A Study on the Recognition of Seabed Environments Employing Sonar Images

by Yasuhiro Tan

Supervisor; Prof. Seiji Ishikawa, Ph.D.

A thesis presented in partial fulfillment of
the requirements for the degree of
DOCTOR OF PHILOSOPHY

Department of Mechanical and Control Engineering
Graduate School of Engineering
Kyushu Institute of Technology, Japan

March 2014

Table of Contents

Acknowledgements

This research has been carried out during the author's Doctoral Degree Program (2010-2014) at the Graduate School of Engineering, Kyushu Institute of Technology, under the guidance of Professor Seiji Ishikawa, PhD, at the Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Japan.

The author wishes to express his deepest gratitude and grateful acknowledgement to his Supervisor, Prof. Seiji Ishikawa for his constant guidance, valuable advice, and constructive criticism with encouragement throughout all the study period at the Doctoral Degree Program in Mechanical and Control Engineering including this research study.

The author also wishes to express his sincere gratitude and acknowledgement to Prof. Tan, who served for her valuable suggestions, comments and guidance during the period of his study. The author records his thanks to all lab members, staff members, other students of Kyushu Institute of Technology and all friends for the continuous help and assistance that enabled him to follow through the courses at the Doctoral Degree Program in Mechanical and Control Engineering.

Finally, grateful acknowledgement is extended to his family, for their eternal love, dedications, constant warm encouragement and support throughout his education and study in the Doctoral Degree Program in Mechanical and Control Engineering.

Yasuhiro Tan

March 2014

Abstract

The ocean accounts for approximately 70% of the area on the earth, and the water as well as coastal areas sustain many species including humans. Ocean resources are used for fish farming, land reclamation, and a variety of other purposes. Seabed resources such as oil, natural gas methane hydrates, and manganese nodules are still largely unexploited on the bottom of the sea. Maps are critical to development activities such as construction, mining, offshore drilling, marine traffic control, security, environmental protection, and tourism. Accordingly, more topographic and others types of mapping information are needed for marine and submarine investigations. Both waterborne and airborne survey techniques show promise for collecting data on marine and submarine environments, and these techniques can be classified into four main categories. First, remote sensing by satellites or aircraft is a widely used technique that can yield important data such as information on sea levels and coastal sediment transport. Second, investigations may collect direct information by remotely operated vehicles (ROVs), autonomous underwater vehicles (AUVs), and divers. While the quality of data obtained from these techniques is high, the data obtained are often limited to relatively shallow and small geographic areas. Third, sediment profile imagery can be used to collect photographs that contain detailed information about the seabed. Lastly, acoustic investigations that use sonar are popular in marine mapping studies, especially in coastal areas. In particular, acoustic investigations that employ ultrasound technology can yield rich information about variations in bathymetry.

Unlike air, water has physical properties that make it difficult for light or electromagnetic waves to pass through. However, sound waves propagate readily in water. Therefore, sound waves are used in a wide range of technical applications to detect underwater structures that are difficult to observe with light-based techniques. In the dark

depths of the ocean, the use of acoustic technology is essential. The development of marine acoustic technology is expanding in modern times. In addition to the basic physics related to acoustic waves, much research has been dedicated to other basic and applied fields such as electronics, physical oceanography, signal processing, and biology. The realization of new sonar systems that utilize advanced detection algorithms can be expected to contribute to major breakthroughs in oceanographic research that require deployment to novel marine environments and other areas of natural resource interest.

In this study, the author focuses on side-scan sonar, which is one of the imaging technologies that employs sound to determine the seabed state, to conduct research on imaging algorithms for discrimination. The proposed method for discrimination was coupled to a high-speed detection method for installed reefs on the seabed. This method is also capable of detecting unknown objects with Haar-like features during object recognition of rectangular regions of a certain size via machine learning by AdaBoost and fast elimination of non-object regions on the cascade structure. Side-scan and forward looking sonars are some of the most widely used imaging systems for obtaining large-scale images of the seafloor, and their application continues to expand rapidly with their increasing deployment on AUVs. However, it can be difficult to extract quantitative information from the images generated from these processes, in particular, for the detection and extraction of information on the objects within these images. Hence, this study analyzes features that are common to most undersea objects projected in side-scan sonar images to improve information processing. By using a technique based on the *k*-means method to determine the Haar-like features, the number of patterns of Haar-like features was minimized and the proposed method was capable of detecting undersea objects faster than current methodology. This study demonstrates the effectiveness of this method by applying it to the detection of real objects imaged on the seabed (i.e., sandy ground and muddy ground).

Attempts are made as well to automate the proposed method for discriminating objects lying on the seafloor from surficial sediments. During undersea exploration, a thorough understanding of the state of the seafloor surrounding objects of interest is important. Therefore, a method is proposed in this study to automatically determine seabed sediment characteristics. In traditional methods, a variety of techniques have been used to collect

information about seabed sediments including depth measurements, bathymetry evaluations, and seabed image analyses using the co-occurrence direction of the gray values of the image. Unfortunately, such data cannot be estimated from the object image itself and it can take a long time to obtain the required information. Therefore, these techniques are not currently suitable for real-time identification of objects on the seafloor. For practical purposes, automatic techniques that are developed should follow a simple procedure that results in highly precise and accurate classifications. The technique proposed here uses the subspace method, which is a method that has been used for supervised pattern recognition and analyses of higher-order local autocorrelation features. The most important feature of this method is that it uses only acoustic images obtained from the side-scan sonar. This feature opens up the possibility of installing this technology in unmanned small digital devices. In this study, the classification accuracy of the proposed automation method is compared to the accuracy of traditional methods in order to show the usefulness of the technology. In addition, the proposed method is applied to real-world images of the seabed to evaluate its effectiveness in marine surveys.

The thesis is organized as follows. In Chapter 1, the purpose of this study is presented and previous studies relevant to this research are reviewed. In Chapter 2, an overview of underwater sound is given and key principles of sound wave technology are explained. In Chapter 3, a new method for detecting and discriminating objects on the seafloor is proposed. In Chapter 4, the possibility of automating the discrimination method is explored. Finally, Chapter 5 summarizes the findings of this study and proposes new avenues for future research.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 General introduction

If an expert engineer observes a seafloor acoustic image such as shown in **Figure 1-1**, the engineer may be able to identify an object present on the seabed and estimate its size at the same time. Furthermore, they may be able to determine if the seafloor is composed of sand or rock. This situation may lead us to think, at first glance, that this is an easy task to accomplish, but in fact, obtaining this level of understanding is not a trivial matter. Actually, such undersea assessments are implemented by the use of complex knowledge and vast experience. Experience plays a crucial role in undersea investigations, for example, when conducting surveys of wrecks, locating unexploded ordnances, or determining whether or not the seafloor is suitable for offshore drilling. Consequently, a thorough understanding of visual information is indispensable to making appropriate decisions. However, problems can arise during the determinations. In other words, uncertainties in the accuracy of the data can vary by the individual engineers employed and the type of technology used. Furthermore, highly accurate surveys can be very costly to implement. Supportive computer vision techniques may help to resolve some of these problems.

Image recognition and computer vision represent a field of research that attempts to use computers and cameras to acquire, process, and analyze information in a manner similar to what can be achieved with human vision. Each of these phases can be difficult to achieve using computers. For example, computers sometimes have difficulty detecting moving objects. Computers can also have difficulty determining the composition and distribution of seabed sediments. This can be especially problematic when there is ambiguity

in the image between the seabed sediments and the target object (i.e., the outline is not clear).

This paper investigates the use of computer vision technology to survey seabed sediments and detect submarine objects. As a starting point, the basics of ocean acoustic technology and side-scan sonars are discussed.

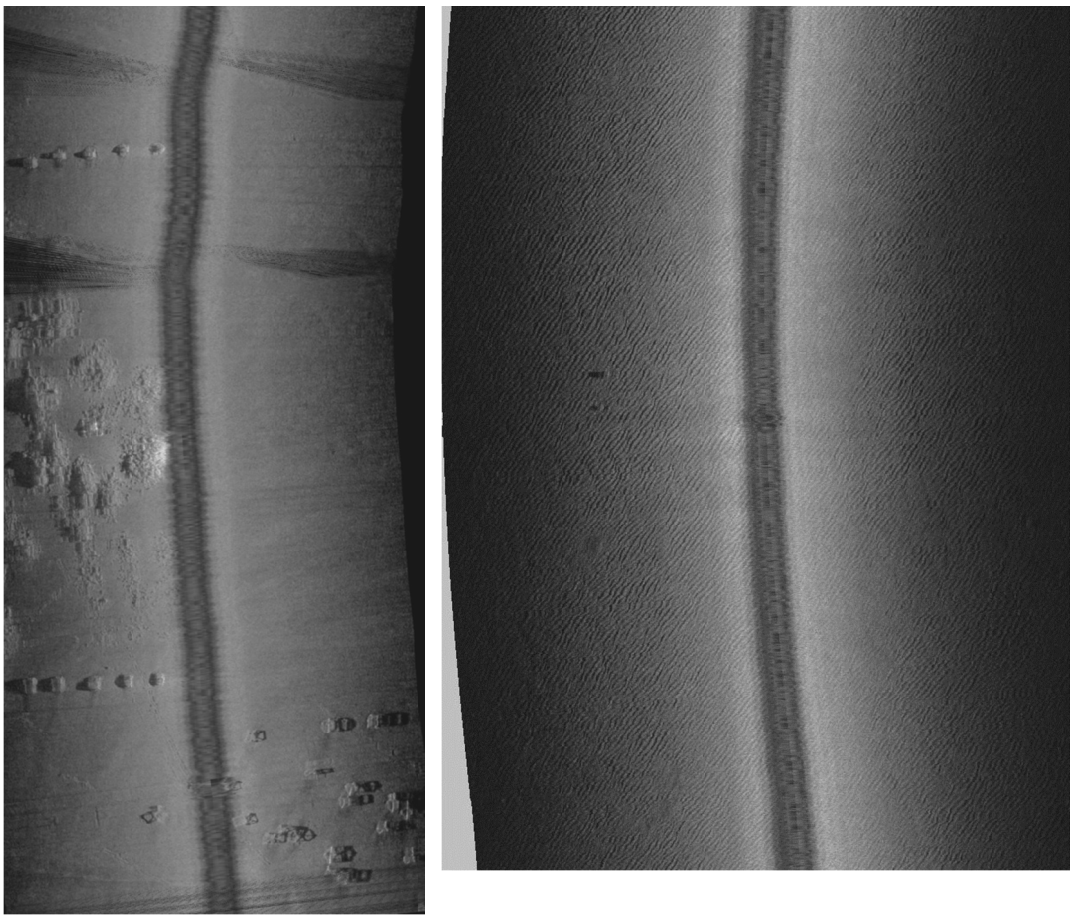


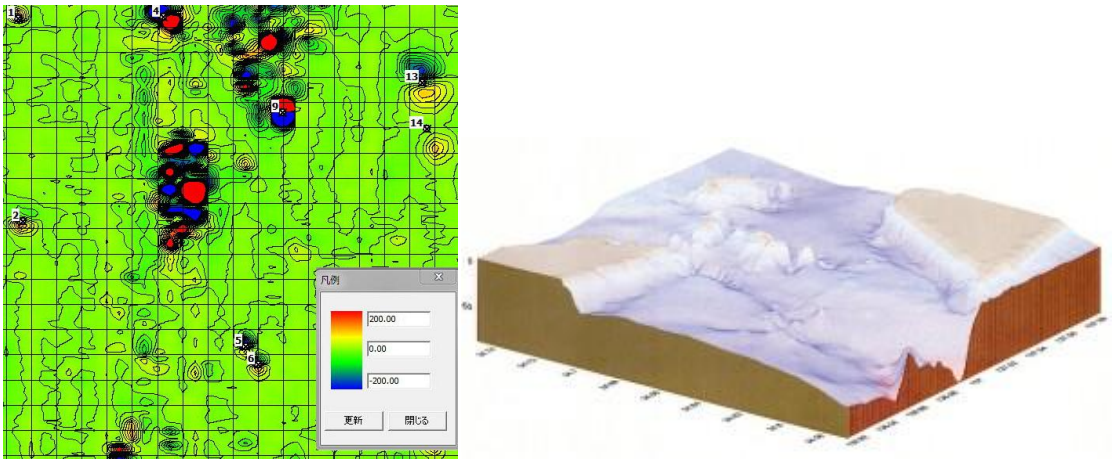
Figure 1-1. Objects present on the seafloor and a seabed sound image.

1.2 Previous studies and problems

Until now, various methods have been considered in the studies of seabed objects detection and seabed sediments classification. Magnetic survey is one of them. This technique measures the magnetic field of the sea, which detects the iron object on the seabed near the surface. In recent years, Sayanagi et al. [1] tried the seabed resource exploration who equipped with a magnetic survey equipment to the AUV. There is a problem in the undersea exploration method based on magnetic survey that, although it is excellent in the ability to detect a submerged object, it is limited to metal objects. In addition, in the detection of the structure of the submarine geology, undersea exploration by the magnetic survey is quite useful in a wide range, but understanding the state of the ground in the local area is difficult. **Figure 1-(a)** shows the result of the undersea exploration using magnetic survey. This figure shows the magnetic force of the band near the bottom of the sea. In the method of magnetic exploration, it is difficult to judge the state of the relief and seabed.

As other method, there is a method for measuring the relief of the seabed by a multi-beam sonar [2, 3]. This is a technique for sounding the acoustic beam in high directivity toward the seafloor, to measure the depth of a large number of points at once. Investigation by the multi-beam sonar requires a variety of equipment such as GNSS, Gyro and Swing sensor. In recent years, this technology has achieved high performance [4-7], but the system is still large scale. **Figure 1-(b)** shows an example of surveying the ocean floor by multi-beam sonar. In either approach, it is necessary to perform post-processing analysis. Furthermore, since the surveying device is a large-scale system, it is not suitable for miniaturization. Hence, the primary purpose of this study is real-time measurement and downsizing of the system.

The author focuses on a relatively small side-scan sonar [8], by which seafloor information (an acoustic image) is easily available.



(a) Magnetic survey **(b) Multi-beam sonar**

Figure 1-2. The seafloor survey resulting in a traditional wide-area exploration: (a) Representing the distribution of magnetic force measurement, (b) represent the relief of the seabed.

1.2.1 Detection of undersea objects

The detection of underwater objects using side-scan sonar imagery includes the exploration of wrecks and mapping of rocks on the seafloor. Most of these investigative strategies make use of the shadows cast by underwater objects [9]. Image generation by side-scan sonar is made possible because objects ordinarily have a higher reflectivity than their background; thus, the return value from object surfaces shows a value higher than that of the background. Sonar images characteristically use this difference to capture images of objects.

Another method used to capture images involves the use of special computer-aided detection/computer-aided classification (CAD/CAC) hardware [10, 11]. This method is effective for decreasing the rate of false detections when a single detection algorithm is used.

More recently, neural networks [12] have been used to detect underwater objects. Sawas et al. [13] proposed a seabed object detection method using Haar-like features. In this method, the identification of countless Haar-like features proved to be problematic because of the lengthy detection times that ensued. Techniques for measuring the magnitude of the submerged object on the basis of the wavelength of the sound waves have also

been reported [14]. In addition, a method to use sonar to estimate the positions of objects on the seabed from positional information collected on the object via Haar-like features was proposed by Aulinas et al. [15]. While both of these proposed methods allow for an easy extraction process, they need to be carried out under the following prerequisites:

- Analyses are limited to simple backgrounds
- The experimental environment is not a real environment
- A landmark needs to be established in a location that is fixed in advance

These assumptions make the results more predictable and also help to clarify the scope of the method. On the other hand, these problems must be addressed in order to make these techniques practical to use in many applications. In the real world, it is not always possible to anticipate the environment under study in advance. Furthermore, in some cases, the seafloor environment may undergo changes from moment to moment during analyses. For these reasons, the conventional methods are unsuitable for real-time detection in the real world.

1.2.2 Automatic identification of seabed sediments

Side-scan sonar is a technique that deploys sound from a remote sensing apparatus to probe the seabed and captures the returning sound wave reflection images of the bottom of the sea. The seabed topography can be visualized through an analysis of the reflected sound waves. The data collected can provide precise surface information on broad areas of the seabed as well as be obtained at a high speed. However, accurate interpretations of acoustic imagery data collected during geological studies of the seabed surface require the knowledge and experience of trained engineers.

A few automated methods for analyzing sound images of the seabed have been developed. Classification of seabed soil from sound images obtained via side-scan sonar was performed by employing texture analysis using a gray level co-occurrence matrix [16, 17], but the results of the analyses were not evaluated using actual seabed soils. Since large calculations were required to determine characteristics such as the frequency and direction

of the reference pixel values, this type of approach converts data to 256-8 gray level by tone correction of the pixel values. During the conversion, it is thought that fine texture information about the seabed is lost, and that this will have an impact on the classification accuracy.

In recent years, Atallah and Smith [18] have proposed a method for analyzing seabed soils using both sound images and depth information collected by depth sensors. However, this method is unsuitable for real-time classifications because the acquisition of depth information is time consuming. For practical purposes, real-time automatic techniques that are developed should follow a simple procedure that results in highly precise and accurate classifications. In addition, field testing (i.e., ground truthing) of research data at 13 points per 1 km^2 range would be desirable.

1.2.3 Configuration and purpose of this study

In this paper, the author proposes a method to perform automatic processing of seabed acoustic data to detect environmental targets in real world situations, which have complex background characteristics. The proposed method is designed to overcome many of the problems associated with conventional methods. The proposed method takes into consideration the following issues.

First, the issue of how to detect high-speed detection of the objects from underwater acoustic imagery is addressed. Shadows of seabed objects can be projected to acoustic images without dependency on shape information. The fact that shadows of seabed objects can be detected from acoustic imagery without using knowledge of specific characteristics such as the size and shape of the target object is highly desirable. This can be accomplished with approaches that use image segmentation. Segmentation is a fundamental and important process in image recognition, and various techniques have been proposed so far [19–24]. However, it can be hard to detect shadows of submarine objects using these methods because the sound images are not clearly divided into “shadow” and “object” data. This causes objects of one region to be divided into many regions, and the locations of objects may be erroneously assigned to different areas. The object detection method discussed in this study focuses its attention on the Haar-like features [25] obtained from

the positional relationship of the object and shadow. Then, to detect an object at high speed, the number of patterns of Haar-like features is minimized.

Second, the issue of how to automatically identify seabed sediments from sound images is addressed. To understand the distribution of offshore resources, seabed sediment information is very important. Distribution maps of seabed sediment are generally developed using a mud vessel technique. Then, once maps have been created, the data can be regarded as being representative of the range of seabed sediments likely to be encountered in the study area. However, if the survey area is wide, the number of observation points increase and the operation becomes enormous, which makes ground truthing difficult to perform. Additionally, conventional methods cannot capture accurate information on changing environments. Hence, a recognition method for seafloor sediment that could be used under real-time environmental conditions would be valuable.

In this study, in consideration of practical uses, a seafloor sediment recognition technique is developed that would be suitable for implementation on current hardware configurations using Higher-order local autocorrelation (HLAC)[26] and that would have other advantageous features such as high classification accuracy and calculations that are easy to perform. Chapter 2 of this paper presents an overview of underwater sound and introduces key principles of sound wave technology. Chapter 3 presents the proposed detection method for objects present on the seafloor, and Chapter 4 describes the details of the automatic sediment classification method. A summary of the findings of this study and proposed new avenues for future research are presented in Chapter 5.

