Acoustic Characterization of Seafloor Sediment Employing a Hybrid Method of Neural Network Architecture and Fuzzy Algorithm

Chanchal De and Bishwajit Chakraborty

Abstract—Seafloor sediment is characterized acoustically in the western continental shelf of India using the echo features extracted from normal incidence single-beam echo sounder backscatter returns at 33 and 210 kHz. The seafloor sediment characterization mainly depends on two important parameters: the number of sediment classes prevailing in the area and the selection of features having most prominent discriminating characteristics. In this letter, a method is proposed using Kohonen’s self-organizing map to estimate the maximum possible number of classes present in a given data set, where no a priori knowledge on sediment classes is available. Applicability of this method at any site is illustrated with simulated data. In addition, another method is proposed to select the three most discriminating echo features using a fuzzy algorithm. The comparison of the results with ground truth at two operating frequencies revealed that this hybrid method could be efficiently used for sediment classification, without any a priori information and applicable for a wide range of frequencies.

Index Terms—Echo feature, fuzzy C-means (FCM), seafloor characterization, self-organizing map (SOM).

I. INTRODUCTION

A
coustic methods are widely used as remote techniques in the field of marine geology, hydrographic, marine engineering, and fisheries sciences to characterize the seafloor sediments. Collection of seafloor sediment samples and their characterizations are tedious and time-consuming tasks even for a small area. Thus, the remote techniques are useful for rapid seafloor characterization in offline and online mode of operations. Usually, the sediment characterization is carried out by comparing the experimental data with theoretical model outputs [1]–[3]. An alternative approach, based on statistical analysis, uses the parameters (derived from backscatter data) called “feature vectors” [4], [5]. The uses of multivariate statistics [6], K-means algorithm [7], neuro-fuzzy classifiers [8]–[10], and neural networks applied to multibeam echo sounder backscatter data [11]–[13] and side-scan sonar images [14]–[16] have been demonstrated for seafloor classification.

In this letter, seafloor sediment is characterized using an unsupervised architecture called Kohonen’s self-organizing map (SOM) [17] and fuzzy C-means (FCM) algorithm [18] based on the echo features extracted from backscatter returns recorded in a normal incidence single-beam echo sounder at 33- and 210-kHz frequencies. The classical uses of SOM are well documented for single-beam [19]–[21] and multibeam backscatter data [11]–[13] to obtain the approximate idea [12] about the number of classes present in a data set based on the inherent discrimination ability of the feature vectors. The approximate number of classes obtained from SOM analysis may vary based on the input vector chosen for unsupervised training, which may influence the subsequent analysis with supervised fine-tuning. Therefore, a SOM-based method is proposed to estimate the possible maximum number of prevailing seafloor classes. The performance of the proposed method is tested with simulated data.

For any pattern recognition problem, selection of the most dominating features is another important task for achieving higher success in classification. However, there is no general rule to select the best feature vectors for a given data set. It is widely accepted that the feature selection is mostly application based and depends on the physical processes under study [16]. Therefore, multivariate analyses are used [6], [7], [10] as preprocessors to decide the dominating feature vectors before applying to the classifier. The fuzzy algorithms, namely, K-means [6], [7] and the FCM [22], have been in use for seafloor sediment classification assuming predefined cluster centers. A method based on FCM algorithm is proposed here to select the best features (environment specific) having dominant discriminating characteristics, using the number of cluster centers obtained from SOM (as mentioned earlier).

II. MATERIALS

Dual-frequency (33 and 210 kHz) echo waveform data were acquired using a hull-mounted Reson Navitronics NS-420 normal incidence echo sounder at 20 locations (Fig. 1) in the western continental shelf of India. The experiments were conducted in calm weather conditions [23], [24]. The seafloor depths vary between 21 and 109 m. For both the frequencies, ~600–800 echoes were acquired at each locations. Good echo waveforms are selected by visual checks [5] for further analysis. To obtain stable acoustic signals, echo waveforms are averaged using 20 successive echoes with 95% overlap. The time adjustment and power compensation of the returned signals are made [25], [26] at a reference seafloor depth of 50 m (approximate average of all spot depths). The echo features are calculated from the normalized averaged echoes.
The sediment samples were collected using the Van Veen grab at the same 20 locations (Fig. 1) to obtain the sediment ground truth. The percentage composition of sand, silt, and clay in surface sediment samples is determined in the laboratory using a particle size analyzer. The sediment type, based on Shepard’s scheme [27], reveals that four types of sediments are available (Fig. 1): clayey silt (CS), silt (Si), silty sand (SS), and sand (Sa).

