A TEXTURE ANALYSIS APPROACH TO IDENTIFYING SABELLARIA SPINULOSA COLONIES IN SIDESCAN SONAR IMAGERY

Richard Harrison 1,2,* Francesco Bianconi 3,† Richard Harvey 2, Wenjia Wang 2

1Gardline Geosurvey Ltd., Gt. Yarmouth, UK
2School of Computing Sciences, University of East Anglia, Norwich, UK
3Dipartimento Ingegneria Industriale, Università degli Studi di Perugia, Italy

ABSTRACT

Offshore wind farms are undergoing unprecedented development as EU member states focus on complying with 2020 renewable energy mandates. However, wind farm site placement requires great care, to avoid compromising protected habitats, such as Sabellaria spinulosa reefs. This paper presents an investigation into the potential of different feature generation methods for identifying sidescan sonar image textures characteristic of Sabellaria spinulosa colonies. We propose an extensible test methodology and carry out a detailed comparison of several textural features. Our results show that Gabor filter bank features yield good (up to 89.4% overall) classification accuracies and often outperform other methods in identifying the Sabellaria spinulosa textural class. A Dual-Tree Complex Wavelet Transform, Ring filters and some statistical methods also produce encouraging results.

Index Terms— Texture analysis, Sonar, Sabellaria spinulosa

1. INTRODUCTION

This paper summarises an experimental investigation into the performance of some Statistical, Signal processing-based and Morphology-based texture features for classifying seabed classes in qualitative sidescan sonar imagery. We are primarily concerned with the detection of the tube-building worm Sabellaria spinulosa (Sabellaria). Sabellaria inhabits sub-tidal and intertidal zones in UK and European waters [1]. Dense accumulations of such worms can form biogenic reefs which are protected under the EU Habitats Directive. Naturally, this has planning and commercial ramifications for the offshore renewable energy industry as well as other construction activities which cause a seabed disturbance.

Sidescan sonar is often the instrument of choice for high-resolution acoustic imaging of the seabed (see for instance [2, 3]). The imagery contains useful textural information which can indicate the presence and extent of Sabellaria colonies [4, 1, 5]. The sonar is usually towed on a cable behind a survey ship. It generates pulsed, fan-shaped lobes of acoustic radiation from a transducer array which doubles as a receiver for returning signals. The acoustic energy undergoes a complex environmental interaction as it propagates through the water column, into the seabed and back to the instrument (figure 1).

Fig. 1. Illustration showing the basic operating principles of a towedsidescan sonar.

Fig. 2. A waterfall image segment is built up progressively, from a sequence of transverse scan lines as the instrument is towed above the seabed.
Returning signals are captured as a time series of amplitudes, comprising the sum of spatiotemporally varying specular components and boundary and volumetric backscatter from the seabed and water body. As the instrument advances, a waterfall image of the seabed (figure 2) is built up from consecutive scan lines. The imagery contains noise and various radiometric and geometric distortions from multiple sources [6, 7]. Image degradation can be exacerbated in the shallow water regimes where Sabellaria are found, due to wave motion and surface noise, such as from rainfall.

Optical imaging of the Sabellaria offers much higher resolution and is sometimes used for localised, non-intrusive ground truthing over small areas. Optical and sonar imaging are both stochastic processes but there are fundamental differences in the imaging physics and the transducers. Underwater optical cameras (still and video) use charged coupled device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) arrays of pixel elements for visible-spectrum imaging [8]. Absorption of electromagnetic radiation by seawater, backscattering due to turbidity and the limited ground coverage make this a less practical option for surveying large areas though. Recent advances in underwater imaging technology are described in [9].

The heterogeneous structure and spatial distribution of the Sabellaria colonies produces diverse textural proxies in sonar imagery. Identification, segmentation and mapping of these textures is predominantly a manual task performed by a human expert. However, this procedure is tedious, subjective and therefore expensive. An automated, objective approach to reliably identifying textural regions characteristic of Sabellaria would assist in the process of mapping the colonies. Further, automation would expedite processing of the large volumes of data currently being acquired in site surveys for offshore wind farm developments.