Chapter 2

Underwater acoustics

Human progress requires actions that are in harmony with the abundance of ocean resources. The depletion of ocean resources beyond sustainable levels is now forcing emergency responses and collaboration on a global scale. To use ocean resources wisely, it is imperative that our society researches and documents the quality and quantity of undersea resources in marine environments. Underwater acoustic technology plays a critical role in mapping marine resources, and the use of this technology has expanded widely in recent years. In this chapter, an overview of underwater sound is given and key principles of underwater acoustic technology are explained.

2.1 Efficient use of sound

Unlike air, water has physical properties that make it difficult for light or electromagnetic waves to pass through. However, sound waves propagate readily in water. Therefore, sound waves are used in a wide range of technical fields to detect underwater structures that are difficult to observe with light-based techniques. Various types of non-destructive inspection techniques using ultrasound can be used to map the oceans. In ocean researches, ocean acoustic tomography is a noted example of this type of technology. Techniques that use such “sounds” are collectively referred to as ultrasound electronics. Sound waves are beneficial to use because they have three key properties that are superior to the properties of underwater light. These properties are as follows:

- Sound waves have good permeability in a variety of substances.

- Sound waves have a high spatial resolution due to their short wavelength.
- Sound waves can be used as a signal in many marine applications.

Light does not reach most areas of the ocean, and in these dark areas, the use of acoustic technology is essential for obtaining information. Marine acoustic technology is also capable of obtaining valuable information in illuminated areas in a cost-effective manner. One of the first historical uses for ultrasound technology was to detect nearby hazards to ships, which was prompted by the sinking of the Titanic in 1912. Since that time, ultrasound technology has been deployed in a variety of applications including those used in marine research and in other scientific fields such as medicine as well. The development of marine acoustic technology has expanded rapidly in modern times. In addition to its use in studying the basic physics of acoustic waves, marine acoustic technology has been valuable to other basic and applied fields of marine research such as electronics, physical oceanography, signal processing, and biology.

In the future, marine acoustic technology is expected to play an important role in studies of global warming, marine pollution, and the discovery of natural resources such as methane hydrates. Moreover, such technology may be advantageous to use in ecological monitoring and the sustainable regulation of fisheries.

2.2 Side-scan sonar

2.2.1 Overview

Side-scan sonar [27] has been used to conduct bathymetric surveys and to search for objects on sea and lake beds. It is the technology of choice in many applications, for example, if a plane were to crash into the water, side-scan sonar would likely be deployed at the crash site to detect the wreckage. With side-scan sonar, bottom topography images can be created easily, and this has greatly aided research in geophysics as well as archaeology. **Figure 2-1** shows an image retrieved using side-scan sonar. This image was collected by mounting a tow-fish to an array suspended from a cable on a research vessel. The sounding was produced laterally by a fan-beam and the device was towed at a constant depth and in as constant of a direction as possible through the sea. During towing, a sound pulse

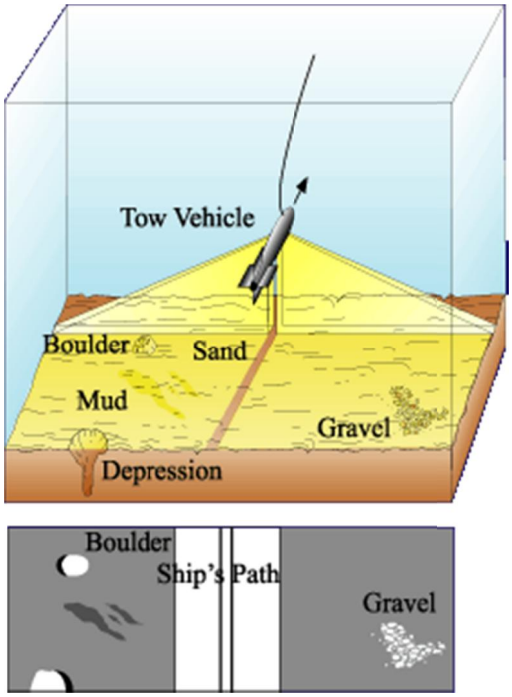


Figure 2-1. Generation of an image by the side-scan sonar.

is emitted at regular intervals from the transducer. The sound waves strike irregularities in the seabed and are then bounced back to the receiver, which records the reflected and scattered waves. Data on the signal strength and the time delay in the received signals can be used to calculate distances. Recording of the received signals is continued until the next pulse is sent. The reflected signal coming back from the seabed by a single pulse is displayed on the recorder as a single line. Bright areas and dark areas of the line represent the intensity of the echo corresponding to the time. Similar to a television screen, meaningful data cannot be obtained by looking only at a single line. Therefore, multiple lines are used. After the process is repeated several times, the lines appearing on the display gradually form a sonogram, which represents a meaningful image. **Figure 2-2** shows a schematic of a side-scan sonar. The schematic shows that not all signals reach the seabed, and in these unreachable areas, a shadow of the sound wave is formed. It is possible to estimate the size and height of seabed protrusions from the length of the shadow. Since differences in the level of scattered waves from mud and sand and direct incoming waves from rock are very large, the receiver has a large dynamic range (i.e., time varied gain or TVG) and change gain time. However, near the center of a protrusion, the transmitted pulse hits just

Table 2-1. Relationship and frequency detection range of a side-scan sonar.

Frequency	Wavelength	Detection distance
100 Hz	15 m	1,000 km-
1 kHz	1.5 m	100 km
10 kHz	15 cm	10 km
25 kHz	6 cm	3 km
50 kHz	3 cm	1 km
100 kHz	1.5 cm	600 m
500 kHz	3 mm	150 m
1 MHz	1.5 mm	50 m

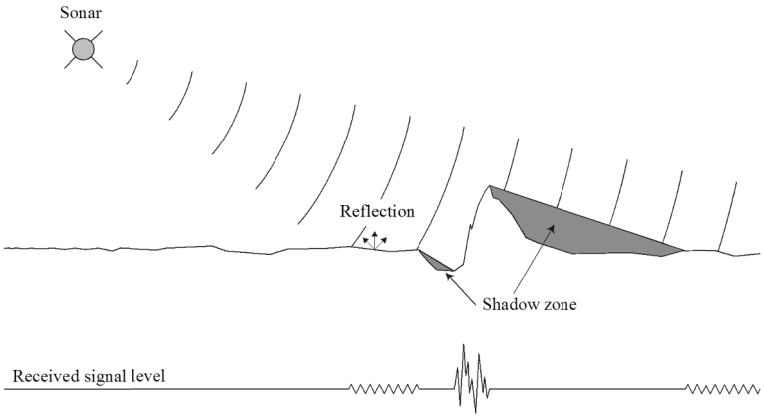


Figure 2-2. Schematic of the side-scan sonar.

from above as shown in **Figure 2-1**. Hence, the received signal is strong. However, the shadow of the sound wave is not well formed; it merely shows a blank column of water. **Table 2-1** shows the frequency-dependent distances that can be detected using side-scan sonar. Acoustic pulses of 100 kHz, which are often used in side-scan sonar systems, are suitable to perform searches at the distances of about 600 m (the resolution can be centi-meters or more).

2.2.2 Configuration of equipment

Figure 2-3 shows a schematic of the equipment configuration for a side-scan sonar system that uses a multi-core cable. A control circuit for adjusting the transmission period of the acoustic pulse is equipped with a controller unit and a display (recorder unit). The trigger pulse and power is sent to the towing cable. The acoustic pulse is transmitted by driving the transducer array by the transmission circuit of the tow-fish in accordance with the trigger pulse. Reflected waves returning from the acoustic pulses are received by the same transducer array and sent to the receiving circuit. The reception signal is amplified through the application of a TVG curve, but in some systems, the TVG adjustments are performed shipboard. Distance to the edge of the scanned range from just below the tow-fish is recorded, and the TVG circuit displays a flat reflection wave. This is used to rapidly increase the gain of the receiving circuit each time a pulse is emitted, and it can be reduced immediately the next time a pulse begins. To estimate the distance to the reflecting object visualized via returning waves, an assumption is made regarding the constant speed of sound in water at the TVG. The gain calculated from the attenuation value corresponding to the distance is changed by the gain nano-wire to the predetermined gain of the receiving circuit. If sand and mud are mixed together at the bottom of the sea, the operator can change the values based on their experience and a correction value curve of the incident complementary angle that forms the basis of the TVG characteristics of the sonar system. It is also possible to improve the sonogram quality further. The received signal is sent to the recorder of the board through conductors in the cable. In a high frequency side-scan sonar, the transducer array can be configured to appropriate horizontal beam widths (1–2 degrees) and vertical beam widths (50–60 degrees). During construction, the array is often arranged as a series of elements. If sacrifices are made in the beam width to suppress side lobes, a weighted electrical unit can be attached to the transmitting circuit. Obliquely arranged elements may also be used in order to reduce side lobes and improve the beam shape [28].

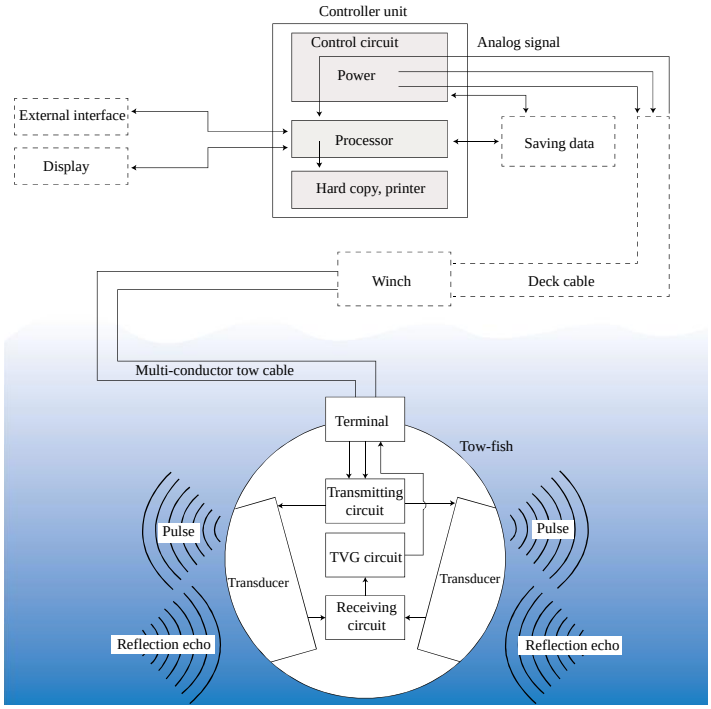


Figure 2-3. Equipment configuration of a side-scan sonar.

2.2.3 Correction of distortion

Instability of the tow-fish

If the tow-fish is not towed stably, the data obtained may be distorted. Instability of the tow-fish is most often caused by the heaving and pitching motions of the research vessel. Distortions caused by low frequency heave from sea swells are the most serious, and objects with a linear shape may appear bent. Sometimes, such distortion is possible to correct by using post-processing techniques that can be implemented with image processing software. However, the fundamental solution is to tow a stable tow-fish. Hence, investigations should be carried out during meteorologically favorable conditions.

Correction by speed

Side-scan sonar systems include a feature called speed correction for changing the chart speed vertical length and width of the drawing to equal the measured speed. Early sonar data were corrected by plotting the speed at different towing speeds and changing the data accordingly. In systems using computers, it has become possible to correct for speed variations in real-time. If the research vessel cannot be equipped with an acoustic

navigation system, a system for measuring and correcting for the speed may be employed using the ship's impeller, which can be effective in calm water. However, if coastal currents are large, speed corrections may need to be based on navigational instruments or perhaps input into the sonar system manually using the keyboard after calculations are performed.

Correction of slant range

Degradation in the accuracy of the distance sonar can be caused by slope distance distortion. **Figure 2-4** shows how slope distance data are distorted by the height of the tow-fish and the timing of the signal return. Specifically, the distance (x) between the targets A and B is the same as the distance between the targets C and D, but the distances can appear different if the slope is not taken into account. Because a tow-fish exists above both A and B, the distance to both targets is approximately equal to the distance to the tow-fish. As a result, the round trip time of the sonar beam is practically the same. The Side-scan sonar displays the data based on the round trip time of the beam, and thus, the two targets are displayed closer than they actually are. For points farther away from the tow-fish, the actual distances are represented more accurately by the sonar data. Even with corrected boat speeds, an operator cannot obtain submarine images by a ratio of 1:1 from the slope distance data. In these cases, the maximum distortion occurs at a location close to the center of the record. For various surveys, a 1:1 linear representation of sonar data is desirable. It is possible to obtain accurate data by correcting the slope distance for each pixel for the entire recording with a personal computer. Corrections to remove the water column from the recording can be performed with high accuracy on the basis of the height data of the tow-fish.

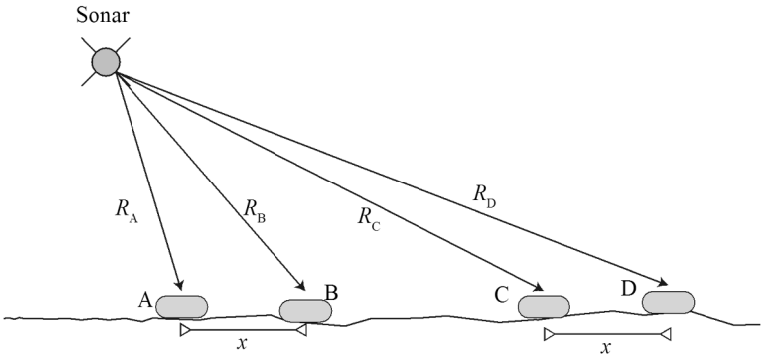


Figure 2-4. Schematic of the distortion caused by the slope distance.

2.2.4 Resolution

Sonar range resolution is the minimum distance that can be used to distinguish two objects in a direction perpendicular to the path of the beam. This resolution has a correlative relationship with the terrain display system. Lower limits to detect different objects can be clearly determined by the pulse width. The range resolution is determined by the projection width of the wave (projected seabed area) and it is determined by the pulse width of the sonar. As shown in **Figure 2-5**, if the difference between two objects is less than the length of the pulse, they appear as a single target. With more narrow pulses, more high-density recordings are generated. On the other hand, the lateral resolution is related to the speed of the ship, pulse transmission interval, and horizontal beam width. Both the ship speed and towing speed is controlled by an operator. However, it is not possible to control the beam width. When two objects are close to the tow-fish, it can be seen clearly that they are individual targets. When the beam spreads gradually, the irradiation area of the pulse is spread out. Therefore, the sonar data obtained far from the tow-fish appear to be from a single target. **Figure 2-6** shows an exaggeration of this effect.

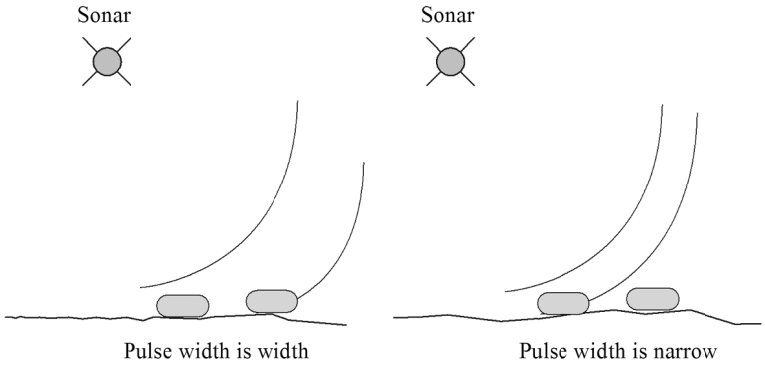


Figure 2-5. Resolution due to the distance of the object and a sonar; (a) wide pulse width, (b) narrow pulse width.

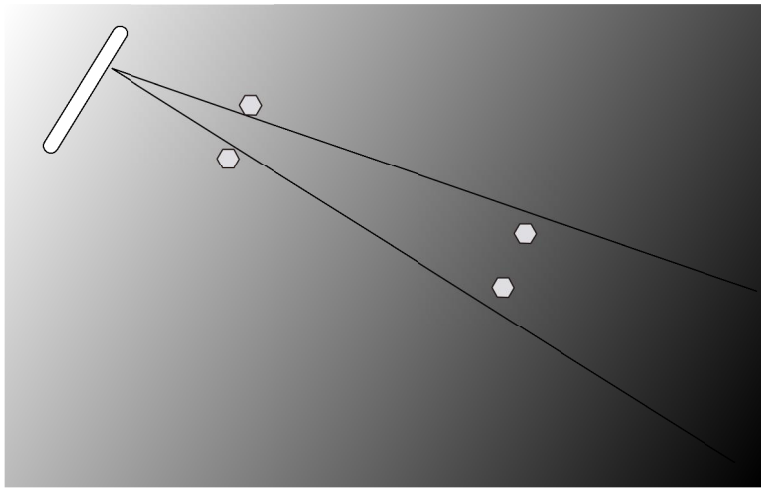


Figure 2-6. Resolution in the horizontal direction and relationship between the distance of the object and the beam width.

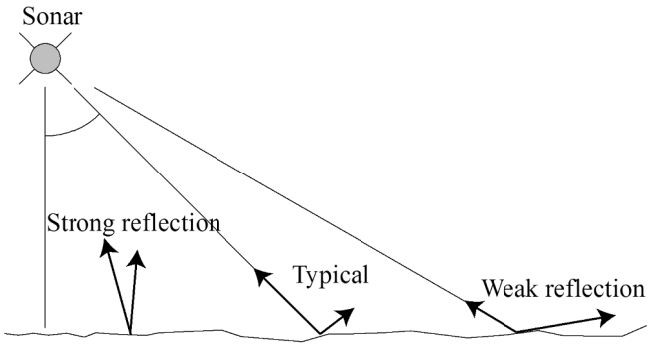


Figure 2-7. Change in the signal intensity due to the angle of incidence.

2.2.5 Judgment of record

Sediment

Reflectance acoustics are determined by the quality of the bottom region of the seabed that is scanned by the sonar. Reflectance of objects is higher than mud, sand, gravel, and rock, which often appear darker in the sonar record. Furthermore, the physical shape of objects can have a significant effect on the intensity of the backscattered waves and reflectivity. Bathymetry is also a factor in determining the energy intensity of the reflected sonar beam. As shown in **Figure 2-7**, the signal intensity changes with the angle of incidence, and this can be influenced by topography. When operators understand how a variety of topographical features can affect sonar records, it becomes easier for them to interpret the results. The impacts that these features have on the sonar record are almost always consistent: Therefore, the impacts are predictable. Skilled operators are often able to easily interpret the sonar record to determine the structure and composition of the seabed.

Shade

In order to produce information in the height direction for sonar records, information on the shading that occurs in the vicinity of objects rising from the seabed is important. The exact interpretation of side-scan sonar records is done by relying on the position of the shadow, the shape, and the intensity. Since the incident angle of the acoustic pulse is reduced for objects standing up on the seafloor, such objects reflect more energy than the surrounding environment as shown in **Figure 2-2**. However, shadows behind objects, in acoustical terms, look bright on the display. These regions are called shadow-zones.