Seven echo features are computed from echo waveforms, namely: backscatter strength (BS) [28], statistical skewness (StatSkew), statistical time-spread (TS), spectral skewness (SpSkew), spectral width (SpWidth), spectral kurtosis (SpKurt) [6], [29], and Hausdorff dimension (HD) [30]. These echo features for all locations together form two required input data files for two operating frequencies. There are 8975 data points (1590 CS, 2141 Si, 2003 SS, and 3241 Sa) and 6762 data points (3337 CS, 970 Si, 918 SS, and 1537 Sa) in the input data files for 33 and 210 kHz, respectively.

III. METHODS AND RESULTS

A. Estimation of Number of Cluster Centers

The unsupervised SOM classifier has the ability to classify input data based on the inherent patterns [19]. Here, SOM is applied to BS data, which is one of the most discriminating echo features [28]. The SOM architecture consists of a one-layer 2-D flat grid. A random weight matrix is initialized within +1 and −1. The output neuron number is chosen as 55. The optimum number of training data points (150) is selected from few trial runs. The normalized input data (within +1 and −1) are segmented successively with 150 data points and 100 points overlap in the moving average sense as 1–150, 51–200, and so on until the end of the data. At each iteration (t), the weights of five neurons (two neighborhood neurons on either side of the winner) were updated using the learning function 0.4/\(t^{0.2}\). The neighborhood neuron number is reduced to 1 during the training process. Training stops when the error goal (the Euclidean distance between the input vector and the neuron weights) of \(10^{-30}\) is achieved or the prespecified maximum iteration number 2000 is reached. Once the training is completed, the SOM is tested with the remaining data segments. A similar training testing process continues using the successive segments until the end of the data points. The excited neurons obtained during every testing process are plotted with respect to the neuron positions in a bar diagram (Fig. 2(a) for 210 kHz for one of the cases). If the testing shows that the winning neuron exists within the range of the trained neuron position, then it is assumed as belonging to the same class, where the data were trained earlier. Otherwise it is considered as a new class. The maximum number of classes is estimated by counting the presence of the number of prominent fired neurons (bars). Here, the highest firing of neurons (e.g., N) occurred at position 21 [Fig. 2(a)] and the surrounding 20% of N occurred at positions 6, 16, and 36. This 20% selection criterion is chosen from a few trial-and-runs. Different training-testing processes produce different numbers of classes. All these numbers are finally plotted in a histogram (Fig. 2(b) for 210 kHz), which shows there are four classes. Similar results are obtained for 33 kHz also. This simple process gives the maximum number of classes (cluster centers) present in any data set without any prior information.