Different methods for the generic textural analysis of sonar images have been studied widely and include: Gabor filters [10]; Wavelets [11, 12] and Co-occurrence matrices [13, 14, 15]. Yet there is no published research on features which could be used for our specific task - the automated identification of Sabellaria in sidescan imagery. Due to data dependent performance, we need to evaluate a set of methods on this novel texture classification problem, as a preliminary stage in the design of an expert system for automatic Sabellaria detection.

The remainder of the paper is organised as follows. In Section 2 we briefly review the methods included in the comparison. Experimental set-up including image acquisition, preparation and accuracy estimation is outlined in section 3. Finally, we present and discuss our results in Section 4.

2. METHODS

During the last 20 years many approaches to texture analysis have been described in the literature. A comprehensive and up-to-date review can be found in [16]. Here, we consider methods from three groups: Statistical, Signal processing-based and Morphology-based. All the approaches included in the comparison are rotation invariant as we do not assume any preferred orientations of natural textures in the imagery. Given the large palette of textural descriptors available, we restrict our subset to include some well-known techniques in the sonar imaging domain and others whose performance on this type of imagery is untested. A salient of all these features in relation to the specific task is unknown though. Thus, the aim of our investigation is to find out which of these feature groups is the most promising for the task. Parameter tuning of the individual methods is a matter for future work.

2.1. Statistical methods

Statistical methods are based on the statistical distribution of greyscale values at predefined relative positions. The methods considered here are: Local Binary Patterns (LBP), Improved Local Binary Patterns (ILBP), Co-occurrence matrices (COOC) and the recently introduced Binary Gradient Contours (BGC). LBP, ILBP and CCR [17] have received significant attention in recent years due to their ease of implementation, low computational cost and high discrimination accuracy in many applications. Co-occurrence matrices are included as a useful benchmark in many comparisons and are appropriate here due to their widespread use in the sonar domain.

The basic version of the LBP (LBP) uses the 256 possible binary patterns that can be defined in a $3 \times 3$ window, taking the central pixel value as a threshold. A rotation invariant version (ILBP) is obtained by replacing the original $3 \times 3$ window with a circular one and considering all rotated versions of the same pattern to be equivalent [18]. This reduces the number of histogram bins (features) to 36.

ILBP [19] is an extension of the LBP. The main differences are: (1) The threshold is the mean value of the nine pixels in the window. (2) The central pixel is included in the definition of the binary patterns (in the LBP it is excluded). Hence, there are 511 possible binary patterns for the basic version (ILBP), since one of the patterns (all black pixels) is impossible by definition. The number of features is reduced to 71 in the rotation invariant case (ILBP), which is obtained as described previously.

The BGC considers the binary gradients defined between pairs of adjacent pixels lying on the peripheral closed path of a $3 \times 3$ grayscale window. This approach, despite discarding the central pixel, has been shown to be potentially superior to the LBP on a theoretical basis [20]. In this paper we used the version referred to as BGC. This descriptor, which generates 255 features in the basic version (BGC), can be easily made rotationally-invariant (BGC) in the same way as the LBP. In this case the number of features reduces to 35.

The CCR [21] differs from the previous methods in the
thresholding approach. Whereas this is local in both LBP and ILBP, it is global in the case of the CCR. This gives 512 possible binary patterns for the basic version (CCR$_{3 	imes 3}$), which reduces to 72 in the rotation invariant version (CCR$_{6 	imes 1}$).

Co-ocurrence matrices [22] estimate the joint probability distributions of gray level combinations for pixel pairs at fixed displacements and orientations. In the implementation used here we consider eight displacement vectors: {(1,0), (1,1), (0,1), (-1,1), (-1,0), (-1,-1), (0,-1), (1,-1)}. The following five features are derived from the distributions: **contrast, correlation, energy, entropy and homogeneity**. Rotation invariance is achieved by averaging the feature values over the orientations.

