark class, 14 unknown with highlights, 92 flat, 104 rippled, 264 sand, and 977 rough (these were rough rock). Excluding images with motion artefacts and outright bad data (caused by instrument jitter or interference by other instruments), which are not a fault of the methodology, then a classification success of 100% could be claimed. This is not the case for other surveys, but it does show what can be achieved for more well ordered seabeds.

A survey with 1568 images returned classes as 442 dark, 42 unknown with highlights, 62 flat, 16 rippled, 7 sand, and 999 rough. Many roughs were caused by a bright surface return line on dark images. They are 'rough' with respect to creating difficulties for minehunting, but not physically rough. The presence of surface returns is to be handled by a preprocessing section prior to the seabed classification software being run.

DISCUSSION

Autonomous means of characterizing actual seabed type from side scan sonar data has only been achieved in basic terms. Human observers can assimilate a wide range of textural, shape, and spatial relationships exhibited by seabeds in side scan sonar imagery which automatic and autonomous methods may not be able to, including recognition of bad data and of portions of good data in an otherwise bad image. The biggest problems in analyzing the present data set were data quality, which varied greatly over the different surveys, the great variation in seabed types, and acoustic ambiguity for the chosen methods of classification. Autonomous surveys of particular areas would likely yield better results, especially if only particular seabed types were of interest, or if some knowledge of seabed types were available. One way to obtain this knowledge is to use the first autonomous mission as a reconnaissance survey.

Other techniques might provide a more refined inference of actual seabed types than achieved, and there are many to choose from, including Convolutional Neural Networks / Deep Learning (Chapple et al 2017). However, this could be some way off, and presently these newer techniques are informed by methodology such as that presented in the present paper.

REFERENCES

- Baraldi, A. & Parmiggiani, F. 1995. An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33, No 2, 293-304.
- Bianconi, F. & Fernández, A. 2014. Rotation invariant co-occurrence features based on digital circles and discrete Fourier transform. *Pattern Recognition Letters* 48: 34-41.
- Blondel, P. 2009. *The Handbook of Sidescan Sonar*. Springer-Verlag. 316pp.
- Blondel, P., Parson, L.M. & Robigou, V. 1998. TexAn: Textural Analysis of Sidescan Sonar Imagery and Generic Seafloor Characterisation. Proceedings 1, 419-423. OCEANS'98, IEEE-OES.
- Chapple, P., Dell, T., Bongiorno, D. 2017. Enhanced detection and classification of mine-like objects using situational awareness and deep learning.
- Clausi, D.A. 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can J Remote Sensing* 28, No. 1, 45-62.
- Clausi, D.A. & Jernigan, M.E. 1985. [A fast method to determine co-occurrence texture features](#). *IEEE Transactions on Geoscience and Remote Sensing* 36(1), 298-300.
- Clausi, D.A. & Zhao, Y. 2003. Grey level co-occurrence integrated algorithm (GLCIA): a superior computational method to rapidly determine co-occurrence probability texture features. *Computers & Geosciences* 29, 837-850.
- Goff, J.A., Olson, H.C., Duncan, C.S. 2000. Correlation of sidescan backscatter intensity with grain-size distribution of shelf sediments, New Jersey margin. *Geo-Marine Letters* 20, 43-49.
- Hamilton, L.J. 2013. Statistical clustering of drifting buoy trajectories around Japan and off Fukushima to identify Lagrangian circulation features. *Methods In Oceanography* 6, 16-32.
- Hamilton, L.J. 2011. Acoustic Seabed Classification For Echosounders Through Direct Statistical Clustering Of Seabed Echoes. *Continental Shelf Research* 31, 2000-2011.
- Hamilton, L.J. 2010. Characterising spectral sea wave conditions with statistical clustering of actual spectra. *Applied Ocean Research* 32(3), 332-342.

- Hamilton, L.J. 2007. Clustering Of Cumulative Grain Size Distribution Curves For Shallow-Marine Samples With Software Program CLARA. *Australian Journal of Earth Sciences* 54, 503-519.
- Hamilton, L.J. and Cleary, J. 2015. Autonomous processing of sidescan sonar imagery for unmanned underwater vehicles. Acoustics Australia Society conference. Hunter Valley, New South Wales.
- Hamilton, L.J. and I. Parnum. 2011. Seabed Segmentation From Unsupervised Statistical Clustering Of Entire Multibeam Sonar Backscatter Curves. *Continental Shelf Research* 31(2), 138-148.
- Haralick, R.M., Shanmugam, K. and Dinstein, R. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3(6), 610-621.
- Kaufman, L. and Rousseeuw, P.J. 1990. *Finding Groups In Data: An Introduction To Cluster Analysis*. John Wiley, New York, 1990.
- Keeton, J.A. 1994. The use of image analysis techniques to characterise mid-ocean ridges from multibeam and sidescan sonar data. PhD Thesis, University of Durham. <http://etheses.dur.ac.uk/1620/>
- Kovalevsky, V. 2006. Code Samples: Tracing and encoding boundaries of subsets in 2D images. [http://www.miszalok.de/Samples/CV/ChainCode/chaincode_kovalev_e.htm]
- Pace, N.G. and C. M. Dyer, 1979. Machine classification of sedimentary sea bottoms, *IEEE Trans. Geosci. Remote Sensing* GE-17, 52–56.
- Reed, T.B. & D. Hussong. 1989. Digital image processing techniques for enhancement and classification of SeaMARC II sidescan sonar imagery. *Journal of Geophysical Research* 94, 7469-7490.
- Sams, T., Hansen, J.L. Thisen, E. Stage, B. 2004. Segmentation of sidescan sonar images. Danish Defence Research Establishment Report DDRE M-21/2004. 30pp.
- Thibault, G., Fertil, B., Navarro, C., Pereira, S., Cau, P., Levy, N., Sequeira, J., Mari, J-L. 2009. Texture Indexes and Gray Level Size Zone Matrix. Application to Cell Nuclei Classification. *Pattern Recognition and Information Processing (PRIP)*: 140–145.