However, some reflected white parts on the sonar record are not just shadows. Sound recorded images without correction, as shown in **Figure 2-1**, assume that reflectors such as fish do not exist around and under the transducer, yet they do. Typically, the energy does not come back through the water column in a manner that displays as pure white in the record, but such energy looks very bright. Semi-transparent objects may produce slight shadows during seabed recordings of many targets. Objects such as rocks and shipwrecks project shadows clearly. In other words, the intensity of the shading contains information about the composition of the objects that are creating the shadows. Causes for the bright parts on the sonar record are as follows:

- Areas of shadow that were blocked by an object
- Locations of the terrain that reduce the backscatter of the sonar beam
- Non-terrain related locations of reduced backscatter

Figure 2-7 shows the change of the signal intensity caused by the angle of incidence. In **Figure 2-7**, the angle of incidence of the acoustic pulse is large where the sediment is sloped downward. Therefore, the energy of the reflection is also reduced. In flat places, reflected energy is greater than that from areas of depressed ground. This will be reflected in the record as dark results. Because of this “incident angle,”, if shading is behind objects and places of high reflectance, the place appears to stick out from the surroundings. If the propagation path from the sonar is relatively straight, as shown in **Figure 2-8**, a triangle with vertices can be drawn along the locations of the tow-fish, the seabed directly below, and the end of the shadow from the target. At the bottom of this triangle by the target, a line can be drawn that intersects the hypotenuse. Similar triangles can be generated from this geometry, and the height of the target can be calculated by use of a simple formula. The height of the targeted object H_F is equal to one divided by the distance R to the end of the shadow product of height H_T from the seabed and the length of the acoustic shadow L_S. The height of the target object is defined by the following equation:

$$H_T = \frac{L_S \times H_F}{R} \tag{2.1}$$

This calculation is very accurate during normal sonar operations. If there is a large sound velocity difference in the propagation path, the sound ray is not necessarily linear. When there is an undulating seabed that is not horizontal, such triangles cannot be used. Therefore, it is not possible in such cases to ensure accurate calculations of the height of objects by the length of their shadows.

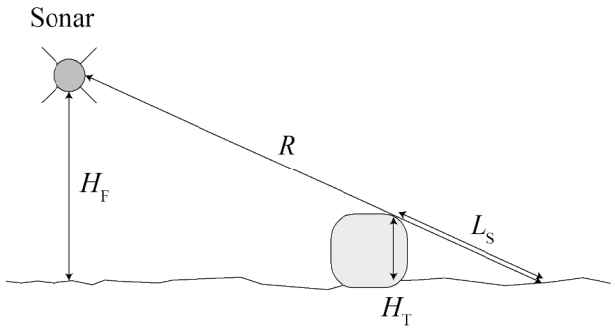


Figure 2-8. Calculation of the height of the seabed object.

2.2.6 Creating a mosaic image

Side-scan sonar mosaic images are created to provide high-resolution images of extensive areas. In order to understand how the vast undersea area looks as a whole, mosaic imaging is a very effective technique to use. However, each acoustic image used to produce the mosaic image needs to be as accurate as possible or the results may not necessarily yield accurate conclusions. To create highly accurate mosaics images, the use of high quality acoustic images is necessary. In order to minimize distortion due to instability of the tow-fish, data collection activities at sea should be performed when the sea is calm. Also, the operator should use a Global Navigation Satellite System (GNSS) to tow the tow-fish at a constant height above the seabed. Navigation errors should be minimized to the greatest extent possible (update rate of about 1 second). For this, it is desirable to maintain a stable boat speed and to proceed in a straight predetermined line. Mosaic image generation requires a navigation display for the operator, and it is necessary

to maintain careful surveillance records of the distance to the tow-fish from the antenna during navigation. The line spacing of the sound images should be set to overlap by 25–50%. Additional survey lines used to create the mosaic image of the submarine environment should use parameters that closely match the conditions used in the initial survey line (i.e., similar vessel speeds and alignments of the directions). By combining small sound imaging datasets, it is possible to extract relevant features of the seafloor from the recordings and determine the actual size and location of objects. Image processing software that comes with the sonar system can be used to calibrate the individual mosaic datasets by using terrain features that exist across the whole image.

In order to determine the gain setting that will best create the final mosaic state, it is ideal to carry out a simple pre-test navigation study of exploration target area. Bright sediments may appear to be pure white in the acoustic images, and it may be possible to improve such images via post-processing. However, if the records used for the mosaic are too dark, there is a risk that results will not be able to be displayed in detail. When the survey begins, the operator should always assume that the acquired data would be used to form a mosaic. Therefore, the setting of the cruising line is an important task to undertake. If the survey ship moves out of the line set, the operator may seek to make changes to the route very slowly. If the sound image acquired is in good condition, the creation of a mosaic may be possible to do shipboard by using the program package installed on the sonar system.

By using computer software, it is possible to construct a three-dimensional mosaic visualization of the bathymetry and target objects. These three-dimensional displays can be used to easily recognize small targets much faster than is possible with conventional sonar systems. Three-dimensional mosaic displays are also useful when searching for objects over a large area of the seafloor. In addition, the sounding data that are created from such systems are very useful for geophysical surveys and geological research.

Chapter 3

Detection of underwater objects based on machine learning

3.1 Introduction

Safe marine navigation depends on the detection of underwater objects and their locations. Methods of detecting such objects fall into three categories: those that use optical devices, those that use ultrasound, and those that employ autonomous underwater vehicles (AUVs). Among these methods, sonar installations that use ultrasound in both commercial and military applications are the most reliable underwater devices, particularly in dark or turbid conditions that restrict the transmission of light through water. If visibility is poor underwater, sonar systems can be employed to visualize the underwater environment at a distance from the vessel and to create clear underwater images. Images obtained in this fashion resemble land-based radar images, and the data can be used to visualize the shape and geological features of the seabed. Detection of underwater objects, which is one of the most common applications for sonar systems, is a major challenge in coastal regions, and this issue is being researched extensively in diverse academic fields including anthropology and marine science. Regardless of the field where the technology is applied, existing analytical methods currently require skilled technicians because of difficulties in automatically detecting objects via side-scan sonar images. These difficulties are caused primarily by the noise that is constantly present in the side-scan sonar images and the significant variability in the images due to the environment.

The detection of underwater objects using side-scan sonar imagery includes the exploration of wrecks and mapping of rocks on the seafloor. Most of these investigative strategies make use of shadows cast by underwater objects. **Figure 3-1** shows how an image is created using side-scan sonar. Because objects ordinarily have a higher reflectivity than their background, the return value from the object surface shows a value higher than that of the background. Sonar images characteristically exploit this difference to create images.

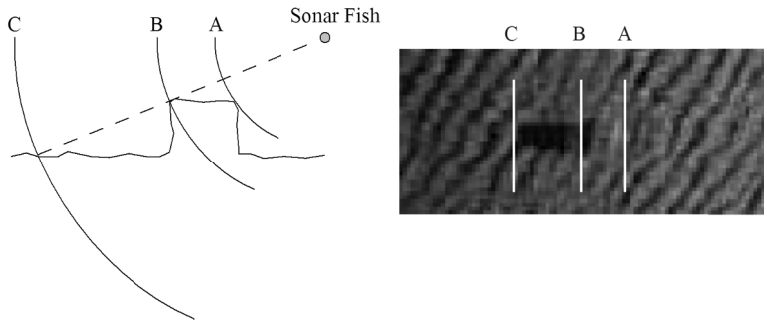


Figure 3-1. Formation of an image by side-scan sonar.

Another method for detection involves the use of special hardware (computer-aided detection/computer-aided classification or CAD/CAC). This method is effective for decreasing the rate of false detections when a single detection algorithm is used. More recently, neural networks have been used to detect underwater objects. Although, this has been problematic because the data generated do not correspond to an object image of undersea objects. On the other hand, Sawas et al. proposed a seabed object detection method using Haar-like features. In this method, there are still drawbacks that need to prepare manually multiple Haar-like features. More recently, a method to estimate the side-scan sonar body positions from the positional information of the objects detected by Haar-like features was proposed by Aulinas et al. However, this approach was not applied in a real environment. For this reason, the conventional method is currently unsuitable for real-time detection in an actual environment.

In this chapter, side-scan sonar imagery was analyzed by the use of a machine learning process with a cascaded AdaBoost and Haar-like features. By minimizing the

number of patterns of Haar-like features, this paper proposes a method for detecting undersea objects faster than previous methods. The effectiveness of the proposed method was verified in a real environmental setting.

Section 3.2 describes the proposed method in further detail and Section 3.3 shows the experimental results obtained for side-scan sonar images from a variety of different seafloor topographies. The conclusions are presented in Section 3.4.

3.2 Detection of underwater objects

3.2.1 Acquisition and preprocessing of underwater images

Side-scan sonar equipment was towed by a research vessel. The equipment beams sonic waves toward the surface of the seabed and the difference in the reflection intensity of the sound waves is displayed as different shadings to visualize the seabed as an image. In the course of sound wave propagation through seawater, acoustic energy is attenuated and converted to other forms of energy. This is called absorption loss [29]. Accordingly, images obtained by a side-scan sonar must be corrected for the attenuation.

To carry out appropriate detections, the obtained images are subjected to preprocessing using an edge-preserving filter that reduces noise on an image while maintaining the edges of an object. Specifically, this technique can be used for calculating the dispersion of gray values in the vicinity of a chosen pixel, smoothing the gray value of the pixel in the direction where the dispersion value is the minimum, and thus reducing noise while retaining edges. **Figure 3-2** shows the effect of the edge retention filter in one-dimensional (1-D) data. **Figure 3-3** shows an example of a pretreated underwater side-scan sonar image. Image (a) in Figure 3-3 shows the original image and image (b) shows the image after the preprocessing.

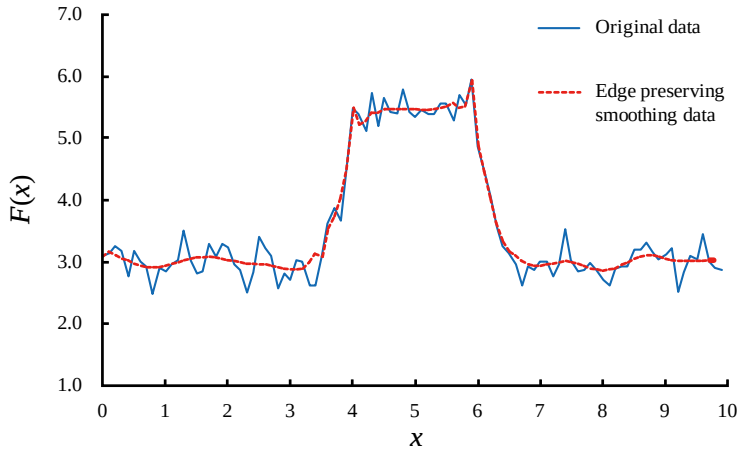
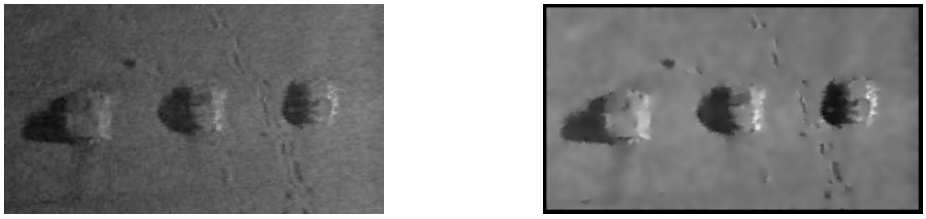


Figure 3-2. Edge preserving smoothing on 1-D edge data.



(a) **(b)**

Figure 3-3. Preprocessing of a side-scan sonar image:
(a) An original image, (b) the processed image.

3.2.2 Haar-like features and machine learning using AdaBoost

In this subsection, using Haar-like features for images and a classifier with a cascade construction for high-speed processing, a method is proposed to detect seabed objects at a high speed by clustering Haar-like features with *k*-means. A seabed object recognition method using Haar-like features and machine learning was first proposed by Sawas et al. [4] and Aulinas et al. The main difference between the proposed method and the previous methods is the specific technique used for selecting Haar-like features.

Extraction of a feature value

Most of underwater objects observed in side-scan sonar images have the following common characteristics.

- The object is paired with its shadow.
- The shadow region is darker than the object.
- The shadow region is darker than the background.
- The object region is lighter than the background.
- A shadow falls only in the direction opposite (horizontal) to the beam emitted from the sonar.

These characteristics are useful to use for object detection. For extraction of these characteristics, with consideration of the beam direction from the sonar, Haar-like features in **Figure 3-4** are used. The value of feature quality is expressed as the difference of the averages of gray values between the white rectangle and the black one.

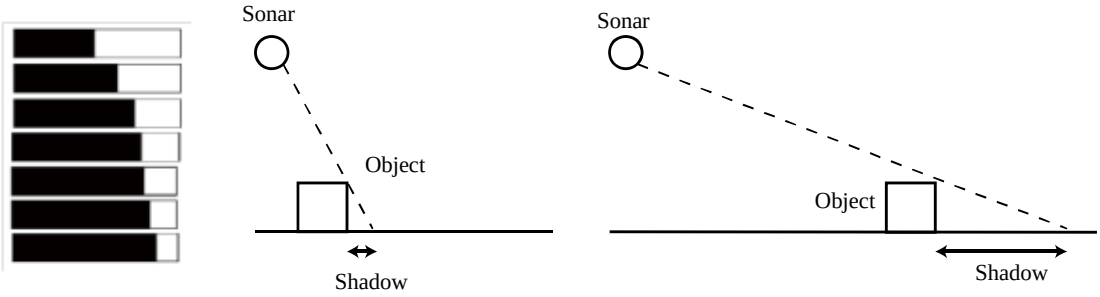


Figure 3-4. The Haar-like feature employed in the proposed method: Schematic formation of the relationship between an object and its shadow.

The weak classifier based on Haar-like features is written as follows:

$$h_t(x) = \begin{cases} +1 & \text{if } p \cdot z(x) > p \cdot q \\ -1 & \text{otherwise} \end{cases} \tag{3.1}$$

Here, z represents a feature value calculated by a single Haar-like feature and q is a threshold value. p is a variable that determines the direction of the inequality sign that compares the feature value z and the threshold value q. It takes the value of either +1 or –1. An integral image [30] $F(i, j)$ defined by **Equation (3.2)** is an image that provides the

sum of gray values from the origin (0, 0) to pixel $p(i, j)$ on image f as shown in **Figure 3-5**.

$$F_p \equiv F(i, j) = \sum_{ii=0}^{i} \sum_{jj=0}^{j} f(ii, jj) \tag{3.2}$$

An integral image is a data structure and algorithm for quickly and efficiently generating the sum of values in a rectangular region in an image.

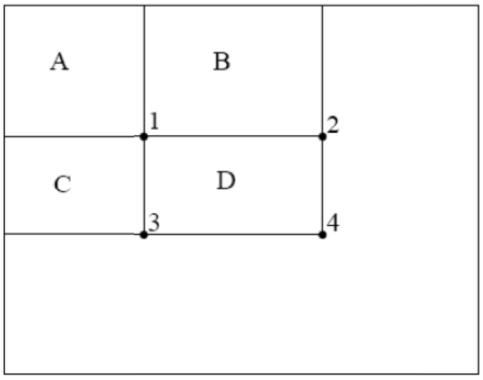


Figure 3-5. Integral image: Using the integral image, the sum of gray values within region *D* can be computed as $F_4 + F_1 - (F_2 + F_3)$.

Determination of Haar-like features

The seabed objects recognition method, which employs machine learning for the Haar-like features and AdaBoost, proposed by Sawas et al. and Aulinas et al. uses various types of Haar-like features that correspond to the shadow extending from the seabed objects. The images of seabed objects have different ratios based on the length of the shadow in the bright and dark parts (**Figure 3-4**). To reduce the number of the Haar-like feature patterns, a histogram of the seabed object images is calculated, and after calculating the resolution of the bright and dark parts, clustering is performed to minimize the number of patterns. The clustering resolution represents the ratio of the bright and dark parts in the Haar-like feature. **Figure 3-6** shows the schematic model for this proposed method.

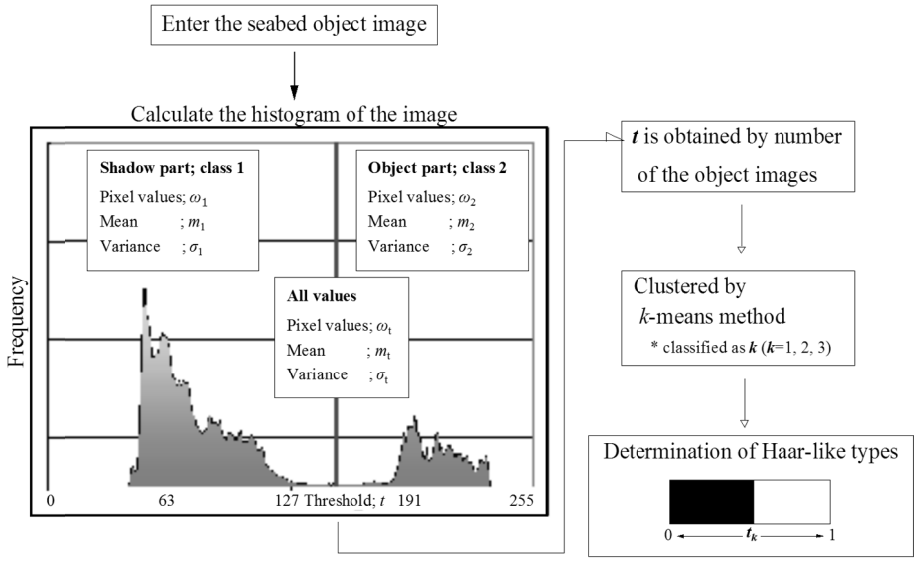


Figure 3-6. Schematic diagram of the proposed method.

In the light and dark parts, the histogram of the seabed object images can be classified by the threshold *t*. A gray value is defined as w_1, which represents the number of pixels in a region smaller than the threshold value, the average is m_1, and the variance is s_1. Similarly, in the region on the other side, the number of pixels is w_2, the average is m_2, and the variance is s_2. Thus, within-class variance s_w^2 can be defined by the following formula:

$$s_w^2 = \frac{w_1 s_1^2 + w_2 s_2^2}{w_1 + w_2} \tag{3.3}$$

The between-class variance is defined as

$$s_b^2 = \frac{w_1 (m_1 - m_t)^2 + w_2 (m_2 - m_t)^2}{w_1 + w_2} \tag{3.4}$$

In this case, when the value s_b^2 / s_w^2 ,the ratio of the within-class variance and between-class variance, is maximized, the threshold value is *t*.

A threshold value *t* is first calculated for all the seabed objects in the training images.

Then, t is clustered by the k-means method. The k-means algorithm is used for clustering the data into k classes that are decided in advance.

The k-means clustering is performed as follows. In the case of clustering to k,

Step1. The k points are chosen from the t points at random and set to the initial value in the center of the cluster. All data points are grouped to the center of the cluster in the shortest distance. Each group is set to the cluster.

Step2. After averaging the data points of each group, it is set to the updated value in the center of the cluster.

Step3. All data points are grouped by separating them into the center of the cluster in the shortest distance. In the case of no change, the k clusters data are out-put and ended. In the case of change, Step 2 is repeated.

The Haar-like features are defined based on the clustered t_k (k = 1, 2, …). **Figure 3-7** shows the Haar-like features of k = 3.

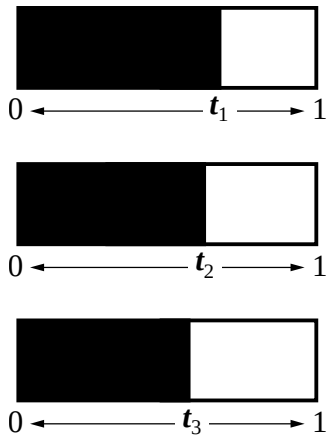


Figure 3-7. Determination of Haar-like features.