### 2.2. Signal processing-based methods

These are usually generated by filtering the image through suitable filter banks and computing global statistics from the filter responses. Herein we consider Gabor filtering, the Dual-Tree Complex Wavelet Transform (DT-CWT) and Ring filters.

Gabor filters are one of the most effective approaches to extracting textural features. It is commonly believed that their effectiveness is related to their ability to model the frequency-orientation decomposition performed by simple cells in the mammalian visual cortex. In our experiments we use two Gabor filter banks: one with four frequencies and six orientations and the other with six frequencies and eight orientations. In both cases the other parameters are: maximum frequency = 0.327, $\eta = 0.5$, $\gamma = 0.5$, frequency ratio = half-octave. These values have been chosen in compliance with the guidelines suggested in [23]. Texture features comprise the mean and standard deviation of the absolute value of the transformed images. Rotation invariance is obtained through DFT normalisation, generating 32 and 60 features for the two banks, respectively.

The DT-CWT has interesting properties, such as moderate redundancy and directional selectivity. It operates on the directions $\pm 15^\circ$, $\pm 45^\circ$, and $\pm 75^\circ$. To achieve consistency with the number of scales and orientations in the Gabor filtering, we use four scales and the six orientations listed above. For each sub-band the mean and the standard deviation of the absolute value of the CWT coefficients are used as texture features. The features generated at each scale are averaged over the six orientations in order to obtain rotational invariance (DFT normalisation is not recommended in this instance, due to differing sensitivities of the complex wavelets). This configuration results in $4 \times 2 = 8$ features.

Ring filters are also well known in texture analysis. Being based on circular Gaussian transfer functions, they are intrinsically invariant against rotation. In the experiments presented in this paper we employed a bank of five filters with centre frequencies 1, 2, 4, 8 and 16 cycles/image. The other filtering parameters are set as suggested in [24]. Texture features are the mean and standard deviation of the absolute value of the transformed images, resulting in a feature vector of dimension 10.

### 2.3. Morphology-based methods

Morphology-based methods extend the classical digital morphological operations to greyscale images. Two well-known methods are considered: granulometry and variogram.

Granulometry [25] is the normalised sum of the pixel values of an image, when transformed with a family of openings and closings, as a function of the size of the structuring elements. In the experiments, we use four linear structuring elements with orientations $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ for opening and closing. The dimension of the elements ranges incrementally, in steps of 4, from 2 to 50 pixels. Rotationally-invariant features are produced by averaging the four granulometry vectors corresponding to each direction, giving $13 \times 2 = 26$ features.

The variogram [26] estimates greyscale difference as a function of the distance between pixels. Given two generic pixels $x_1$ and $x_2$, the average greyscale difference is defined as $1/2E\{g(x_1) - g(x_2)\}^{1/2}$, where $g$ represents the greyscale value and $E$ the expected value. In the experiments, we considered eight variograms corresponding to the displacements $(d_1, d_2) = ((n,0), (n,n), (0,n), (-n,n), (-n,0), (-n,-n), (0,-n), (n,-n))$, for $n = 1, \ldots, 20$. For rotation invariance we computed the average variogram over the eight displacements, which yields a 20-dimensional feature vector.

### 3. EXPERIMENTS

#### 3.1. Data acquisition and preparation

An Edgetech$^1$ 500 kHz towed sidescan sonar was used to capture the acoustic imagery. Range on port and starboard channels was set at 100m and along-track samples acquired at approximately five pings (scan lines) per metre. A subset of survey line waterfall segments were downsampled in the across-track direction at five pixels per metre and slant range corrected in Coda GeoSurvey$^2$ processing software. No further geometric or radiometric corrections were applied and the waterfall image segments were output as greyscale [0, 255] Tiffs, with a fixed (software specific) five-bit radiometric resolution.

For the classification experiments, 40 image regions (20 Sabellaria and 20 non-Sabellaria) each $256 \times 256$ pixels (approximately 50 m $\times$ 50 m ground coverage) were extracted from the processed waterfall segments. The non-Sabellaria regions are further sub-classified as *Sand, Mussels, Bedforms* and *Boulders*. These classes are defined in Table 1 and the number of class instances listed. Typical appearances of the five textural classes are shown in figure 3.