B. Simulation Study and Performance of the Method

To establish the effectiveness and quantify the performance of the proposed method (in Section III-A) under controlled conditions [12], it is tested with simulated data. Although several sediment classification schemes exist to represent the sedimentary environment based on the relative proportions among different grain size particles, for simplicity the Wentworth mean grain size (\(\phi\)) grade scale [31] is chosen for simulation purposes. Accordingly, five different environments (i.e., five sediment classes) comprising coarse- to fine-grain sediments are considered: 1) coarse and medium sand (mean grain size chosen as 1 \(\phi\)); 2) fine and very fine sand (3 \(\phi\)); 3) coarse and medium silt (5 \(\phi\)); 4) fine silt (6.8 \(\phi\)); and 5) clay (8.6 \(\phi\)). The echo waveform simulations are performed at 50-m depths for 33 kHz using the theoretical model [3] (based on the composite roughness concept [1]). The pulse duration and beam width are assumed as 0.97 ms and 20°, respectively. The spectral strength (\(w_2\)), the volume scattering coefficient (\(\sigma_v\)), spectral exponent, density ratio, and sound speed ratios are estimated from the published relationships [3] based on grain size. However, in a real scenario, for a given sea bottom, the echo waveforms can vary from one data ensemble to the other [3]. To incorporate this aspect, the values of \(\phi\), \(w_2\), \(\sigma_v\) are varied within small standard deviations ±0.1 for \(\phi\) (phi units), ±0.00005 to 0.0005 for \(w_2\) (in centimeters to the fourth power), and ±0.0002 to 0.001 for \(\sigma_v\) (in per meter) around their mean values. Five sets of echo waveforms (850 echoes in
Fig. 3. (a) SOM result for various training-testing process at 33 kHz using the simulated data. (b) Histogram showing the maximum five numbers of classes obtained from simulated data at 33 kHz.

each set) are simulated. Since the calculation of BS from echo waveforms involves the echo sounder transceiver characteristics [28], the echo energies estimated from the simulated waveforms (of 10-ms duration) [6] are used for testing the algorithm. Altogether, 4250 echo energy data points, extracted from the simulated normalized waveforms, are subjected to the proposed technique (in Section III-A). The output neuron numbers 55 is sufficient to accommodate at least nine classes with five neurons in each class. Fig. 3(a) shows that the different segments of the data set, when subjected to SOM, produces different numbers of maximum classes. The histogram [Fig. 3(b)], using the numbers of maximum classes obtained from different training and testing, shows that there are five classes. This demonstrates that if the maximum number of prevailing classes in a given data set is estimated from any one of the training segments, the result may not be correct. The five simulated regions in the data set are indicated by the \( \varphi \) value at the top of Fig. 3(a).

C. Selection of Discriminating Echo Features

The selection of features is having a great importance in a wide range of classification problems and can serve as a preprocessor to select a subset of relevant features among the available ones for achieving better results. Although the use of nonlinear techniques and discriminant analysis [32] are available for features selection, principal component analysis (PCA) [33] is widely used as preprocessors for features selection [10], [34], [35]. In this section, a data-driven approach based on the FCM algorithm (using Fuzzy Logic Toolbox in MATLAB 7 [36]) is proposed to select the three best features in a given data set.

If three echo features are selected at a time, out of seven features (without regard to order), there exist 35 cases (from binomial coefficient). The FCM algorithm is used for all the 35 cases/sets with three different echo features at a time. Various parameters used in the FCM algorithm are: exponent for the membership function matrix—2.0, maximum iterations—200, minimum improvement value—1e-5. To compare the FCM output with the ground truth, the percentages of data points falling in the available classes are computed. The results for different combination of echo features are shown in Fig. 4. These analyses indicate that the success (to classify sediments correctly) is high only for the echo features set numbers 13, 14, and 15. These three feature sets (13, 14, and 15) represent the feature combinations [BS, TS, StatSkew], [BS, HD, TS], and [BS, StatSkew, HD], respectively. The overall average percentages of correct classification for feature sets 13, 14, and 15 are (86%, 84%, and 78%) and (86%, 91%, and 89%) at 33 and 210 kHz, respectively. This shows that feature sets 13 and 14 give maximum success at 33 and 210 kHz, respectively (Table I). From these trials, it can be inferred that the most dominant discriminating feature BS, when used in combination...
with TS, HD, and StatSkew, produces much better results than any other feature sets. Although the sediment classification based on echo spectral features showed very good results in other areas [29], these features could not produce good results in the present study area.

