Extraction of visual texture features of seabed sediments using an SVDD approach

Yan Li\textsuperscript{a}, Shijie Liu\textsuperscript{a,b}, Puqiang Zhu\textsuperscript{a}, Jiancheng Yu\textsuperscript{a}, Shuo Li\textsuperscript{b,⁎}

\textsuperscript{a} State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China
\textsuperscript{b} University of Chinese Academy of Sciences, Beijing, China

\section*{A R T I C L E I N F O}

\textbf{Keywords:}
Support vector domain description
Classification of seabed sediments
Texture features analysis
Perception of seabed environment
Gray-level co-occurrence matrix
Self-organizing map

\section*{A B S T R A C T}

Perception of the seabed environment is an important capability of autonomous underwater vehicles. This paper focuses on defining and extracting robust texture features from visual images that lead to useful and practical automated identification of the types of seabed sediments. The visual texture features are described by using a gray-level co-occurrence matrix (GLCM) and fractal dimension, after which an unsupervised learning method, self-organizing map (SOM), is adopted to evaluate the validity of features descriptors on three types of seabed sediments. Subsequently, a kernel-based approach that exhibits robustness versus low numbers of high-dimensional samples, named support vector domain description (SVDD), is applied to classify the types of seabed sediments. In comparison with state-of-the-art classifiers, the experimental results demonstrated the effectiveness of the SVDD on the classification of seabed sediments.

\section*{1. Introduction}

Over 70 percent of the surface of the earth is covered by oceans where extremely large amounts of mineral and ecological resources are deposited. Approximately 3 trillion tons of polymetallic nodules of the global reserves are proved to exist in the ocean, according to previous studies and field surveys. More than half of these polymetallic nodules are located in the Pacific Ocean, which is shared by numerous countries over the world. To reduce conflicts in developing marine resources and protect marine ecological environments, marine spatial planning initiatives are being conducted globally. Exploring marine resources is necessary and investigation and mapping of the seabed structure across multiple spatial scales are the key issues to enhance policies of effective oceanic environmental management and to guide the development of a strategy for marine resources.

Autonomous Underwater Vehicles (AUVs) is one of the essential tools to explore the ocean and are widely applied to ocean surveys including seabed mapping and oceanographic measurements (Feng et al., 2011). Submersible AUVs can provide both high-quality and high-resolution environmental data comparing with ship-based collection. The advantage of investigation using AUVs is that these vehicles could be equipped with various sensors (e.g., sonar, camera, CTD, or biosensors) for \textit{in situ} monitoring of the underwater environment. An imaging camera could provide high-resolution visual information as an intuitive tool for analyzing environmental parameters on the sea floor, such as seabed structures. Image-based identification of seabed sediments could extend the ability of AUVs for autonomous seabed surveying. The accurate and fast identification of seabed sediments is a preliminary study to achieve autonomous perception for AUVs. The automatic image identification methods could also be applied to other surveying tasks carried out by AUVs.

In previous studies, most seabed surveys were conducted by utilizing acoustic techniques based on sonar systems such as single beam, side-scan, multi-beam and acoustic ground discrimination systems (AGDS) (Collier and Brown, 2005; Ji et al., 2013; Ji and Liu, 2015). In addition, physically motivated echo features were proposed to identify seabed types (Rodriguez-Pérez et al., 2014). Orlowski et al. introduced two energies they termed the roughness index and hardness index as features to classify the type of seabed (Orlowski, 1984). Subsequently, these two variables and other indexes built from them were successfully applied to sea floor studies for various reasons (Strong and Service, 2011; Grave et al., 2000; Greenstreet et al., 1997; Siwabessy et al., 2000; Bates and Whitehead, 2001; Satyanarayana et al., 2007; Serpetti et al., 2011). A numbers of computational methods were presented to extract acoustic features, such as the echo duration, skewness, wavelet coefficients, and fractal dimension, for seabed segmentation (Biffard et al., 2010; van Walree et al., 2005; Tegowski et al., 2003). Furthermore, multivariate statistical analysis methods were utilized to achieve supervised or unsupervised classifications of the seabed based on the acoustic features.
However, acoustic features are difficult to interpret and require a ground-truth, which is often used with the acquisition of sample grabs or underwater photography to establish the true seabed type (Blondel and Murton, 1997). Nevertheless, acoustic systems have disadvantages such as high cost, and a lack of color and textures.