AdaBoost

AdaBoost is a learning/discrimination algorithm that is based on the concept of connecting a multiple number of weak classifiers whose individual discrimination capabilities are not very high to create a strong classifier. A strong classifier can be made

by combining several weak classifiers that are weighted according to their degree of importance. Initially, uniform weights are given to all the samples. In the process of learning of weak classifiers, the weight of a sample that was correctly classified by a weak classifier is reduced, whereas the weight of a sample that was not correctly classified is increased. By using the samples with increased weights, the weak classifiers that correctly classified the samples can be chosen in the next stage.

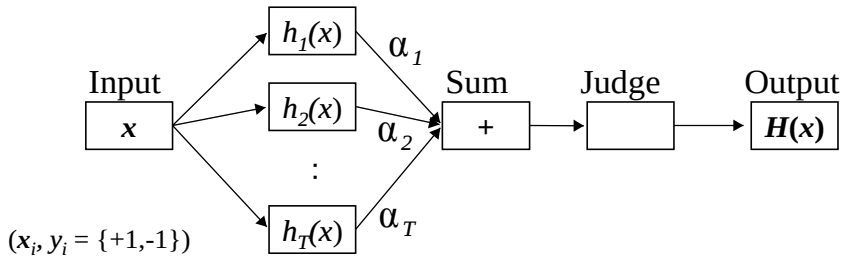
Assuming that a weak classifier is $h_t(x)$ (t = 1, 2, …, T), α_t is expressed as $a_t = \log\frac{1-e_t}{e_t}$ using the error ε_t of the weak classifiers determined during the learning process. A strong classifier is then defined by the following formula:

$$H(x) = sign\left(\sum_{t=1}^{T} a_t h_t(x)\right) \tag{3.5}$$

where

$$sign(w) = \begin{cases} 1 \ \ldots \ w > 0 \\ -1 \ \ldots \ w \le 0 \end{cases} \tag{3.6}$$

When $H(x)$ is +1, the data are judged as seabed objects, and when $H(x)$ is -1, the data are judged as non-seabed objects (**Figure 3-8**).



x : Calculated from the Haar-like features of one.
α : Reliability
$h(x)$: Weak classifiers

Figure 3-8. Classifier obtained by AdaBoost.

Cascade structure

As shown in **Figure 3-9**, a classifier having a cascade structure has strong classifiers ($H_n(\boldsymbol{x})$: n = 1, 2, …, N) connected in a straight line, with N representing the number of the strong classifiers. The feature values or the input data extracted from image data in regions surrounded by a search region (sub-window) is used to identify whether or not an object is contained inside the search region (sub-window). The nth strong classifier $H_n(\boldsymbol{x})$ can be defined by the following formula:

$$H_n(x) = sign\left(\sum_{t=1}^{T_n} a_{n,t} h_{n,t}(x) - q_n\right) \tag{3.7}$$

Here, $h_{n,t}(\boldsymbol{x})$ is the t^{th} weak classifier selected at stage n; T_n is the number of weak classifiers selected at stage n; $\alpha_{n,t}$ is the weight of the weak classifier $h_{n,t}(\boldsymbol{x})$; and θ_n is the threshold value at stage n.

The role of the classifier placed in the first stage of the cascade is to reject simple background regions. To distinguish an object from background regions, only a very small number of weak classifiers are needed. Hence, the majority of images can be rejected with a very low calculation cost and the process can move on to the identification of the window in the next position. The classifiers used in the second stage need an even greater number of weak classifiers to distinguish an object from any complex background that was not rejected in the former stage. Only a window that has passed the classifier H^N on the final stage is judged to be an object. To detect an object, the image must be scanned thoroughly. However, most regions contain no objects. Therefore, since most of the background regions contain patterns that clearly differ from those of an object, it is possible to expedite the process dramatically by adopting this immediate rejection strategy. Let us denote the probability that an image with a correct object passes a single cascade by D_r ($0 < D_r < 1$). Let us also denote the probability that an image with an incorrect object passes it by F_p ($0 < F_p < 1$). Then, the probability that an image with a correct object passes n successive cascades is D_r^n, whereas the probability that an image with an incorrect object passes them is F_p^n. For D_r = 0.999, F_p = 0.5, and n = 40, e.g., D_r^n =

$0.999^{40} \approx 0.96$ and $F_p^n = 0.5^{40} \approx 10^{-13}$. This means that this cascade-type classifier classifies almost all the images containing objects correctly, whereas it rejects almost all the images containing incorrect objects.

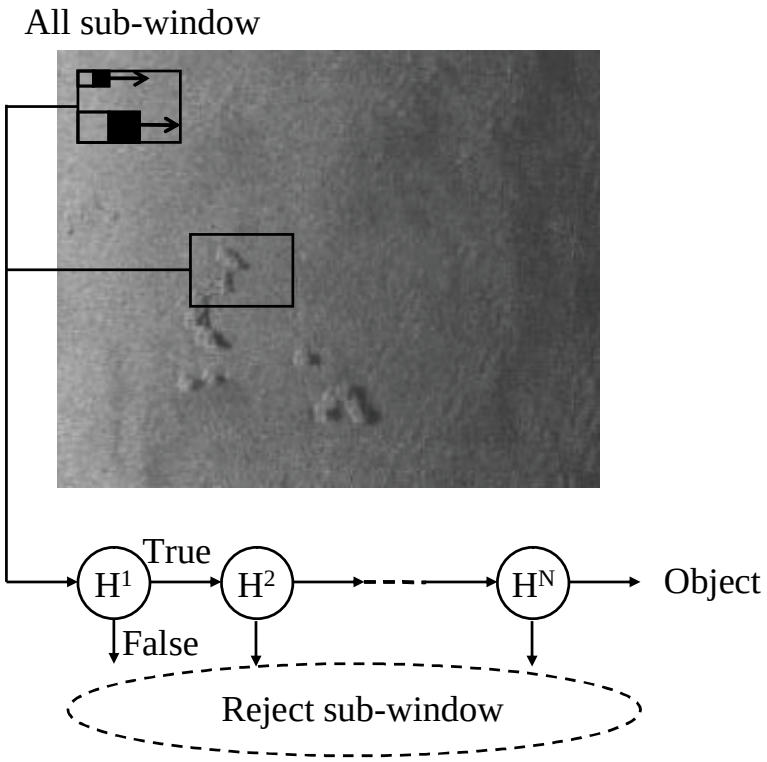


Figure 3-9. Schematic model of the cascade classifier.

3.3 Experimental results and discussion

For the experiments, side-scan sonar images obtained in actual environments are used. Two sites were used for exploration: a muddy seabed and a sandy seabed.

After lowering the sonar device into the water from the stern of the research vessel, an area about 300 meters wide and 2,000 meters long was explored and the images of the seabed were obtained. The water within the scope of the exploration area was about 20 meters deep. The area was inspected by adjusting the towing depth of the sonar to a set interval of 12–14 meters from the sea surface. The total navigation time was approximately 8 hours.

This exploratory activity revealed that the seabed was generally flat. **Figure 3-11(a)** shows the images of the seabed taken using the side-scan sonar.

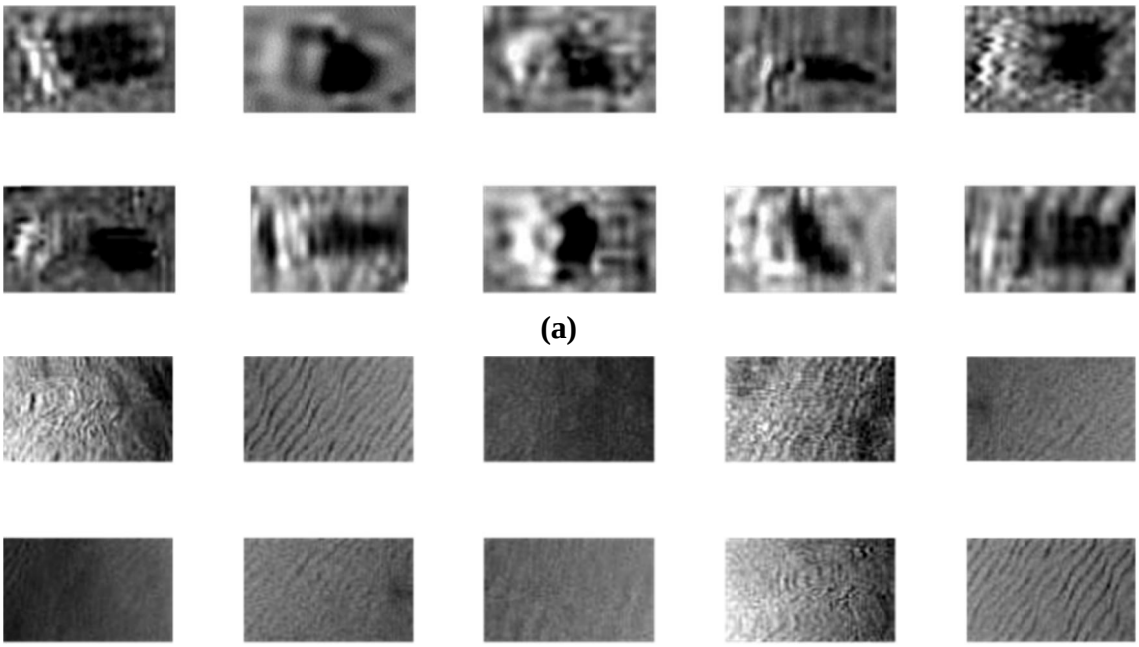
3.3.1 Detection of objects

As a result of the calculation, the number of clusters *k* by the *k*-means method is 3. This compares well with the 4–7 Haar-like features found using traditional research methods, and the number was reduced to 3 with the newly proposed method. **Table 3-2** shows the values determined by *k*-means.

The number of training images needed by cascade classifiers for detecting objects is 4,000 object images (24 × 32 pixels) and 1,000 non-object images. Since it is practically difficult to prepare for 4,000 images of underwater objects, the number of object images were artificially increased by changing the luminosity of some of them. Specifically, the gray values of selected original images were varied within the range of -10% to 10%. **Figure 3-10** shows examples of the object images. Since there is a need to exhaustively detect all seabed objects, whatever their size, The size of the sub-window is varied in 0.1-fold increments varied from 0.8-fold to 1.2-fold. The size of the image is determined for learning as 24 × 32 pixels, and the window size is changed every 0.1-fold. The computer used for the calculation had a 4 GB memory and a 2.66 GHz Intel Core 2 Duo processor. It took about 60–95 milliseconds to detect seabed objects. **Figure 3-11(b)** and **Figure 3-11(c)** show the results of the detection. **Figure 3-12** shows representative examples of detection results in a mud and a sandy bottom area. Detection was favorable for objects both on the muddy seabed and on top of sand-waves (i.e., ripples) on the sandy seabed. **Table 3-3** shows the results of object detection in these two sea areas. Within the first sea area, manual exploration revealed 124 objects existent on the seabed. Out of these objects, 110 were detected correctly using the proposed method. Therefore, the seabed object detection rate was 88.7%. Within the second sea area, 13 objects were detected correctly out of a total of 17 objects; the accuracy rate was 76.5%. These high accuracy rates suggest that the proposed method is effective for detecting seabed objects.

Table 3-1. Specifications of sonar.

Basic Performance Parameters	
Frequency	340 kHz
Horizontal beam-width	0.9 deg
Vertical beam-width	60 deg
Range resolution	10 cm
Max operating depth	100 m
Max cable length	200+ m
Materials	Stainless steel, PVC and Polyurethane



(a)

(b)

Figure 3-10. Snapshot of a seabed: Examples of (a) object images, (b) non-object images.

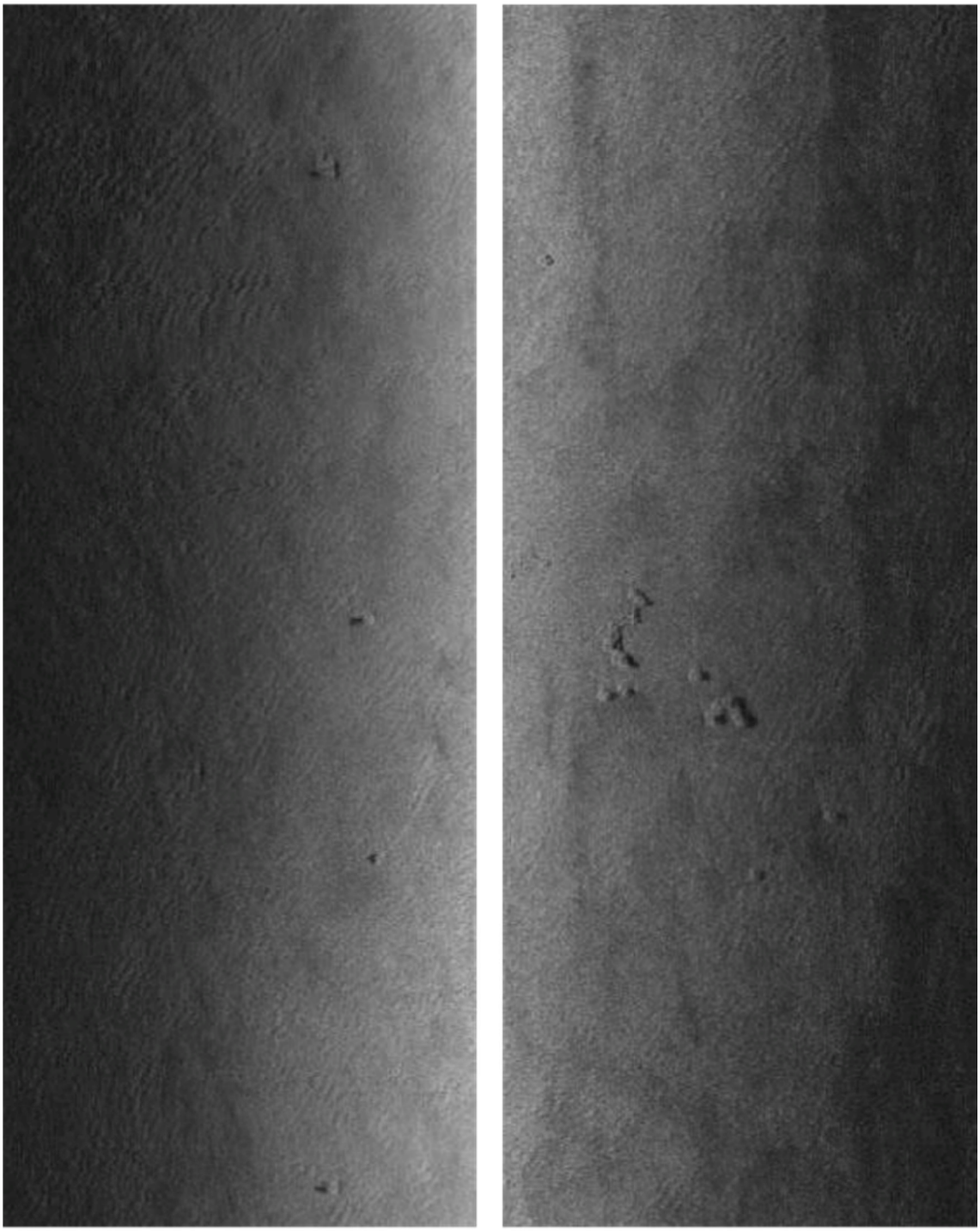


Figure 3-11(a). Seabed objects on a side-scan sonar image and the result of the detection: *Original image*.

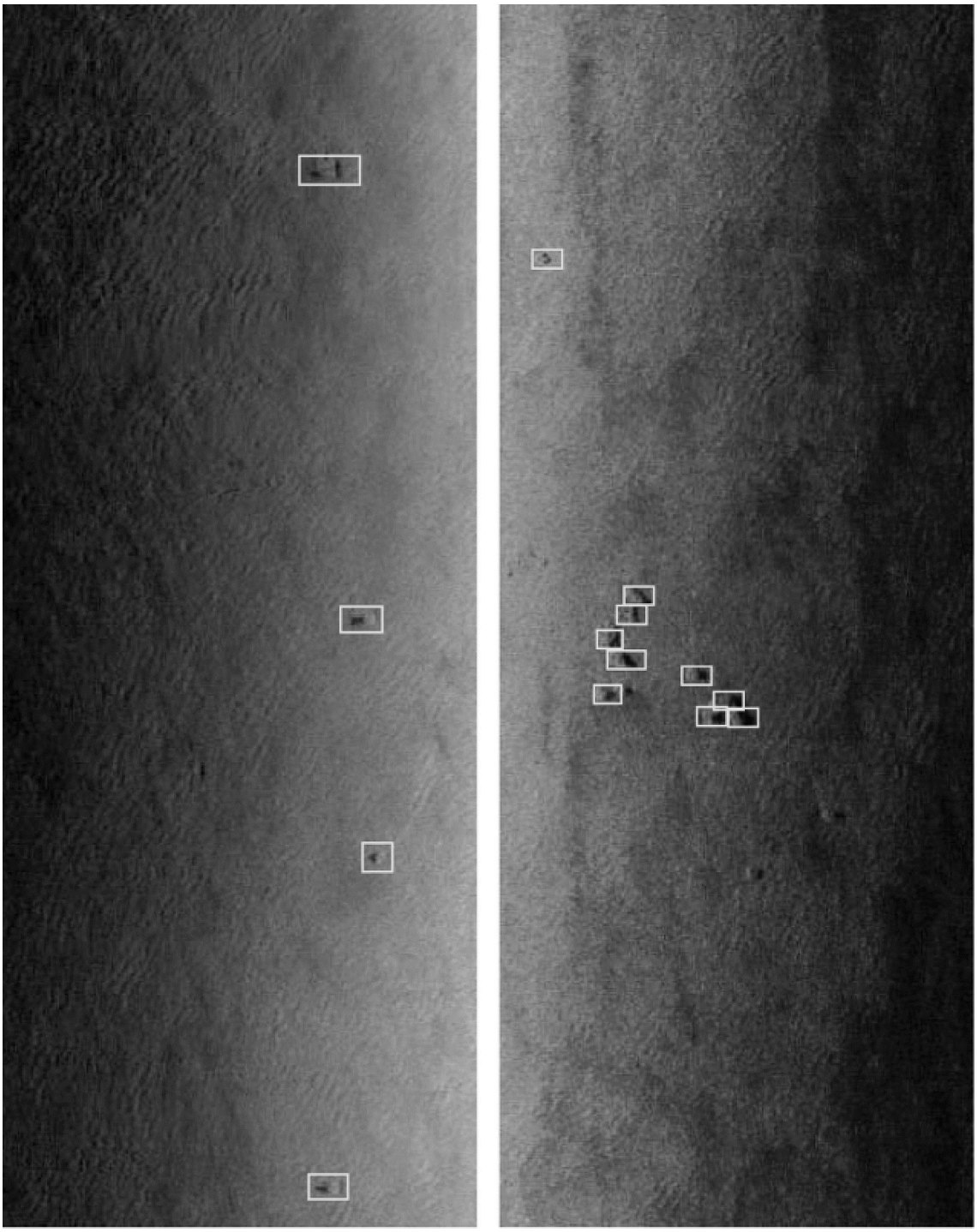


Figure 3-11(b). Seabed objects on a side-scan sonar image and the result of the detection: *Detected objects*.