---

Table 1. Summary of textural regions.

<table>
<thead>
<tr>
<th>Class</th>
<th>Image region description</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabellaria</td>
<td>Moderate Sabellaria aggregations</td>
<td>20</td>
</tr>
<tr>
<td>Sand</td>
<td>Sandy and mixed sandy sediments</td>
<td>6</td>
</tr>
<tr>
<td>Bedforms</td>
<td>Linear dunes, wavelength $\approx 2 - 5$ m</td>
<td>6</td>
</tr>
<tr>
<td>Mussels</td>
<td>Mussel beds on a sandy substrate</td>
<td>6</td>
</tr>
<tr>
<td>Boulders</td>
<td>Boulders on a sandy substrate</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 3. Examples of the five textural classes.

Table 2. Overall multi-class classification accuracy and class-specific accuracy (sensitivity) for Sabellaria

<table>
<thead>
<tr>
<th>Method</th>
<th>Overall multi-class accuracy (Sub-image size (pixels))</th>
<th>Sabellaria-specific accuracy (Sub-image size (pixels))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32 64 128 256</td>
<td>32 64 128 256</td>
</tr>
<tr>
<td>BGC1$^{L_1}$</td>
<td>62.0 78.3 89.4 82.5 70.1 87.5 95.8 83.3</td>
<td>62.0 78.3 89.4 82.5 70.1 87.5 95.8 83.3</td>
</tr>
<tr>
<td>LBP$^{L_1}$</td>
<td>62.0 72.0 75.0 62.5 78.3 83.4 86.3 75.0</td>
<td>62.0 72.0 75.0 62.5 78.3 83.4 86.3 75.0</td>
</tr>
<tr>
<td>ILBP$^{L_1}$</td>
<td>66.7 80.9 85.6 75.0 82.3 91.6 97.5 90.0</td>
<td>66.7 80.9 85.6 75.0 82.3 91.6 97.5 90.0</td>
</tr>
<tr>
<td>CCR$^{L_1}$</td>
<td>63.7 74.1 87.5 70.0 80.9 91.3 98.8 85.0</td>
<td>63.7 74.1 87.5 70.0 80.9 91.3 98.8 85.0</td>
</tr>
<tr>
<td>CGOC</td>
<td>58.2 56.9 51.9 45.0 70.9 67.8 63.8 65.0</td>
<td>58.2 56.9 51.9 45.0 70.9 67.8 63.8 65.0</td>
</tr>
<tr>
<td>Gabor 4-6</td>
<td>66.8 82.7 89.4 72.5 77.5 90.3 93.8 90.0</td>
<td>66.8 82.7 89.4 72.5 77.5 90.3 93.8 90.0</td>
</tr>
<tr>
<td>Gabor 6-8</td>
<td>68.9 83.4 88.8 82.5 79.4 89.1 92.5 95.0</td>
<td>68.9 83.4 88.8 82.5 79.4 89.1 92.5 95.0</td>
</tr>
<tr>
<td>Ring filters</td>
<td>75.5 84.1 86.3 60.0 88.0 95.8 91.7 66.7</td>
<td>75.5 84.1 86.3 60.0 88.0 95.8 91.7 66.7</td>
</tr>
<tr>
<td>DT-CWT 4-6</td>
<td>67.5 74.2 76.3 77.5 80.7 86.6 83.8 85.0</td>
<td>67.5 74.2 76.3 77.5 80.7 86.6 83.8 85.0</td>
</tr>
<tr>
<td>Granulometry</td>
<td>62.9 68.8 67.5 60.0 75.6 80.6 81.3 85.0</td>
<td>62.9 68.8 67.5 60.0 75.6 80.6 81.3 85.0</td>
</tr>
<tr>
<td>Variogram</td>
<td>59.4 69.5 66.6 65.0 76.7 87.5 86.3 85.0</td>
<td>59.4 69.5 66.6 65.0 76.7 87.5 86.3 85.0</td>
</tr>
</tbody>
</table>

Fig. 4. Variation in overall accuracy with sub-image size.