Visual techniques provide visual systems that deliver abundant information that could enhance the autonomous perception ability of intelligent robot systems (Tang et al., 2009). The rapid development of imaging techniques has resulted in robotic vision systems becoming common exploration devices installed on underwater vehicles such as AUVs and ROVs (Remote Operated Vehicles). This has enabled these vehicles to complete various missions for engineering and science studies, e.g., object detection and tracking (Kia and Arshad, 2005), ecological studies (Armstrong et al., 2006; Singh et al., 2004), localization and navigation (Hao et al., 2006; Gracias et al., 2003) and seabed mapping (Rzhanov et al., 2000).

However, the visual recognition of seabed sediments has not been extensively studied. In recent years, artificial intelligence techniques were successfully applied to identify submarine targets. Especially, support vector machines (SVMs) are particularly appealing for the data classification owing to their ability to generalize well even with limited training samples and producing higher classification accuracy than conventional methods (Mountrakis et al., 2011). Support vector domain description (SVDD), which is inspired by SVMs (Vapnik, 1999), is one of the most well-known support vector learning methods for the one-class problem. A one-class classifier can be used to solve a two-class classification problem, where each of the classes has a special meaning (Pan et al., 2009). It can also be generalized to multi-class classification by assigning one of the classes as target and the remaining ones as outliers. Compared with the hyperplanes of SVMs, the SVDD only uses training data from the class of interest and aims to calculate an optimal hypersphere, which ensures the capability to reject samples from any of the other classes (outliers) (Tax and Duin, 1999, 2004; Tax and Juszczak, 2003). Since hyperspheres only express a limited class of regions, kernels are frequently utilized in SVDD to enhance the effectiveness by mapping objects from the input space to a high dimensional feature space.

A number of applications have been developed since the emergence of SVDD. In (Lee et al., 2006), the SVDD method is extended for low-resolution face detection. The data belonging to the given prototype facial images are used for training SVDD hyperspheres that represent normal faces in feature space. For a new input low-resolution facial image, feature vectors are projected onto a spherical decision boundary of the trained SVDD. This method can process either images that existed in the set of training images or entirely new images. In (Sanchez-Hernandez et al., 2007), SVM- and SVDD-based approaches were applied to process remote sensing images to identify habitats of the European Union. The results indicated that the SVDD-based approach requires less training resources and higher overall classification accuracies (95.2%) are obtained compared with the overall accuracy by SVM-based approaches (92.0%). Therefore, a modified classification scheme of SVDD is proposed to solve multi-class detection problems for the identification of seabed sediments. A gray-level co-occurrence matrix (GLCM) and fractal dimension are adopted as the feature descriptors for seabed images. The proposed identification scheme based on an SVDD classifier demonstrated shows more accurate performance in identifying three types of seabed sediments compared with SVM classifiers.

The sections of this paper are organized as follows: In Section 2, texture features extractions based on GLCM and fractal theory are introduced. An overview of SVDD and the scheme used to recognize the three seabed types are presented in Section 3. The experiments and results are described in Section 4 and the paper is concluded in Section 5.

2. Texture features extraction and evaluation methods

2.1. Image acquisition

In the deep sea, the seabed is covered by many types of sediments that mainly consist of polymetallic nodules, sand, and benthic communities (Fig. 1). These three types of seabed sediments have their respective visual characteristics. In general, images of sediments often lack the sharp pronounced features that are useful for feature extraction. The regions of benthic communities are often desultorily and densely covered with shellfish and crabs, which are the most common organisms on the deep seabed. Compared with the regions of benthic communities, the structure of regions of polymetallic nodules is relatively simple, with numerous cobblestone-like polymetallic nodules scattered over the seabed. The regions of sand have no obvious visual structure.