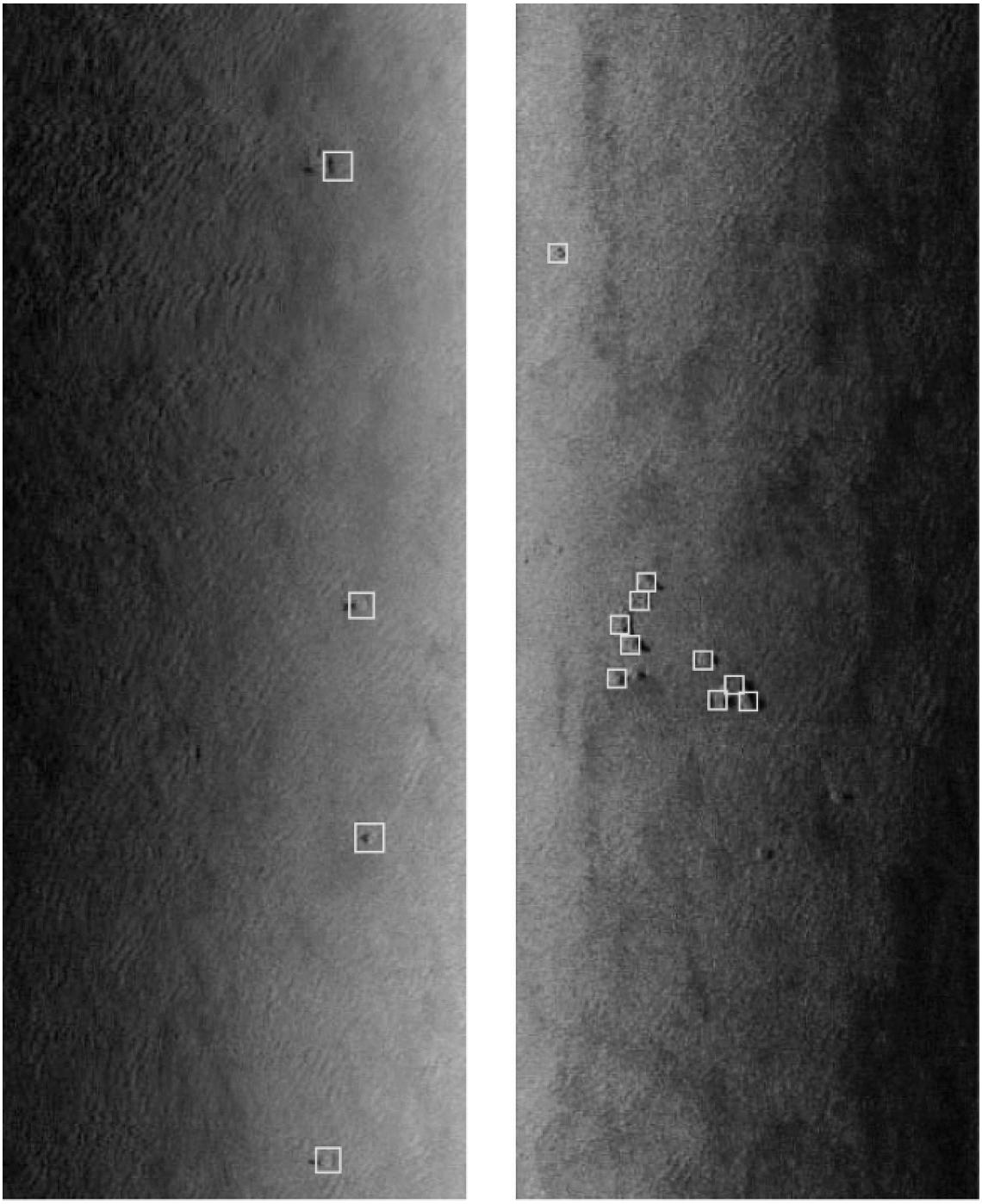


Figure 3-11(c). Seabed objects on a side-scan sonar image and the result of the detection: *Rectangles indicating only the objects*.

Table 3-2. Determined value of the threshold (t_k).

Threshold	Value
t_1	0.6857
t_2	0.6357
t_3	0.5714

The shortage of the shadow length of the seabed objects caused the length of the shadow of the undersea object to be insufficient. Haar-like features are characterized by the amount of difference between white areas and black areas. Since the object image contains a gray area, differences of gray values do not come out well in the prepared Haar-like pattern. **Figure 3-13** shows an example of a failed detection and a successful detection. In addition, **Table 3-4** shows the comparison of detection times by the proposed method and the method described by Sawas et al. Detection times were faster when using the proposed method (65–90 milliseconds) than when using the method of Sawas et al.(141–205 milliseconds). This result shows that one can realize shorter detection time without compromising detection capability. Therefore, the proposed method is effective for fast detection of undersea objects, and it is intended to lead to real-time processing of seabed imagery.

(a) **(b)**

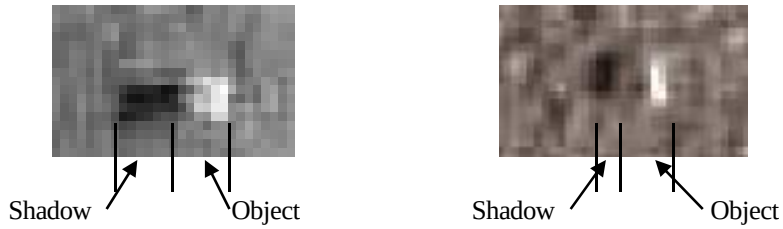
Figure 3-12. Detection results for seabed objects on the mud flat board and the sandy area board: (a) Seabed object on the mud, (b) seabed object on the sandy bottom.

Table 3-3. Performance of the detector.

Endpoints	Sea area 01	Sea area 02
Total objects	124	17
Detected true objects	110	13
False objects	14	4
Accuracy rate	88.7%	76.5%
Average	**82.6%**	

Table 3-4. Comparison of the detection time of objects.

Method	Detection time (ms)
J. Sawas, et. al.	141-205
Proposed method	**65-90**



(a) **(b)**

Figure 3-13. Examples of successful and failed detections: (a) Successful, (b) failed.

3.3.2 Accuracy of the proposed method

To use the proposed method in an actual environment, positional accuracy will come into question. Since the research vessel that tows the sonar device is equipped with a Global Navigation Satellite System (GNSS), it can easily obtain information on the location. The vessel's location information is used as the basis for comparing the position of known objects with the position of the newly detected seabed objects and investigated its degree of accuracy. When performing this comparison, seabed objects were visually confirmed from the images. Installation position information was acquired from the position information of the object at the center of the sound images. The objects on the seabed consisted of a reef (a cube of about 3 meters). As the rectangle that surrounds the

detected object includes its shadow, a rectangle that surrounds only the object is needed. Note that the characteristics of a beam emitted from the sonar device are reflected only at the side of an object, and that its width roughly coincides with the vertical direction of the rectangular region. Therefore, a single point on the lateral side of a seabed object found inside a rectangular region is used as a starting point. Then, the length of the rectangle's horizontal direction is matched to the length of the rectangle's vertical direction to create a square. However, since this square do not necessarily encompass the contour of the seabed objects accurately, The edges of the object images contained inside the square is highlighted, and then further re-shape the square to enclose the edges of the contour. The result is an accurate rectangle surrounding the object. The results are shown in **Figure 3-11(c)**. This technique detects the same objects in adjacent survey lines. However, the objects are not simultaneously detected. When outputting the detection result separately on both sides, the detection coordinate value is determined to have an error range within ± 1.5 meters. Then, the merge process is performed. **Table 3-5** shows the results of the comparison of accuracy. The average error between the positions of the detected objects and their known positions was 0.501 meters, and the maximum error was 0.615 meters. The GNSS, which is commonly used for seabed exploration, show a positional accuracy (catalog value) of 1–2 meters in the horizontal direction in Sea Area 01. In Sea Area 02, the average error was 0.398 meters, and the maximum error was 0.499 meters. The positional error of the proposed method can thus be regarded as satisfactory. However, it does not take into account distortions caused by waves. Therefore, investigations in stormy weather are not recommended. **Figure 3-14** shows the location of the detected objects and the trajectory of the research vessel. The blue line shows the trajectory, and the red squares indicate the position of the seabed objects detected by the proposed method. The numbers on the *x* and *y* axes represent the World Geodetic System (JGD2000; I) scale, i.e., the geodetic coordinates that divide the Japanese archipelago into 19 datum.

Table 3-5. Errors between the positions of the detected objects and their installed positions (units are meters).

Sea Area 01		Sea Area 02	
Number	Error value	Number	Error value
1	0.259	1	0.336
:		:	
:		:	
:		:	
109	0.188	12	0.426
110	0.547	13	0.298
Average	0.501	Average	0.398
Max error	0.615	Max error	0.499

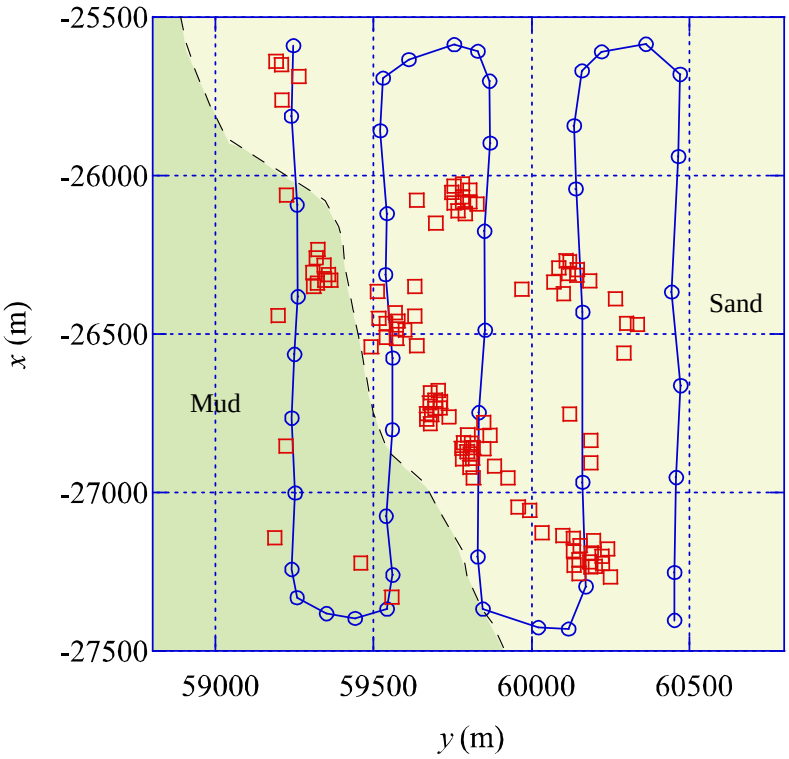


Figure 3-14. Location of the detected objects on the seabed and the trajectory of a research vessel. The line shows the trajectory, whereas the dot indicates the objects. The *x* and *y* axes show the Japan Geodetic Datum 2000.

3.4 Conclusion

In this chapter, the author proposed a method of detecting seabed objects based on machine learning that uses Haar-like features and cascaded AdaBoost to analyze side-scan sonar images. Furthermore, a method for performing object detection using Haar-like patterns of a small number of features via *k*-means clustering was proposed. The effectiveness of the proposed method was investigated by applying this technique to the detection of real objects on seabeds consisting of both sandy and muddy ground. The results show that the new method (i.e., use of Haar-like features characterized by *k*-means clustering) was faster at detecting objects than traditional methods. The new method opens up the possibility for processing seabed object images in real-time. According to past references, seabed object detection results for different types of the seabed should be confirmed with known data. This is also important for the results generated by the proposed method. Because Haar-like features of seabed shape objects contain the shadows of seabed images, the technique was conducive to detecting objects underwater. However, it may be necessary to consider another approach for seabed objects that do not form shadows (e.g., plate-like objects). For the proposed method using *k*-means, the classification class number *k* was determined manually from the information present in the sonar images. In the future, the clustering method should be modified to perform automated unsupervised classification of data from the group classification classes (e.g., self-organizing maps or SOM) [31]. In addition, there may be a need to use Principal Component Analysis (PCA) or Independent Component Analysis (ICA), either separately or together [32], to examine pretreatments for detecting seabed objects that do not depend on the luminosity and contrast of the original image.

Chapter 4

Automatic classification of seabed sediments using HLAC

4.1 Introduction

The ocean accounts for approximately 70% of the area on the earth. Ocean resources are used for fish farming, land reclamation, and a variety of other purposes. Seabed resources such as oil, natural gas methane hydrates, and manganese nodules are still largely unexploited on the bottom of the sea. Nowadays, seabed resources are becoming more accessible to a large extent because of recent, rapid technological developments. Information on the quality of the seabed surface is important to understand the distribution of seabed resources and to minimize the environmental impacts of extractive practices. Currently, distribution maps of seabed resources are generally made using a bottom sampler, which samples the surface of the seabed area of interest point by point. These point-wise data are regarded as being representative of the seabed area under investigation. Obviously, this type of methodology gives less exact information on the surface of extensive seabed areas even with a large number of sampling points. To cover wide geographic areas, acoustic technology is often employed. Acoustic technology can be especially valuable to use in areas where light does not reach the sea bottom because sound waves have a very good permeability through media irrespective of whether the media is a liquid, gas, or solid.

A side-scan sonar can be used to create cost-effective sound wave reflection images of the sea bottom during short-term investigations. **Figure 4-1** summarizes how images

are generated with side-scan sonar. The side-scan sonar equipment irradiates a sound wave to the seabed and the seabed topography can be visualized in an image that is formed by analyzing the reflected sound waves. This technique can be used to provide precise surface information over broad areas of the seabed at a high speed. However, geological studies of seabed surfaces that employ acoustic imagery largely depend on knowledgeable and experienced engineers, which are not always available. A few automated methods for analyzing acoustic images of the seabed exist. Texture analysis was performed by Yamamoto et al. [17] to classify seabed soil from sound images obtained with side-scan sonar, but the results of the analysis were not evaluated using actual seabed soils. In recent years, Atallah and Smith [18] proposed a method of analyzing seabed soils using both sound images and depth information provided by a depth sensor. However, this method is unsuitable for performing classification in real-time because of the delays that occur during the acquisition of depth information. For practical use, any newly developed techniques should constitute a simple procedure and yield highly precise classification.

This chapter proposes an automatic method for classifying seabed soils from sound images based on HLAC (higher-order local auto-correlation) features [26]. HLAC is a basic as well as general image feature processing method that directs attention to the co-occurrence nature of texture patterns. It has been employed for the recognition of faces [33,34], letters [35], gestures [36], and other objects of interest and it has been used in diverse applications such as medical image analysis and the evaluation of remote sensing data. This paper shows the effectiveness of using the proposed method in oceanographic research by applying it to real seabed images. The author also compared the proposed method to other methods such as the one that uses a gray level co-occurrence matrix developed by Peckinpaugh and a Local Binary Pattern (LBP) technique [37] that is used widely in texture analysis.

In the following text, Section 4-2 explains the proposed method in further detail and Section 4-3 shows the experimental results that were obtained using side-scan sonar images of various types of seafloor topography. The discussion is presented in Section 4-4 and Section 4-5 concludes the paper with recommended directions for future work.

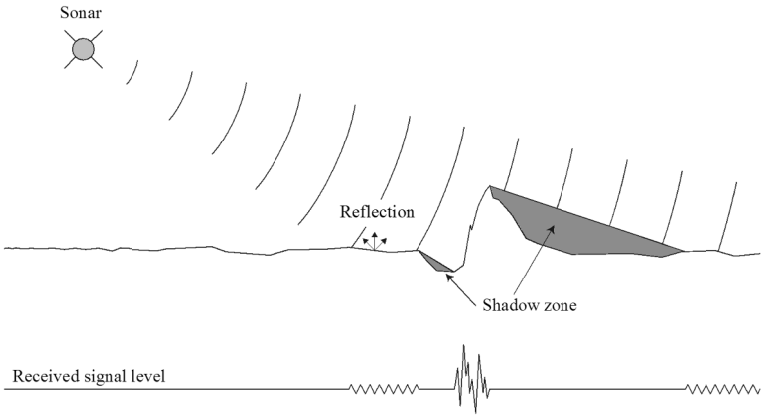


Figure 4-1. Schematic of a side-scan sonar.

4.2 Classification of the seabed sediments

Sonic waves are beamed toward the surface of the seabed and the difference in the reflection intensity of the sound waves is displayed as different shadings to visualize the seabed with images. In the course of sound wave propagation through seawater, acoustic energy is attenuated and converted into other forms of energy. This is called absorption loss. The amount of attenuation of sound waves *Np* (dB) can be calculated by the following formula:

$$\begin{aligned} Np &= 20\log r + ar \\ a &= 0.22f + 0.000175f^2 \end{aligned} \tag{4.1}$$

where r is the distance from the sound source to the object (km) and α is a coefficient that depends on the frequency of the sound wave f.

4.2.1 Flow of the proposed method

The proposed method creates a feature space from training data that are derived from sound images of the seabed. An unknown sound image of the seabed is classified into one of several classes representing different seabed sediments using the feature space. The HLAC feature, which extracts co-occurrence data, is employed for describing the texture of the seabed sediments. The subspace method is employed to create the feature space and

to perform the classification. A schematic diagram of the classification process is given in **Figure 4-2**.

The HLAC feature vectors are computed, in advance, from training image data of all seabed sediments. Then, a Principal Component Analysis (PCA) is applied to the vectors to compress the dimension and to form a subspace. Given an unknown sound image, the HLAC feature vector is calculated from every scanned window on the image and the distance of the vector to each subspace is evaluated for classification. The vector is used to classify unknown sound images into the class where the distance is the minimum.

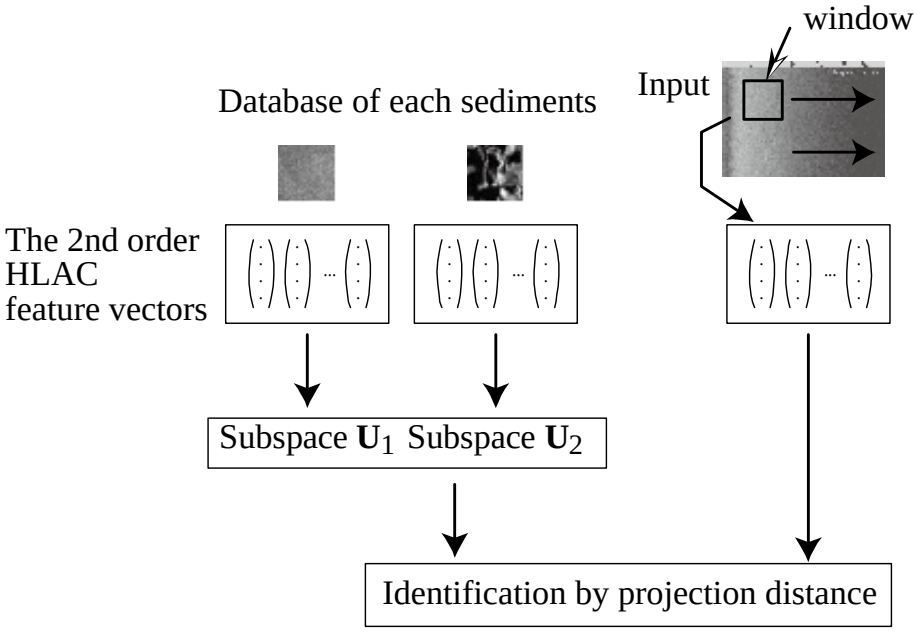


Figure 4-2. Flowchart of classification.