3.2. Classification and accuracy estimation

To assess the effectiveness of the methods described earlier, a multi-class supervised classification experiment was carried out using a nearest-neighbour (1-NN) classifier with $L_1$ distance. Accuracy estimation is based on leave-one-out cross validation. The choice of 1-NN is motivated by its absence of tuning parameters, its ease of implementation and by the asymptotic property that its error is bounded by twice the Bayes error as the number of samples tends to infinity.

4. RESULTS AND DISCUSSION

A potentially useful feature generation method for this task should yield good classification accuracies on relatively small sub-image sizes. Table 2 shows mean classification accuracies (%) for each method on the different sized sub-images. Results are partitioned into overall multi-class accuracies (left) and Sabellaria-specific accuracies (right) - i.e. sensitivity with Sabellaria as the target.

We define a ranking of the methods through the following pairwise comparison rule: A method outperforms another method if its accuracy is greater on each sub-image size. Assigning $+1$ for a win, $-1$ for a loss and $0$ for a tie we obtain the results shown in Table 3. The overall accuracy as a function of sub-image size for the top five and bottom ranked texture descriptors is shown graphically in figure 4.

These results clearly indicate the superiority of signal processing methods on this texture classification task. Both Gabor and ring filters produce good results, with an overall classification accuracy as high as 89.4% and an accuracy for Sabellaria-specific images exceeding 90%. These two methods also yielded some of the highest accuracies on specific classes of Sand and Bedforms. This is very useful, as in their natural habitat, Sabellaria require sand to construct their tubes and they are often found in regions where relatively strong currents (hence bedforms) exist. Features producing good classification accuracies on the Sabellaria, Sand and Bedform classes should also be effective for discriminating between these classes and performing the desired image segmentation.
The classification could be extended to include distinguishing between different periodicities in bedform textures using a Fourier Transform approach. Variations in the spatial-frequency can occur over short length scales as shown in figure 3(c). In this instance, wavelengths range from approximately 2 - 5 metres over the region, so a short-time, windowed Fourier Transform with good spatial and spatial-frequency localisation would be required. Our choice of window function for the transform would therefore be Gaussian as this optimises the joint spatial and spatial-frequency localisation. Further, we would want to analyse at multiple frequencies. Clearly though, a Gaussian windowed short time Fourier Transform applied at different centre frequencies is a Gabor expansion. The filter banks are computational implementations of this expansion and therefore have the added versatility of being able to perform this type of periodicity discrimination, should it be needed.

Among the statistical methods investigated, the ILBP and BGC1 gave good results too. As far as we know, these two methods have not been applied to sonar imagery before. Their performance here suggests that they could be useful additions to the pool of methods currently being used in the domain. Surprisingly the LBP does not appear to perform particularly well in this context. We believe this could be a consequence of the image formation process. Downsampling and radiometric compression will result in a loss of some high-frequency components that are vital for methods based on small scanning windows such as the LBP.

The COOC features were ranked bottom out of all the methods considered. Most methods exhibited a peak overall accuracy at a sub-image size of 128 pixels but as can be seen in figure 4, accuracies from the COOC features decreased monotonically by 13% over the range of sub-image sizes tested. This is most likely due to the small magnitude of the displacement vectors in relationship to the scales of the textural variations and also noise in the imagery. It is possible to achieve good results with co-occurrence features on sonar data but this requires careful parameter and feature selection for the specific data set and texture discrimination tasks. Further details of an approach to co-occurrence parameter optimisation can be found in [27].

Having investigated these feature groups, we now expect to improve classification accuracies on even smaller sized sub-images, by parameter tuning and using better machine learning methods. Importantly, this paper has demonstrated that the identification of Sabellaria, which is so surprisingly linked to future offshore power generation in many countries, is a tractable task for an automated system.

5. REFERENCES