In this study, these three typical kinds of sedimentary types are selected to verify the feasibility of the method proposed. The images of the seabed sediments were captured in a top-down view of the seabed by a high-definition camera (Kongsberg maritime* OE14-502D) installed on a 7000 m-class manned submersible named Jiaolong on a scientific expedition in the South China Sea (Cui, 2013). The computational cost for the practical seabed survey in the field, was reduced by normalizing the images to the size 320 × 240 pixels.

2.2. Description of texture features

The texture features of sediments were defined and extracted by GLCM and fractal theory. GLCM as an estimate of a joint probability density function of gray level pairs in an image can be expressed as:

\[ P_{ij}, (i, j = 0, 1, 2, ..., L - 1) \]  

where \( i, j \) indicate the gray level of two pixels of the image, \( L \) is the gray level and \( \mu \) is the positional relation of the two pixels. Usually, we can choose \( 0, 45, 90 \) and \( 135 \) as the four directions. Haralick proposed 14 statistical features extracted from GLCM (Haralick et al., 1973).

Considering the computational complexity of the classification of seabed sediments, four GLCM features are chosen in this work including energy (also known as the angular second moment), correlation, entropy and contrast. The definitions of the four features are obtained by the following equations.
Energy: 
\[ f_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P^2(i, j) \]  
(2)

Correlation: 
\[ f_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (ij)P(i, j) - \mu_x \mu_y \sigma_x \sigma_y \]  
(3)

where
\[
\begin{align*}
\mu_x &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j), \\
\mu_y &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - \mu_x)^2 P(i, j), \\
\sigma_x &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (j - \mu_y)^2 P(i, j), \\
\sigma_y &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (j - \mu_y)^2 P(i, j)
\end{align*}
\]

Entropy: 
\[ f_3 = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j) \log(P(i, j)) \]  
(4)

Contrast: 
\[ f_4 = \sum_{i=0}^{L-1} \left\{ \left( \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j) \right)^2 \left| j - i \right| = n \right\} \]  
(5)

In addition to GLCM, the fractal dimension is also adopted to represent the texture feature of seabed images. Self-similar objects and phenomena can be described by a non-integer dimension known as the fractal dimension to show the irregular structure of objects (Mandelbrot, 1983). The fractal dimension measures the degree of irregularity and complexity of an object to describe the self-similarity of the object (Tricot, 1995). Therefore, the concept of fractal dimension can be useful in the analysis of the texture of images.

The box-counting approach, which was defined in (Russell et al., 1980), is the widely used and common way to estimate the fractal dimension (Sarkar and Chaudhuri, 1994). The fractal dimension \( FD \) of an image can be derived from the relation (Tricot, 1995):
\[ FD = \frac{\log(N_r)}{\log(1/r)} \]  
(6)

where, \( r \) is the size ratio and \( N_r \) is the number of boxes with size ratio \( r \) inside which at least one point from the attractor lies.

### 3. Classification method based on SVDD

SVDD is a one-class classifier based on the principles of SVM, which is particularly effective in the presence of incomplete training data. A brief review of the SVM algorithm is provided in the following part to explain the fundamental concept of the proposed scheme. A detailed description and the process of the SVM can be found in the literature (Wallraven et al., 2003). SVM is typically a supervised classification and the process of the SVM can be found in the literature (Tax and Duin, 2004).