4.2.2 Extraction of the features

The proposed method uses the HLAC features to extract texture features. If the gray value of the target image at the point r is denoted by $f(r)$, the Nth autocorrelation function can be defined by the following equation:

$$x(a_1, a_2, \mathbf{L}, a_N) = \int f(r) f(r + a_1) \mathbf{L}\ f(r + a_N) dr \qquad (4.2)$$

The HLAC feature is a basic image feature based on this function. Since the local correlation between nearby pixels is more important in image data, we set $N = 2$ and considers the domain of local 3×3 pixels around the reference point r, that is, correlation of up to three points in the local domain is taken into account. Then, the HLAC feature performs the calculation using 25 mask patterns of 3×3 pixels (1 zero-dimensional, 4 one-dimensional, and 20 two-dimensional patterns) as shown in **Figure 4-3**(a). As one mask pattern produces a single value from a window, a feature vector is defined by a 25-dimensional vector. The 0th order, the 1st order, and the 2nd order HLAC features are calculated by the following equations, respectively:

$$\begin{aligned} x(a_0) &= \int f(r)dr \\ x(a_1) &= \int f(r)f(r+a_1)dr \\ x(a_2) &= \int f(r)f(r+a_1)f(r+a_2)dr \end{aligned} \tag{4.3}$$

The calculation of the HLAC feature proceeds by multiplying the gray values corresponding to the dark pixels of each mask pattern and then summing them in the window of interest.

In the present research, masks of the size 3×3, 5×5, 7×7, and 9×9 pixels are also employed as shown in Figure 4-3(b) in order to extract texture features of larger domains. Note that the mask of each size has 25 patterns similar to the mask of the 3×3 example. Therefore, the dimension of the actual feature vector is 100. These mask patterns with larger sizes can extract lower frequency features compared to the 3×3 mask, as they can calculate the correlation among two or three mutually distant pixels. This is equivalent to applying a 3×3 mask pattern to a lower resolution image made by reducing pixels of the original image in a pyramidal way. The 3×3 mask pattern extracts the high frequency component of the image. By considering variation in the mask sizes, low frequency features as well as high frequency features can be extracted, which may be advantageous to the seabed soil analysis.

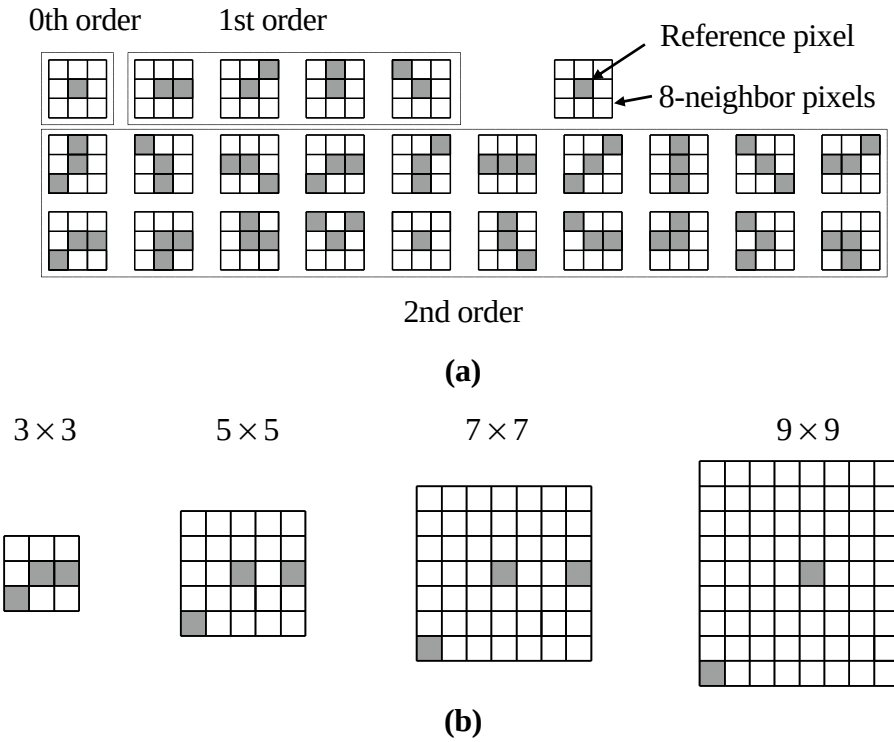


Figure 4-3. Local patterns for calculating HLAC features:

(a) Local 3 ´ 3 masks up to the second order, (b) used 4 sizes of the mask patterns.

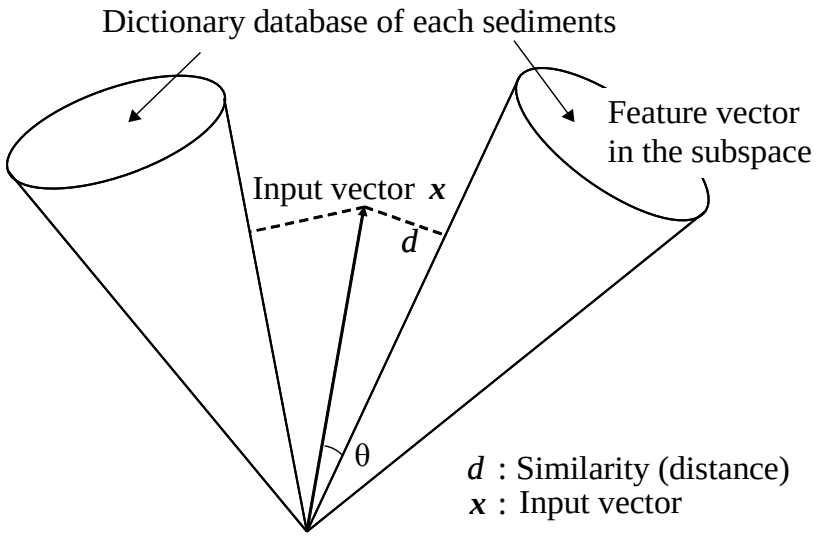


Figure 4-4. Schematic of the subspace method.

4.2.3 Recognition by the subspace method

Figure 4-4 shows the concept of the subspace method [38]. The subspace method is a method of class discrimination. The proposed method classifies the seabed sediments by the subspace method in which each class of the seabed sediment has its own subspace and an unknown image is classified into one of those subspaces by a distance measure. Each class is learning the subspace of low-dimensional representations of the class. By determining the unknown patterns that are to be approximated best in each subspace, the classes of unknown patterns can be identified.

Given N learning images of sediments, the HLAC feature vectors $\boldsymbol{x}_i$ (i = 1, 2, …, N) can be calculated. The covariance matrix C is calculated using the feature vectors as follows:

$$C = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{x}_i - m)(\boldsymbol{x}_i - m)^T \tag{4.4}$$

Here, m is an average vector of $\boldsymbol{x}_i$. Eigen values l_i and the corresponding eigenvectors $\boldsymbol{u}_i$ are calculated from the covariance matrix. The eigenvalues are arranged in descending order. Compression of the feature space is done using the accumulated contribution ratio h_K ($K \le N$) defined by

$$h_K = \frac{\sum_{i=1}^{K} l_i}{\sum_{i=1}^{N} l_i} \tag{4.5}$$

where N is the number of dimensions and K eigenvectors $\boldsymbol{u}_1, \mathbf{L}, \boldsymbol{u}_K$ are chosen, where K is the minimum value satisfying $h_K > t$. Here, t is a threshold determined experimentally. These K eigenvalues create a subspace.

If we denote $\boldsymbol{U}_K = (\boldsymbol{u}_1, \mathbf{L}, \boldsymbol{u}_K)$, the projection matrix is expressed as $\boldsymbol{U}_K \boldsymbol{U}_K^T$. Then, the distance d between an unknown feature vector $\boldsymbol{x}$ and the subspace defined by $\boldsymbol{U}_K$ is given by

$$d = \boldsymbol{x}^T \boldsymbol{x} - \boldsymbol{x}^T \boldsymbol{U}_K \boldsymbol{U}_K^T \boldsymbol{x} \tag{4.6}$$

The value d is the index for the classification in the proposed method.

If there are M kinds of seabed sediments, we define M subspaces. Given an unknown vector, the distances d_m to a sediment class m (m=1, 2, ... , M) are examined to find the minimum value. In case the unknown vector does not belong to any class, it is rejected. The rejection occurs if the minimum distance exceeds a certain threshold, or the following inequality holds with respect to any m and a given threshold q ;

$$\min_K \left\{ \frac{\boldsymbol{x}^T \boldsymbol{U}_{m,K} \boldsymbol{U}_{m,K}^T \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \right\} > q \tag{4.7}$$

where $U_{m,K}$ defines the subspace of the sediment class m and $\boldsymbol{x}$ is an unknown vector.

4.3 Experimental results

4.3.1 Experimental setup

For the experiments, side-scan sonar images obtained in actual environments are used for the analysis. Two sites were used for exploration: a muddy seabed and a sandy seabed. After lowering the sonar device into the water from the stern of the research vessel, an area about 2000 meters wide and 2,000 meters long was explored and sound images of the seabed were obtained. The water within the scope of the exploration area was about 40 meters deep. The area was inspected by adjusting the sonar's towing depth to a set interval of 15 meters from the sea surface. The total navigation time was approximately 8 hours. **Table 3-1** shows the specifications used for the side-scan sonar. The acoustic reflectance was decided by the quality of the bottom of the range scanned by the side-scan sonar.

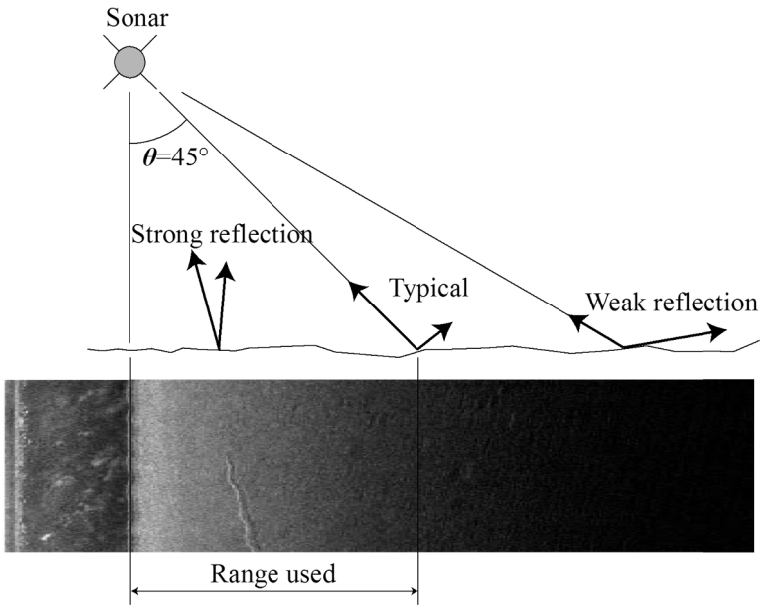


Figure 4-5. The reflection intensity by irradiation angle.

The reflectance of rock and gravel is higher than sand and mud, and reflectance from rock and gravel show up brightly on the side-scan sonar records. Furthermore, the shape of the individual components of such material has a large influence on the reflectance and strength of backscattering. Similarly, the seafloor topography is a factor that influence acoustic reflectance. Assuming that the terrain the sonar unit is facing represent a plane view, when the incidence angle of the irradiated sound from the plane seabed become large, the reflected sound waves show up brighter on the records. In contrast, return values become extremely weak when the incidence angle is smaller than 45 degrees. Therefore, the reflection intensity showed a change of the signal strength according to the degree of acoustic irradiation. Hence, the angle of the sound wave irradiation by side-scan sonar is limited up to 45 degrees in this study (**Figure 4-5**).

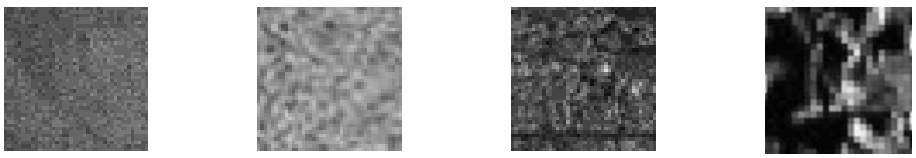


Figure 4-6. Categories of seabed sediment images.

4.3.2 Result of classification

Classification was performed with respect to four types of representative sediments, namely mud, sand, gravel, and rock, from the sound images obtained from side-scan sonar in a real seabed environment. In addition, the relationship was examined between the classification rate and the size of the window used for calculating the HLAC features. Examples of the seabed sediments are shown in **Figure 4-6**. The images were separated into four types of mesh regions (each having 12×12, 24×24, 36×36, 48×48 pixels) and the feature vector at each region was calculated to classify it. The classification rate is defined by the following formula:

$$\text{Classification rate} = \frac{\text{Correctly classified regions}}{\text{All the divided regions}} \tag{4.8}$$

100 images were prepared for with respect to each of the four sizes of windows (12×12, 24×24, 36×36, 48×48 pixels) and with each of the four types of sediments (mud, sand, gravel, and rock) to make the image database. In total, this amounts to 1,600 images. The computer used for the calculations has a 4 GB memory and a 2.66 GHz Intel Core 2 Duo processor. **Table 4-1** shows the classification rate for each sediment type in relation to the size of the mesh region. As shown in **Table 4-1**, the sand region had the highest classification rate (93.2%). This may be because the sand region was spread uniformly and gave a high reflectance of the acoustic wave, which resulted in high contrast ultrasound images. The low classification rate in gravel and rock regions may have been the result of the fact that gravel and rocks are often mixed with sand and those regions do not always contain only gravel or rocks. The result of the classification with respect to the dimension of the subspace is given in **Figure 4-7**. The mesh region with 36×36 pixels achieved the highest classification rate.

Table 4-1. Relation of the classification rate with each seabed sediment type and the size of the mesh region along with the value *K* giving the maximum classification rate.

[%]

Endpoint	Pixels			
	12×12	24×24	**36 ´ 36**	48×48
Subspace dimension *K*	15	17	**25**	24
Mud region	57.9	75.3	**84.6**	85.0
Sand region	65.3	83.1	**93.2**	90.1
Gravel region	45.3	58.0	**63.5**	61.8
Rock region	42.1	45.9	**51.6**	48.2
Average rate	52.7	65.6	**73.2**	71.3

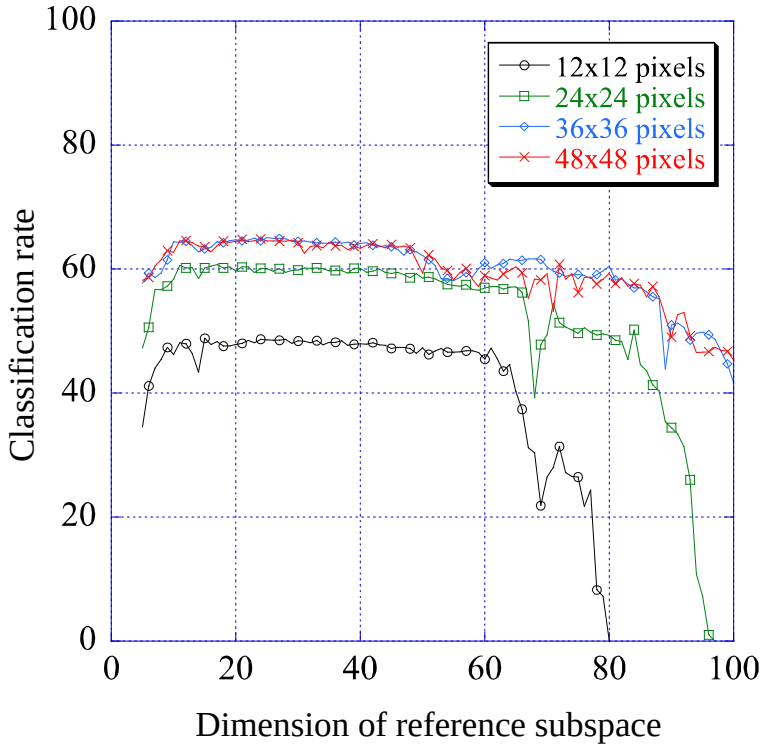


Figure 4-7. Results of the classification. The abscissa is the dimension of the subspaces and the ordinate is the classification rate. A graph is shown with every size of the mesh region.

Figure 4-8 shows the overall classification result for the seabed sound images by the proposed method. In this figure, the very fine mud region is shown by green, the sand region is yellow, the gravel region is red, and the rock region is blue. Directly under the research vessel, the transmission pulse from the side-scan sonar is emitted almost vertically to the seabed and the shadow of the sound wave is hard to discern: These areas are displayed as blank areas on both sides of the black line in **Figure 4-8**. Other blank areas are those where the acquisition of seabed images is difficult because of the weak reflection intensity of the sound wave. The seabed sediments in this part of the water column between the main body of the sonar and the bottom of the sea could not be classified. On the other hand, some of the domains where the reflection strength was relatively weak were able to perform the classification because of smaller incidence angles (less than 45 degrees) of the sound wave.

Although misclassification sometimes occurred, the results are satisfactory for the classification according to the information from sea charts based on geological background data. The classified regions were distributed uniformly as well as continuously.

Table 4-2. Comparison of the classification rate with previously used methods and the proposed method.

Identification Technique	Classification Rate
Higher-order local autocorrelation	**73.2**
Gray level co-occurrence matrix	64.7
Local binary patterns	40.3

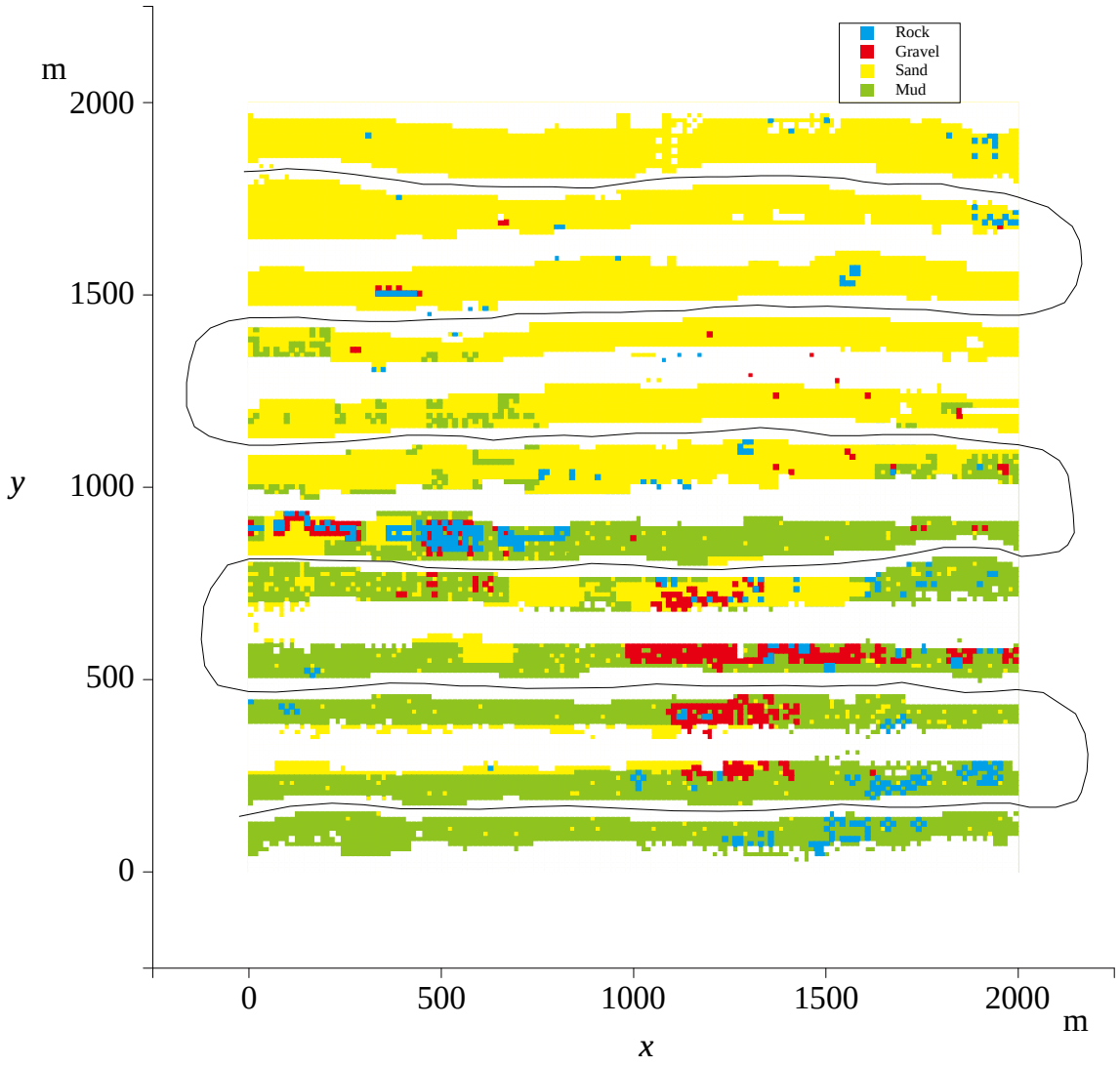


Figure 4-8. Visual display of the result of the classification: The mud region is shown by green, the sand region yellow, the gravel region red and the rock region blue.