To compare the performance of the proposed feature selection method, PCA is concurrently performed after standardization of the data set with zero mean and unit variance. It is observed that the first three principal components (PCs) together explain 94% and 89% of the total variance in the data for 33 and 210 kHz, respectively. For 33 kHz, the relative contributions of SpSkew, SpKurt, StatSkew, and SpWidth to the first PC are having comparable magnitude. The second and third PCs are highly dominated by BS and HD, respectively. For 210 kHz, the first PC is having comparable contribution from SpSkew, SpKurt, StatSkew, and SpWidth, whereas the second and third PCs are having prominent contribution from TS and HD, respectively. The FCM algorithm is applied to the first three PCs. The results show that clayey silt sediments are well classified (overall average 93%) compared to silt, silty sand, and sand sediments (with averages 51%, 55%, and 56%, respectively) at 33 kHz. For 210 kHz, silt sediments are very well classified (average 97%), whereas clayey silt, silty sand, and sand sediments are poorly classified (with averages 36%, 57%, and 56%, respectively). The comparison demonstrates that the proposed data-driven approach can be used as preprocessor to increase the efficiency of any classifier.

### D. Comparison of Results With Ground Truth

The 3-D plots of FCM results for feature sets 13 and 14 are shown in Fig. 5. The results (Table I) show that the clayey silt samples are well classified for both the frequencies except at station 8, where majority of sediments are misclassified as silt at 33 kHz and partly at 210 kHz. For silty sands, both the frequencies show good results. For silty sand samples, the results at stations 11 and 17 are comparatively better than that of stations 18 and 19. For sandy sediments, the agreements with the ground truth are good at all stations, except at stations 16 and 20. At station 20 for 33 kHz and at station 16 for 210 kHz, the sediments are partly misclassified as silty sand. The overall results for silt and sand samples are comparatively better than that of mixed sediments (clayey silt and silty sand).

In this letter, the sediment type, based on Shepard’s scheme (one of the available schemes), is used as ground truth data. Since this classification scheme does not take into consideration the acoustical characteristics of sediments, the deviations of the results from the ground truth need to be viewed in the light of acoustic seafloor sediment characterization.

### E. Comparison With Another Hybrid Method

The performance of any neural networks depends on the careful selection of training sets having sufficient information to distinguish the different classes present in a given data set [16]. The applications of multistage neuro-fuzzy classifier [9], incremental fuzzy neural network and fuzzy decision trees [10] and neural networks utilizing simulated reverberation data [11], multibeam backscatter data [12], [13], and side scan imagery [14]–[16] have been demonstrated successfully for seafloor classifications. The waveforms from a single-beam echo sounder in the western continental shelf of India are for the first-time successfully classified using SOM with moving average as the preprocessor [19] and hybrid tool based on SOM and learning vector quantization (LVQ) [20]. To compare the performance of the proposed technique, the hybrid method [20] based on SOM, and the first kind of LVQ (called LVQ1), is implemented.

Following [20], the echo waveforms (250 numbers at each locations) are divided into three segments, i.e., 1–70, 71–150, and 151–250 for classifier training and testing. For SOM training, three different echo waveforms (in the aforementioned segments) such as 20th, 80th, and 170th are chosen. In the SOM architecture, a 200 × 1 input grid (input waveforms resampled to 200 points) and 20 × 1 output grid (sufficient for mapping four sediment classes with three neurons) are chosen optimally. For each sediment type, data from two different locations are used for SOM training. To improve the performance of the classifier, the supervised architecture LVQ1 is implemented using SOM trained weights. A total of 70 waveforms (20 from first two segments and 30 from the last segment) are used for LVQ1 training. The overall averages of the three segments of correct classification obtained from SOM and LVQ1 testing (except training samples) are given in Table II. Asterisks (*) indicate the waveforms from these locations are used for training and the remaining waveforms are used for testing. The sediment types obtained from the hybrid method are given in