In order to allow the possibility of outliers in the training set, the distance from \( x_i \) to the center \( a \) should be strictly smaller than the minimum radius \( R \) of the hypersphere, but larger distances should be penalized. Therefore, a slack variable \( \xi_i \) should be introduced and the minimization problem changes into:
\[ F(R, a, \xi) = R^2 + C \sum_i \xi_i \]  
(8)

with constraints that almost all objects are within the sphere:
\[ \|x_i - a\|^2 \leq R^2 + \xi_i, \quad \xi_i \geq 0, \quad \forall \ i \]  
(9)

where, the parameter \( C \) controls the trade-off between the volume of the sphere and the number of target points excluded by the sphere as errors.

To solve this constrained optimization problem, constraints (9) can be incorporated into (8) by using Lagrange multipliers:
\[ L(R, a, \alpha, \gamma, \xi) = R^2 + C \sum_i \xi_i - \sum_i \alpha_i \times \]  
\[ (R^2 + \xi_i - (\|x_i\|^2 - 2a \cdot x_i + \|a\|^2)) - \sum_i \gamma_i \xi_i \]  
(10)

with the Lagrange multipliers \( \alpha_i \geq 0 \) and \( \gamma_i \geq 0 \). \( L \) should be minimized with respect to \( R, a, \xi \) and maximized with respect to \( \alpha_i \) and \( \gamma_i \). Setting partial derivatives of \( L \) to zero gives the constraints:
\[ \frac{\partial L}{\partial R} = 0: \quad \sum_i \alpha_i = 1 \]  
(11)

\[ \frac{\partial L}{\partial a} = 0: \quad a = \sum_i \alpha_i x_i \quad \sum_i \alpha_i = 1 \]  
(12)

\[ \frac{\partial L}{\partial \xi_i} = 0: \quad C - \alpha_i - \gamma_i = 0 \]  
(13)

From (13), we obtain \( \alpha_i = C - \gamma_i \). Considering the constraints \( \alpha_i \geq 0 \) and \( \gamma_i \geq 0 \), a new constraint \( 0 \leq \alpha_i \leq C \) and Lagrange multiplier \( \gamma_i \) also can be removed. With this new constraint and substituting (11)–(13) into (10) results in:
\[ L = \sum_i a_i (x_i, z) - \sum_i \alpha_i a_i (x_i, x_j) \]  
(14)

For a one-class classification, it is possible to calculate the center of the hypersphere \( \alpha \); thus, it is easy to decide whether a new test sample data is accepted inside the description. This requires the calculation of the distance from the test sample data \( z \) to the center of the hypersphere. Test sample data \( z \) is accepted within the hypersphere when this distance is smaller than or equal to the radius:

\[ ||z - \alpha||^2 = (z \cdot z) - 2 \sum_i a_i (z \cdot x_i) + \sum a_i (x_i x_i) \leq R^2 \]  
(15)

where \( R \) is the distance from the center of the hypersphere \( \alpha \) to the boundary. Support vectors that fall outside the description \( (\alpha = C) \) are excluded (Tax and Duin, 2004).

In this work, the classification targets are three types of sediments. This problem could not be solved by the one-class classification scheme. Consequently, a scheme of integrating multiple one-class classifiers of SVDD is presented to solve the multi-classes classification problem. Here, three separating hyperspheres are be constructed by training each type of sediments data as target data respectively.

When implementing the three one-classifiers that were built based on SVDD to classify the types of seabed sediments (Fig. 4), it is conceivable that most sample data fall inside the single hypersphere. This indicates that most sample data could be classified correctly. However, it is important to note that some singular data may fall outside of all hyperspheres or inside of more than one hypersphere. It signifies that these few data points are excluded by all classes or accepted by more than one class. Therefore, a discrimination function is presented to calculate the minimal distances from the data points to the centers of the hyperspheres. For a given sample data point \( z \), the minimum distance to the center of hypersphere is used to decide whether \( z \) belongs to one class.

\[ z \in \text{Class}_i, \quad s. t. \quad \min ||z - \alpha_i|| \]  
(16)

4. Experimental results and discussion

In this work, nine texture factors were selected in the classification experiments to identify the type of seabed sediments. These factors include the mean and the standard deviation of four texture features (energy, correlation entropy and contrast) extracted using GLCM in four directions \( (0, 45^\circ, 90^\circ, 135^\circ) \). Furthermore, the fractal dimension of images of seabed sediments is regarded as another factor.