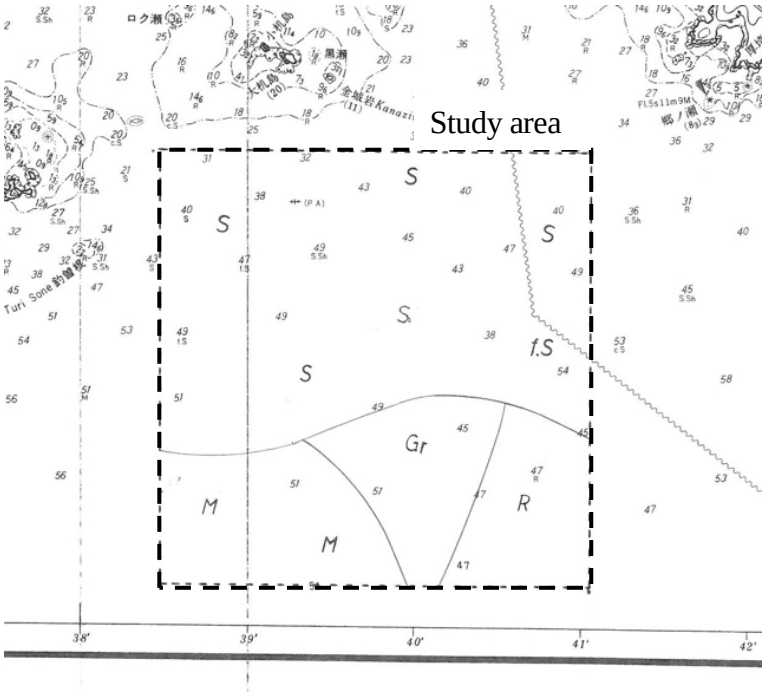


Figure 4-9. Coastal chart of the study area.

4.4 Discussion

Although ground-truth data are necessary for determining the precision of the classification, it is actually difficult to use such data during seabed sediments classification. Therefore, the charts provided by the Japan Hydrographic Association (JHA) were used to evaluate the experimental results. The JHA charts classify seabed sediments based on the geological background. **Figure 4-9** shows a JHA chart of the study area.

As shown in **Figure 4-7**, classification results are better in large mesh regions compared to small mesh regions. This is because the amount of information is reduced in smaller sized images. Since there was hardly any difference between the cases obtained with 36×36 pixels and with 48×48 pixels, the most suitable domain was determined to be 36×36 pixels; this yielded the best classification rate of 73.2%. Since the HLAC feature is an integral feature, larger regions may decrease the difference of the HLAC feature among the set regions. Therefore, there may be a region of optimal size with respect to the recognition area.

Some other kinds of texture feature extraction were performed to compare the results with the proposed method. The examined features are the gray level co-occurrence matrix

and the LBP, which are used widely in texture analyses. In the texture feature extraction by the gray level co-occurrence matrix, entropy, local homogeneity, moment, and contrast at various angles (θ = 0, 45, 90, 135) and a set distance (d = 1) were employed. The LBP is a method for comparing the gray value of a pixel with its 8-neighbors and it expresses large or small correspondences by numbers of 1 or 0 to make an 8-bit binary number. **Table 4-2** compares the classification rate in the entire study area between the proposed method and the other methods. The classification made by LBP were the lowest at 40.3%. The reason for this may be that the LBP compresses the original image into a 256 level gradation image regardless of the actual gradation of the original image. Therefore, it is likely that important data on small shadow differences of mud and sand in the original image are lost during the normalization. **Table 4-2** shows that the proposed method achieved the best classification rate in comparison to the two other methods.

The classification rate for the proposed method of 73.2% was not very high. It is, however, a reasonable result, considering that there are hardly any clear borders among the seabed sediments. It should also be noted that, to the best of our knowledge, this is the first practical result obtained for the automatic classification of seabed sediments on the acoustic image. The fact that the classification rate reached about 90% in the sandy region suggests that further improvements of the proposed method in other regions are possible.

There were several false-classifications and false-rejections between mud and sand, and sand and gravel. One of the main reasons for this misjudgment is that the sound images obtained from the border between mud and sand, for example, contained a mixture of both of the sediment types. Hence, larger number of training images for such sediment borders may be necessary to resolve this difficulty.

4.5 Conclusion

In this chapter, the author proposed a method for classifying seabed sediments using ultrasound images provided by a side-scan sonar. The proposed method classifies the seabed sediments using the HLAC feature and the subspace method. It has been confirmed throughout various experiments on texture classification that the HLAC feature is effective in classifying seabed sediments from their ultrasound images, although the classification precision depends on the size of the texture region. Since calculation of the HLAC feature is simple, high-speed operation is possible with appropriate hardware. Because of this fact, the proposed method may offer a highly practical and useful method for future research.

In order to improve the precision of the classification, accumulation of sample data to be used as training data will be necessary. In particular, the exact classification of regions along the border of different types of seabed sediments represents a problematic issue that should be solved in future research.

Chapter 5

Conclusion

In this paper, the author has proposed a method that uses marine acoustic images obtained from a side-scan sonar to extract the information on the seabed. The method can be used to identify and characterize objects that may be present on the seafloor. The method does not require that other sensors be used to extract the information. The contents of each chapter and the major findings of this research are described as follows.

In Chapter 2, an overview of underwater sound was given and key principles of sound wave technology were explained. In particular, side-scan sonar technology, which was the focus of this study, was described in detail. This study notes that the development and use of marine acoustic technology has expanded in modern times. In addition to the basic physics related to acoustic waves, much research has been dedicated to other basic and applied fields such as electronics, physical oceanography, signal processing, and biology. The realization of new sonar systems that utilize advanced detection algorithms are expected to contribute to major breakthroughs in oceanographic research that require deployment to novel marine environments and other areas of natural resource interest

Chapter 3 described a technique for detecting objects on the seabed from acoustic images where there are multiple objects present. Based on images obtained from a side-scan sonar, the author proposed a method for detecting objects at high speed. The method uses a machine learning process that employs cascaded AdaBoost and Haar-like features. By using a technique based on the *k*-means method to determine the Haar-like features, the number of patterns of Haar-like features was minimized and the proposed method was capable of detecting undersea objects faster than current methodology. This

study demonstrated the effectiveness of this method by applying it to the detection of real objects imaged on the seabed (i.e., sandy ground and muddy ground). This method opens up the possibility of conducting analyses in real-time.

Chapter 4 described an automated technique for using acoustic images obtained by side-scan sonar to perform classification of the seafloor sediment. The proposed method classifies the seabed sediments using higher-order local auto-correlation (HLAC) features and the subspace method. It has been confirmed throughout various experiments of texture classification that the HLAC features are effective to use for classifying seabed sediments from ultrasound images, although the classification precision depends on the size of the textured region. Since calculation of the HLAC features is simple, high-speed operations are possible via current hardware. Because of this fact, the proposed method represents a highly practical and useful method for oceanographic research.

The proposed method would benefit from additional research in the future. Because the seabed object shapes contained shadows, the proposed method was effective for detecting objects underwater. In the future, it will be necessary to consider another approach for undersea objects that do not form shadows. The proposed technique used a *k*-means method, and the classification class number *k* was determined manually. In the future, it would be desirable to automatically categorize the data from a group of classification classes, and consider pre-processing techniques for undersea objects that do not depend on the brightness and contrast of the original images. In order to improve the precision of the classification, accumulation of sample data that can be used as training data will be necessary. The improved classification of regions along the border of the seabed sediments is also a problem that needs to be solved in the future.

One of the most important features of the proposed method is that it uses only acoustic images obtained from a side-scan sonar. Therefore, data from additional sensors are not required in marine surveys. This feature opens up the possibility of installing this technology in unmanned small underwater vehicles.

In conclusion, the author hopes that this study will contribute to the development of advanced marine technology in the future.

Appendix A

Evaluation of features for seabed object detection

Section 3.2.2 of the main text described seabed object detection using Haar-like features. This Appendix A calculates Gabor features and wavelet features and evaluates the use of these features for seabed object detection. This study used Haar-like features for seabed object detection, which relies on the positional relationship between an object and its shadow projected on a seabed acoustic image. A Gabor filter, which also can be used for seabed object detection, is said to imitate the human visual system and is robust to changes in illumination. Therefore, a Gabor wavelet and a Haar wavelet were used to extract features and the results were compared with those of the proposed method in terms of the accuracy and speed of seabed object detection.

A.1 Wavelet transformation

Wavelet transformation is a feature extraction algorithm that can simultaneously extract color, texture, shape, and positional information of images as well as flexibly handle dilatation, rotation, and translation of images. In this section, a Haar wavelet was used as a mother wavelet for wavelet transformation. Although the Haar wavelet is problematic in that it creates distortion during image processing, it is commonly used because of its ease of implementation.

The discrete wavelet used in this section can be expressed as follows:

$$y_{j,k}(t) = 2^{-j/2} y(2^{-j} t - k) \tag{A.1}$$

where the shift orthogonality is given by

$$\left\langle y(t-k), y(t-n) \right\rangle = \begin{cases} c, & n = k \\ 0, & n \neq k \end{cases}. \tag{A.2}$$

When c = 1 in Equation (A.2), the vectors are orthonormal. The orthogonality of scaling is given by

$$\left\langle 2^{-m/2} y(2^{-m}t), 2^{-n/2} y(2^{-n}t) \right\rangle = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}. \tag{A.3}$$

The simplest wavelet satisfying Equations (A.2) and (A.3) is the Haar wavelet. Its equation and waveform are shown in Equation (A.4) and **Figure A-1**:

$$y(t) = \begin{cases} 1, & 0 \le t < 1/2 \\ -1, & 1/2 \le t < 1. \\ 0, & \text{otherwise} \end{cases} \tag{A.4}$$

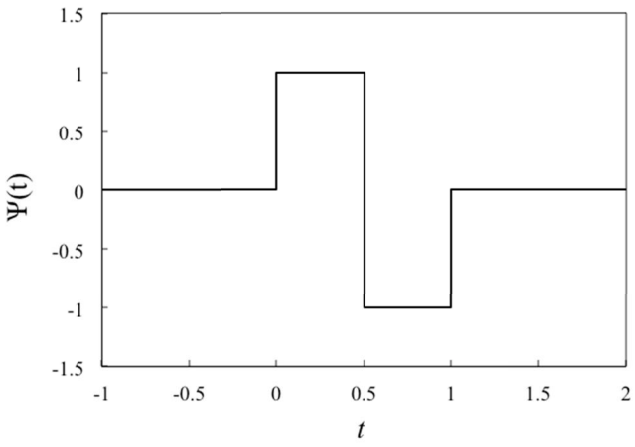


Figure A-1. Haar wavelet waveform.

Wavelet feature vectors are obtained by sub-band decomposition. That is, an image of seabed objects ($2^n \times 2^n$) is subjected to one-dimensional wavelet transformation in the horizontal direction, which gives a low-frequency component and a high-frequency component. Subsequently, each component after horizontal transformation is subjected to one-dimensional wavelet transformation in the vertical direction. This procedure is shown in **Figure A-2**. Finally, the image is split into four components, as shown on the far right in **Figure A-2**. By placing vectors of the LL, LH, HL, and HH components of the transformed image in a column, the feature vector of wavelet transformation $\boldsymbol{x}_i = (\boldsymbol{f}_{LL}, \boldsymbol{f}_{LH}, \boldsymbol{f}_{HL}, \boldsymbol{f}_{HH})^T$ will be obtained. When the image subjected to wavelet transformation is split into four components (see **Figure A-2**), the upper left component represents reduction information of the original image and the remaining three components represent shape (edge) information of the original image. Figure A-3 shows some results of feature extraction from images of seabed objects by Haar wavelet transformation.

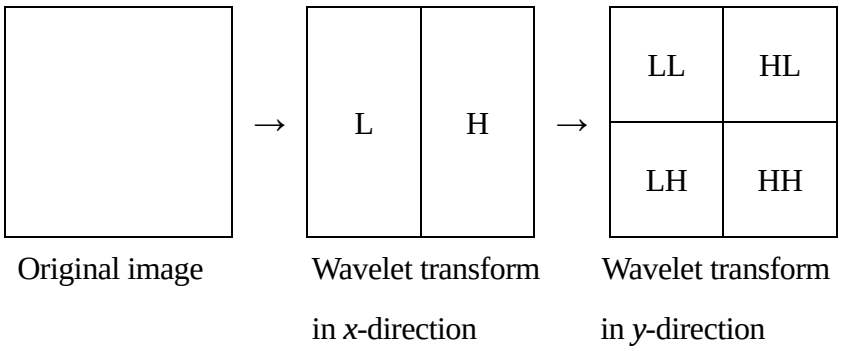


Figure A-2. Wavelet transformation

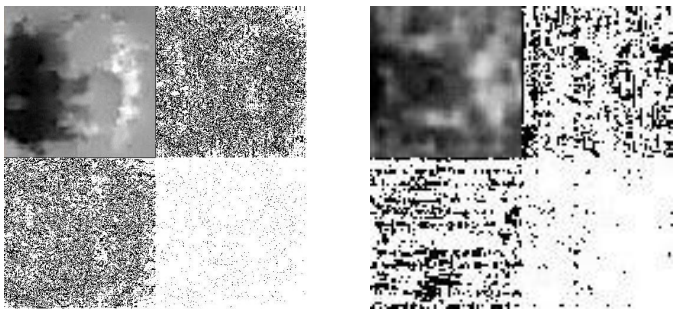


Figure A-3. Images obtained by feature extraction of seabed objects via Haar wavelet transformation.

A.2 Gabor wavelet filtering

Gabor wavelet filtering is one of the continuous wavelets contained in Gabor functions used for window Fourier transforms. Like direction-selective neurons in the primary visual cortex of an organism, the Gabor wavelet filtering can extract specific directions and widths of the edges; therefore, Gabor feature-based feature extraction has been widely used as a recognition and detection method in recent years.

Gabor wavelet transformation can extract both periodicity and directionality of the gray level around a feature point as features by convolving Gabor filters with different resolutions and directions, and the Gabor filters used for feature extraction are known as a model imitating characteristic of the primary visual cortex. A Gabor filter that can extract information at a local area is defined by

$$\psi_{\boldsymbol{k}}(\boldsymbol{x}) = \frac{\boldsymbol{k}^2}{\sigma^2}\exp\left(\frac{-\boldsymbol{k}^2\boldsymbol{x}^2}{2\sigma^2}\right)\left[\exp\left(i\boldsymbol{k}\boldsymbol{x}\right) - \exp\left(\frac{-\sigma^2}{2}\right)\right] \tag{A.5}$$

where the parameters are as follows:

$$\begin{aligned}
&\boldsymbol{x} = (x, y)^{\mathrm{T}} \\
&\boldsymbol{k} = k_v \exp(i\phi) = (k_v \cos\phi, k_v \sin\phi)^{\mathrm{T}} \\
&k_v = \frac{k_{\max}}{f^v} \\
&\phi = \frac{\pi\mu}{8}
\end{aligned} \tag{A.6}$$

While the above mentioned features imitate the human visual system, features called four directional features are known to be effective in common recognition techniques. Four directional features are expressed by assigning an edge gradient of each pixel to one of four directions, and these are like the simplified Gabor features mentioned above. Specifically, the four images are created by classifying edge gradients of pixels obtained by the Prewitt operator into four (horizontal, vertical, and two diagonal) directions on the

basis of the edge direction and resolution of the images is reduced by applying a Gaussian filter. Even after reducing the resolution, the four directional features retain the edge direction information and become robust to changes in edge shape. In addition, the number of features can be reduced by resolution reduction. **Figures A-4** and **A-5** show a real part and an imaginary part, respectively, of the Gabor filter when $\Phi = \pi\mu/8$, $\nu = 0$, $\mu = \{0, \cdots, 7\}$, $\sigma = \pi$, $k_{max} = \pi/2$, and $f = \sqrt{2}$. **Figure A-6** shows examples of seabed object images to which the Gabor filters were applied.

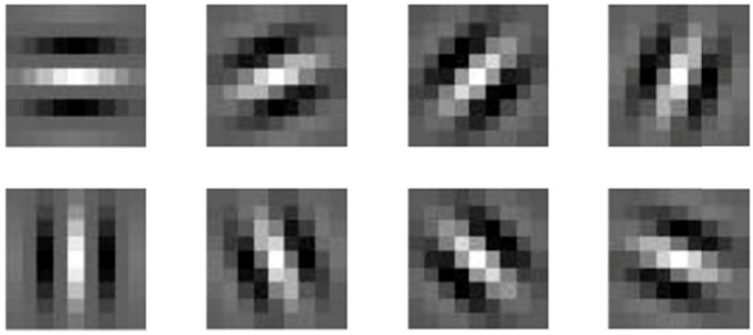


Figure A-4. Real part of the Gabor filter when ν= 0 and μ = {0,...,7}.

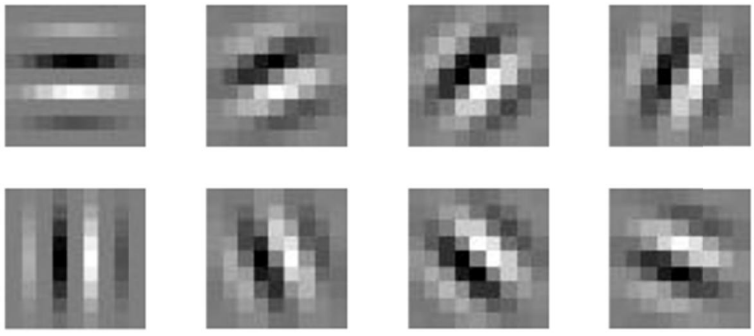


Figure A-5. Imaginary part of the Gabor filter when ν= 0 and μ = {0,...,7}.