---

**Table II: Results Using Hybrid Architecture (SOM and LVQ1) Utilizing Echo Waveforms for 33 and 210 kHz**

<table>
<thead>
<tr>
<th>Stn No.</th>
<th>True Sed. Type</th>
<th>Percentage Classification (SOM)</th>
<th>Percentage Classification (LVQ1)</th>
<th>Sediment Type Obtained from the Hybrid Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>33 kHz</td>
<td>210 kHz</td>
<td>33 kHz</td>
<td>210 kHz</td>
</tr>
<tr>
<td>1</td>
<td>CS</td>
<td>33*</td>
<td>10*</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>CS</td>
<td>54</td>
<td>7</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>CS</td>
<td>88</td>
<td>13</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>CS</td>
<td>54*</td>
<td>100*</td>
<td>87*</td>
</tr>
<tr>
<td>5</td>
<td>CS</td>
<td>56</td>
<td>8</td>
<td>72</td>
</tr>
<tr>
<td>8</td>
<td>CS</td>
<td>54*</td>
<td>18*</td>
<td>99*</td>
</tr>
<tr>
<td>13</td>
<td>CS</td>
<td>23</td>
<td>51</td>
<td>43</td>
</tr>
<tr>
<td>9</td>
<td>Si</td>
<td>2</td>
<td>73</td>
<td>5*</td>
</tr>
<tr>
<td>10</td>
<td>Si</td>
<td>27*</td>
<td>44*</td>
<td>40*</td>
</tr>
<tr>
<td>14</td>
<td>Si</td>
<td>6*</td>
<td>42*</td>
<td>23</td>
</tr>
<tr>
<td>11</td>
<td>SS</td>
<td>42</td>
<td>14</td>
<td>32</td>
</tr>
<tr>
<td>17</td>
<td>SS</td>
<td>16*</td>
<td>5*</td>
<td>43</td>
</tr>
<tr>
<td>18</td>
<td>SS</td>
<td>21</td>
<td>22*</td>
<td>46*</td>
</tr>
<tr>
<td>19</td>
<td>SS</td>
<td>5*</td>
<td>2*</td>
<td>70*</td>
</tr>
<tr>
<td>6</td>
<td>Sa</td>
<td>23*</td>
<td>10</td>
<td>55</td>
</tr>
<tr>
<td>7</td>
<td>Sa</td>
<td>9</td>
<td>20</td>
<td>59</td>
</tr>
<tr>
<td>12</td>
<td>Sa</td>
<td>6</td>
<td>14*</td>
<td>57</td>
</tr>
<tr>
<td>15</td>
<td>Sa</td>
<td>7</td>
<td>10</td>
<td>85*</td>
</tr>
<tr>
<td>16</td>
<td>Sa</td>
<td>10*</td>
<td>8*</td>
<td>89*</td>
</tr>
<tr>
<td>20</td>
<td>Sa</td>
<td>20</td>
<td>4</td>
<td>24</td>
</tr>
</tbody>
</table>
the last two columns of Table II. The neuron firing patterns at locations 11, 17, 18, 7, and 12 show mixed response among the four available classes and denoted by “Mixed.” The percentage misclassifications for stations 9, 10, 14, 11, and 20 are also mentioned in the last two columns in Table II.

Although the results of the hybrid method [20] show average success in classifying sediments, the applicability of the method is limited in a scenario where ground truth information is available to carry out the unsupervised training and supervised fine tuning. Therefore, in that scenario (in the absence of ground truth), the proposed method is having great advantages and can be considered as another efficient tool.

IV. CONCLUSION

A hybrid method is proposed here to characterize seafloor sediments in the absence of any ground truth information. A method using SOM is illustrated to estimate the presence of maximum possible number of classes. A data-driven approach based on a fuzzy algorithm is then proposed to select three dominant discriminating features. The results revealed that BS when used in conjunction with StatSkew, TS, and HD provides better results. The performance of the classifier using 210 kHz is marginally better than that of 33 kHz.

The method utilizing SOM can be used first to estimate the number of cluster centers, corresponding to different unknown seafloor types and subsequently sampled at few distinguished locations to ascertain their physical nature (ground truth). Further classification can be extended to other similar area using the proposed hybrid method. The results obtained at two different operating frequencies illustrate its applicability to a wide range of frequencies usually available in normal incidence echo sounder. The proposed feature selection process can also be efficiently used in conjunction with other classification tools.

In near-real-time data processing, the proposed method can be used without doing any multivariate analysis to save the processing time.

REFERENCES