4.1. Feasibility testing of texture factors

To examine the feasibility of these proposed texture factors of seabed sediment images for classification purposes, an unsupervised learning model in the form of a self-organizing map (SOM) is employed. Compared to other neural network models, the SOM performs a nonlinear projection of data onto a two-dimensional space and can provide a patterned map of input data trained with unsupervised learning (Kohonen, 2001). Through preliminary training, nodes sized \( 8 \times 6 \) were used in this study. A detailed description and the process of the SOM can be found in the literature (Kohonen, 2001).

Partial image data sets of sediments were randomly selected to test the texture features for sediments classification. The selected data set consists of 18 images of polymetallic nodules, 19 images of benthic communities and 44 images of sandy regions (labeled as M1-M18, B1-B19, and S1-S44, respectively). After the randomly selected data with nine texture factors were trained by the SOM, sample data belonging to the same type was clustered and classified into three groups (marked with ellipses), as shown on the map (Fig. 5). The cluster analysis with SOM indicates that the proposed texture features present a feasible method for the classification of seabed sediments.

In total, data consisting of 1500 images of seabed sediments including 500 images in each category were used in the experiment. These image data were manually labeled as belonging to three categories before the experiments were carried out. Considering the advantageous ability of a support vector to generalize well even with a limited number of training samples, the training and test data sets were constructed as in Table 1.

The SVDD method was used to classify and identify the type of seabed sediment of interest in the marine exploring mission. To validate the performance of the proposed method, three one-classifiers based on SVDD were built for each respective type of seabed sediment respectively. For each one-classifier, the target data is defined as the 200 training data belonging to the same category, whereas the remaining 400 training data belonging to the other categories are considered as outliers. The SVDD allows for some target data to be excluded from the sphere description, the fraction of the target that is rejected is set as 0.1, and the RBF kernel is set as 0.5. Subsequently, all of the test data were sent to every one-classifier based on SVDD, the overall performance is higher than 82% and the average accuracy rate

### Table 1

<table>
<thead>
<tr>
<th>Types</th>
<th>Benthic community</th>
<th>Polymetallic nodule</th>
<th>Sand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training data</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Test data</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>
is 88.9% (in Tables 2, 92.2%, 93.0%, and 91.6%, for benthic communities, polymetallic nodules, and sand, respectively).

In this study, a combination strategy is proposed to carry out the multi-classification on the types of seabed sediments (Fig. 6). The test data is pre-classified using the three one-classifiers built for each type of seabed sediment, respectively. Although most of the test data would be classified correctly and fall inside of the three trained hyperspheres, some singular data may fall outside of all hyperspheres or inside of more than one hypersphere. As introduced above, these singular data will be re-classified with the discrimination function by comparing their distances to the centers of hyperspheres. For each separating hypersphere of SVDD, the radius of the hyperspheres could be calculated. In addition, the distances from each testing data to the centers of hyperspheres (noted as $R_0$, $R_1$, and $R_2$ for benthic communities, polymetallic nodules, and sand, respectively) are also obtained as shown in Table 3. For the test data belonging to the type benthic community, the average distance to the center of the hypersphere built with training data belonging to the type benthic community is 0.881, which is shorter than to the centers of the other hyperspheres. Moreover, the average distance is very close to the radii of the hyperspheres. The same phenomenon is found for the test data belonging to the other two types of seabed sediments. These results illuminate that the data from different types of seabed sediment are clustered using SVDD.