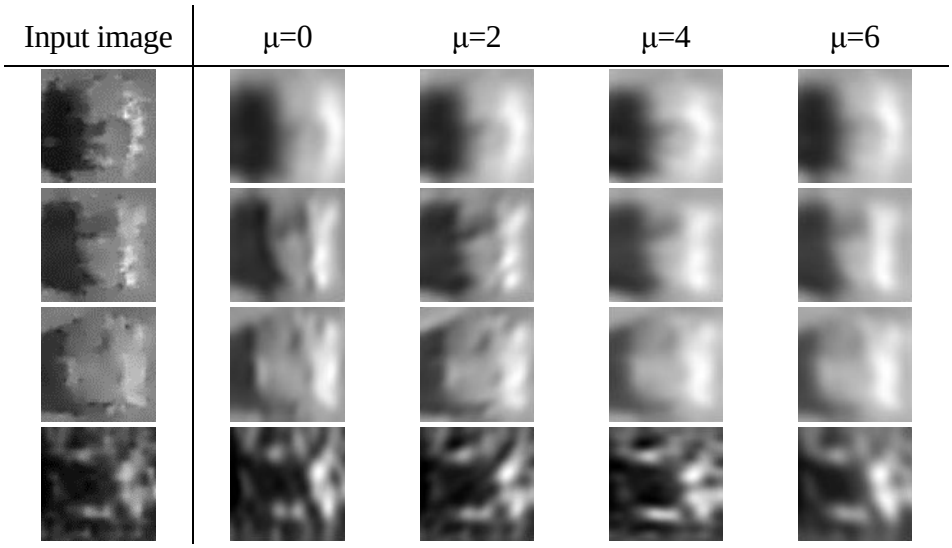
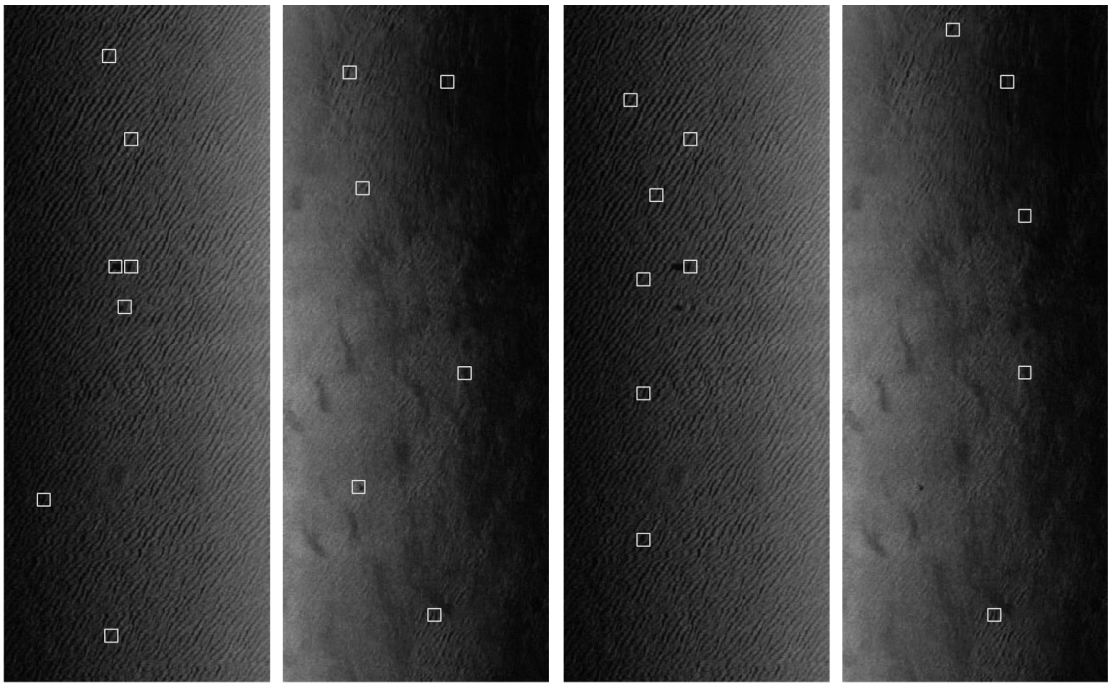


Figure A-6. Seabed object images to which the Gabor filters were applied.

A.3 Experimental results

Figure A-7 shows results of seabed object detection using the Haar wavelet features and the Gabor features mentioned above. As compared to the proposed method, the number of misdetections increased in sandy places (sand waves). A likely reason why the misdetection number increased is that these alternative techniques detected even the patterns formed by sand waves because the Haar wavelet features and the Gabor features used local features. **Table A-1** shows a comparison of the detection rate accuracy and detection speed for the proposed method and alternative techniques. In the case of the proposed method, the detection speed was able to be increased while retaining the accuracy, whereas the accuracy was poorer when the other two methods were used. This may be because the Haar-like features use overall features. Overall, these results demonstrate that the proposed method is more useful for seabed object detection than the other two methods.



(a) Haar wavelet features (b) Gabor features

Figure A-7. Results of seabed object detection using the Haar wavelet features and the Gabor features.

Table A-1. Comparison of the accuracy of seabed object detection by the proposed method and the other features.

Method	True Positives [%]	False Positives [%]	Detection Time [ms]
Proposed method	**82.6**	**17.4**	**65–90**
Sawas et al [13].	81.9	18.1	141–205
Haar wavelet	24.0	76.0	82–98
Gabor filter	23.1	76.9	61–69

Appendix B

Experiments with scale-space seabed classification

Section 4.2.2 of the main text described the classification of seabed soils on the basis of higher-order local auto-correlation (HLAC) features that use features from high-frequency components to low-frequency components. This Appendix B describes the experimental results of the classification of seabed soils that were obtained by calculating HLAC features based on scale space (multi-resolution pyramid).

Scale space (multi-resolution pyramid) consists of a set of several images with different, from high to low, resolutions. **Figure B-1** shows the correspondence of pixels among different levels in an image pyramid.

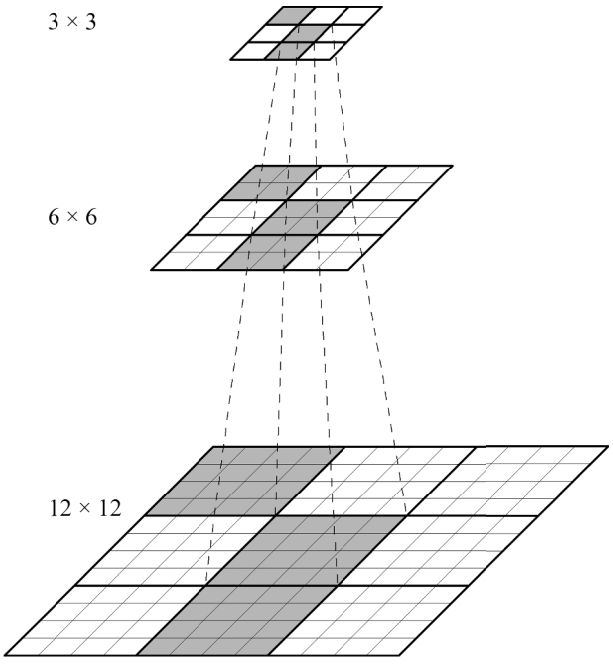


Figure B-1. Schematic of a multi-resolution pyramid.

Scale space (multi-resolution pyramid) is composed by splitting an original image into small regions that do not overlap each other, replacing each small region by the average of the small region, and repeating a series of the operation. When extracting features based on HLAC from images of scale space (multi-resolution pyramid), a set of the features contains information ranging from a detailed one to a rough one. In addition to that, the HLAC features are taken over by the set of the features.

Figure B-2 shows the results of the classification of seabed soils by scale space (multi-resolution pyramid). The recognition area for this classification was 36 × 36 pixels. As compared to the proposed method, the sand region was enlarged and the other regions decreased. Specifically, the region that had been classified as mud by the proposed method was reclassified as sand by the scale-space-based method and it increased in size. Moreover, the region that had been classified as gravel or rock by the proposed method was reclassified as mud by the scale-space-based method and its size increased. This may be because detailed information on seabed soils (the texture feature) had been lost during the averaging process that was conducted in low resolution mode; hence, differences in features among the seabed soils were diminished.

Table B-1 shows a comparison of the classification rates between the proposed method and the scale-space-based method. Classification rates were lower during the application of scale space in all regions; in particular, the classification rate of the rock region decreased significantly. This fact also indicates that the proposed method is more useful for seabed classification than the method that uses features based on the multi-resolution pyramid.

Table B-1. Comparison of classification rates for the proposed method and when scale space was applied.

	Classification Rate	
Endpoint	Proposed Method	Application of Scale Space
Mud region	84.6	78.6
Sand region	93.2	89.2
Gravel region	63.5	50.1
Rock region	51.6	33.8
Average	**73.2**	**62.9**

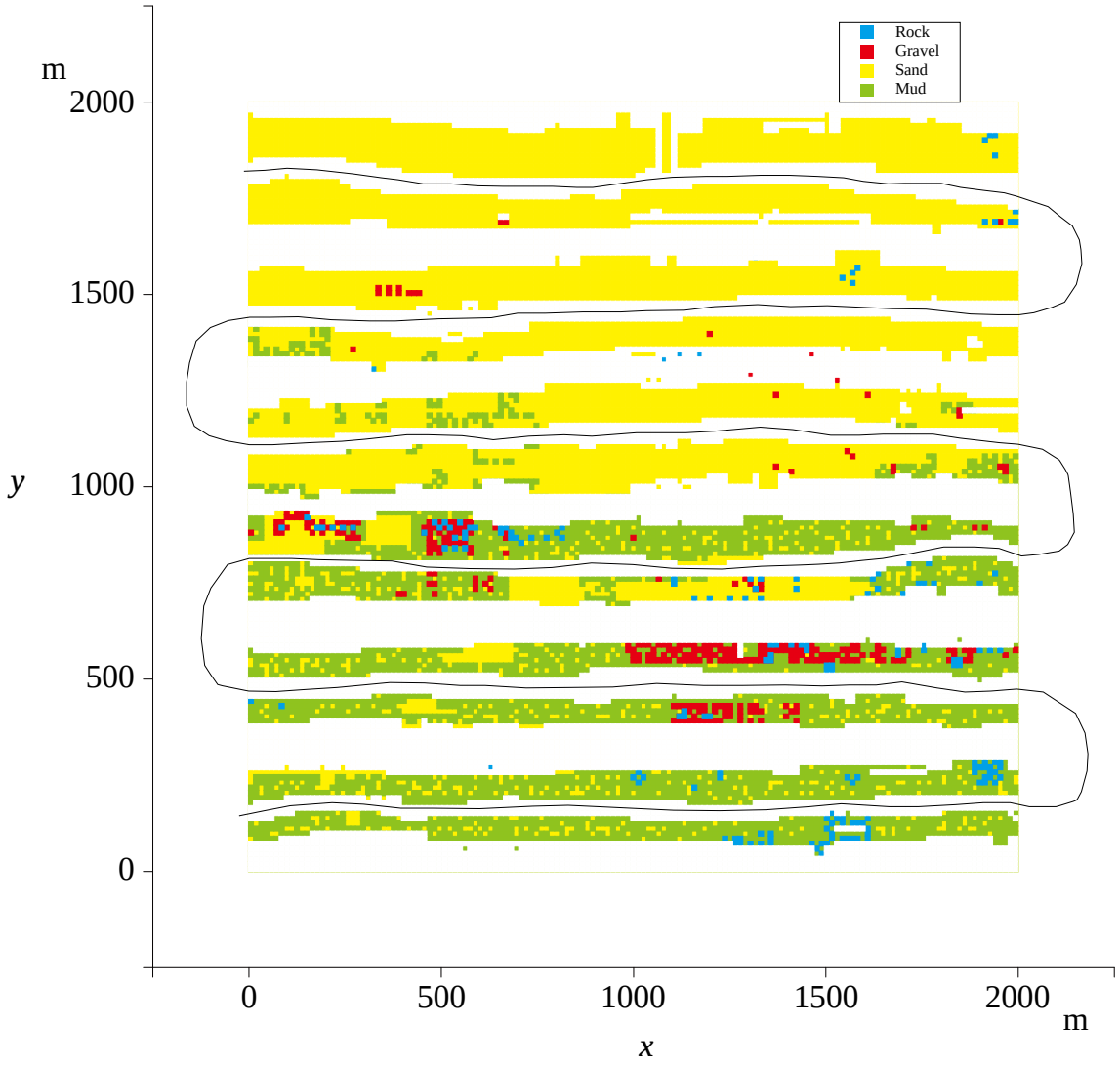


Figure B-2. Results of the classification using HLAC features based on scale space.

References

[1] K. Sayanagi, et. al., " Development of precise magnetic exploration system for seabed resources: Preliminary result of test in sea area using AUV," available at http://www.eqh.dpri.kyoto-u.ac.jp/CA/2010/Sayanagi_et_al_CA2010.pdf, accessed on October 19, 2012.

[2] H. K. Farr, "Multi-beam bathymetric sonar: SEA-BEAM and HYDROCHART", Marine Geodesy, vol. 4, pp. 77-93, 1980.

[3] W. J. Capell and K. C. Kiesel, "An extend capability multibeam bathymetric mapping system", Proceedings of IEEE Conference on OCEAN'S 89, pp1124-1136, 1989.

[4] USC and GS Scientific and Technical Publications Group, "Narrow beam transducer sounding system yields excellent results in operation aboard U.S.C and G.S ship Surveyor", International Hydrographic Review, vol. XLIII, no. 1, pp. 37-42, 1966.

[5] K. Talukdar, W. Capell, and C. Zabounidis, "Sidescan survey results from a multibeam sonar system-SEA BEAM 2000", Marine Geodesy, vol. 15, pp. 63-79, 1992.

[6] A. Asada, "Sea Beam 2000: Bathymetric surveying with interferometry", Sea Technology, vol. 33, no. 6, pp. 10-15, 1992.

[7] A. Asada, "New bathymetric surveying and processing system based on SeaBeam2000", *Report of Hydrographic Research*, vol. 29, pp. 1-32, 1993.

[8] J. P. Fish and H. A. Carr, *Sound Underwater Images: A Guide to the Generation and Interpretation of Side Scan Sonar Data*, Lower Cape Pub, 1990.

[9] J. M. Bell, E. Dura, S. Reed, Y. R. Petillot, and D. M. Lane, “Extraction and classification of objects from sidescan sonar,” available at http://dsor.isr.ist.utl.pt/Projects/Exocet/Priv/Bibliography/CollectedByUAlg/NLNGSP_02.pdf, accessed on August 8, 2011.

[10] C. M. Ciany, W. C. Zurawski, I. Kerfoot, “Performance of fusion algorithms for computer-aided detection and classification of mines in very shallow water obtained from testing in navy Fleet Battle Exercise-Hotel 2000,” Proceedings of SPIE, vol. 4394, pp. 1116–1122, 2001.

[11] B. Zak, " The Search and Identification of Underwater Object by Sonar and Vision systems of Underwater Vehicle", Proceedings of the 11th WSEAS International Conference on SYSTEMS, pp. 238-243, 2007.

[12] S. W. Perry and L. Guan, “Pulse-length-tolerant features and detectors for sector-scan sonar imagery,” IEEE Journal of Oceanic Engineering, vol. 29, no. 1, pp. 138–156, 2004.

[13] J. Sawas, et. al., “Cascade of boosted classifiers for rapid detection of underwater objects,” ECUA 2010 Istanbul Conference, 2010.

[14] V. Murino and A. Tucco, " Three-dimensional image generation and processing in underwater acoustic vision", Proceedings of the IEEE, vol. 88, no. 12, pp. 1903-1948, 2000.

[15] J. Aulinas, et. al., “Robust automatic landmark detection for underwater SLAM using side-scan sonar imaging,” Proceedings of the 11th International Conference on Mobile Robots and Competitions, pp. 21–26, 2011.

[16] S. H. Peckinpaugh, “An improved method for computing gray-level cooccurrence matrix based texture measures,” CVGIP: Graphical Models and Image Processing, vol. 53(6), no. 5B, pp. 574–578, 1991.

[17] F. Yamamoto, T. Takeuchi, H. Tokuyama, K. Suehiro, and A. Taira, “Classification of seafloor characteristics using IZANAGI side scan sonar imagery,” Transactions of Institute of Electronics, Information, and Communication Engineers, vol. 25, pp. 29–35, 1994.

[18] Atallah, L., and PJ. Smith. "Automatic seabed classification by the analysis of sidescan sonar and bathymetric imagery," IEE Proceedings-Radar, Sonar and Navigation, Vol. 151(5), pp.327-336, 2004.

[19] T. Kanade, “Survey on region segmentation: Signal vs semantics,” Computer Graphics and Image Processing, vol. 13, pp. 279–297, 1980.

[20] D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Inc., 1982.

[21] R. Nevatia, “Image segmentation,” in T. Y. Young and K.-S. Fu, eds., Handbook of Pattern Recognition and Image Processing, chapter 9, Academic Press, 1986.

[22] Y. Shirai, *Three-Dimensional Computer Vision*, Springer-Verlag, 1987.

[23] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

[24] N. R. Pal and S. K. Pal, “A review on image segmentation techniques,” Pattern Recognition, vol. 26, no. 9, pp. 1277–1294, 1993.

[25] P. Viola and M. J. Jones, “Rapid object detection using a boosted cascade of simple features,” 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 511–518, 2001.

[26] N. Otsu, “Mathematical studies on feature extraction in pattern recognition,” Researches of the ETL, vol. 818, 1981 (in Japanese).

[27] J. P. Fish and H. A. Carr, Sound Underwater Images: A Guide to the Generation and Interpretation of Side Scan Sonar Data, Lower Cape Pub, 1990.

[28] J. Naoi, H. Saitou, T. Shimura, H. Monma, and T. Tsuchiya, “Search for objects on the deep-sea floor using side scan sonar with a tilted arrangement transducer array,” Japanese Journal of Applied Physics, vol. 40(1), pp. 3777–3781, 2001.

[29] M. A. Ainslie and J. G. McColm, “A simplified formula for viscous and chemical absorption in sea water,” Journal of the Acoustical Society of America, vol. 103, no. 3, pp. 1671–1672, 1998.

[30] P. Viola, M. Jones, "Robust real-time object detection," Proc. of Workshop on Statistical and Computational Theories of Vision, 2001.

[31] T. Kohonen, Self-Organizing Maps, Springer, 2000, ISBN 978-3540679219.

[32] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, “Face recognition by independent component analysis,” IEEE Transactions on Neural Networks, vol. 13, no. 6, pp. 1450–1464, 2002.

[33] F. Goudail, et al., “Face recognition system using local autocorrelations and multiscale integration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18(10), pp. 1024–1028, 1996.

[34] T. Kurita, N. Otsu, and T. Sato, “A face recognition method using higher order local autocorrelation and multivariate analysis,” Proceedings of the 11th IAPR International Conference on Pattern Recognition, Vol. II, Conference B: Pattern Recognition Methodology and Systems, IEEE, pp. 213–216, 1992.

[35] H. Cheng, et al., “Spam image discrimination using support vector machine based on higher-order local autocorrelation feature extraction,” IEEE Conference on Cybernetics and Intelligent Systems, pp. 1017–1021, 2008.

[36] T. Kurita and H. Satoru, “Gesture recognition using HLAC features of PARCOR images and HMM based recognizer,” Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 422–427, 1998.

[37] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24(7), pp. 971–987, 2002.

[38] E. Oja and M. Kuusela, “The ALSM algorithm—an improved subspace method of classification,” Pattern Recognition, vol. 16(4), pp. 421–427, 1983.