The discrimination function is used to evaluate our proposed multi-classification strategy based on SVDD, by conducting the classification tests on the same test data by the proposed multi-classification strategy and SVM. The SVM model is trained using the same training data as indicated in Table 1. In total, 900 images of the three types of sediments are tested and the classification accuracies are presented in Table 4. Most of the test data were correctly categorized into three classes with both SVM and the proposed method. After pre-classification by the three respective one-classifiers, 630 data images were identified correctly, and the accuracy rate of pre-classification is 70%. In addition, 91 data images fell outside all hyperspheres and 160 data images fell inside of two hyperspheres. These singular data are re-classified with the discrimination function in the proposed strategy, thereby reducing the number of misclassified data images to 60 whereas the accuracy rate is increased to 93.3%, i.e., 3.3% higher than SVM.

The misclassified image samples were investigated and the reasons leading to misclassification were analyzed. On the one hand, since the seabed sediments show various texture features and complex physical composition, more than one type of seabed sediment could be combined in one image, which would cause the identified type to possibly differ from subjective human judgment. On the other hand, the texture features of partial misclassified images are similar with both types of seabed sediments, i.e., the image of sparse shells lying on the sand is very similar in texture to an image of polymetallic nodules lying on the sand.

### Table 2

<table>
<thead>
<tr>
<th>One-classifier based on SVDD</th>
<th>Target</th>
<th>Test data</th>
<th>Misclassification</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benthic Community</td>
<td>200</td>
<td>400</td>
<td>160</td>
<td>82.2%</td>
</tr>
<tr>
<td>Polymetallic Nodule</td>
<td>200</td>
<td>400</td>
<td>63</td>
<td>93.0%</td>
</tr>
<tr>
<td>Sand</td>
<td>200</td>
<td>400</td>
<td>76</td>
<td>91.6%</td>
</tr>
<tr>
<td>Average</td>
<td>N/A</td>
<td></td>
<td></td>
<td>88.9%</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Sediment Type</th>
<th>$R_0$</th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$R_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benthic community</td>
<td>0.880</td>
<td>0.912</td>
<td>1.069</td>
<td>0.895</td>
</tr>
<tr>
<td>Polymetallic nodule</td>
<td>0.751</td>
<td>0.897</td>
<td>0.939</td>
<td>0.918</td>
</tr>
<tr>
<td>Sand</td>
<td>0.580</td>
<td>0.692</td>
<td>0.759</td>
<td>0.689</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Classification results by SVM and proposed strategy.</th>
<th>SVM</th>
<th>Proposed strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample amount</td>
<td>900</td>
<td></td>
</tr>
<tr>
<td>Accurate classification</td>
<td>810</td>
<td>840</td>
</tr>
<tr>
<td>Accurate rate</td>
<td>90%</td>
<td>93.3%</td>
</tr>
</tbody>
</table>

5. Conclusions

The aim of this study was to develop an effective classification strategy to recognize the three types of seabed sediments. The visual texture features extracted from images by using GLCM and fractal dimension contributed to the excellent performance of the classification of the three types of seabed sediments. The feasibility of feature definition was proven by the SOM model. Moreover, the proposed strategy based on SVDD with a discrimination function was demonstrated as a feasible method in classifying the three types of seabed sediments with a limited number of images as training data and was superior to SVM. According to the experimental results, SVDD with a discrimination function shows outstanding ability to minimize the misclassification rates and is the optimal option for recognition of seabed sediments given limited prior knowledge. The experimental results could serve as practical environmental perception for AUVs in an ocean exploring survey. In future work, combining visual sensing with acoustics data should be considered to improve the accuracy and stability of underwater surveys.
Acknowledgements

This work was in part supported by the National Key Research and Development Program of China under Grant 2016YFC0300801 and in part supported by the project titled “R & D Center for Underwater Construction Robotics,” funded by the Ministry of Ocean and Fisheries (MOF) and Korea Institute of Marine Science & Technology Promotion (KIMST), Korea (PJ200539). The work of this article is also in part supported by National Natural Science Foundation under grants 61233013, 41376110 and a self-sponsored project of the State Key Laboratory of Robotics at Shenyang Institute of Automation under grant 2013-Z13.

References


