

基于浅地层回波信号的底质分类技术研究

罗宇¹, 郑旭¹, 施剑^{1*}, 吴逸凡^{1,2}, 李斌¹

(1. 山东科技大学 测绘学院, 山东 青岛 266590; 2. 自然资源部第二海洋研究所, 浙江 杭州 310012)

摘要: 海底底质分类与识别具有十分重要的意义。利用浅地层剖面仪信号频率低、底质穿透力强的特点, 本文对浅剖原始数据进行处理, 分析了不同底质回波信号的能量序列分布特征。根据回波能量序列特点, 利用阈值检测方法提取出含有丰富海底浅层沉积物特征的回波能量序列, 减小了水深对分类精度的影响, 并利用支持向量机 (Support Vector Machine, SVM) 进行海底底质自动分类。利用该方法对实验水槽采集的淤泥、砂砾、水泥 3 种底质数据进行支持向量机训练, 并对分类模型的关键参数 g 与惩罚系数 c 进行优化, 优化后的模型分类精度达到 96%。实验结果表明, 该方法可有效应用于海底底质分类。

关键词: 底质分类; 支持向量机; 浅地层剖面仪; 回波能量序列

中图分类号: P229 **文献标识码:** A **文章编号:** 1003-2029 (2020) 06-0042-06

海底底质类型是重要的海洋环境参数, 是进行海底科学研究的基础, 海底底质分类对于海洋工程建设、海底科学研究、海洋资源开采具有重要的科学与实际意义。传统的底质分类依靠机械进行直接采样来确定沉积物的类型及分布, 该方法效率低、采样数据离散, 不适合进行大面积的调查分类。声学方法遥测海底类型根据海底底质的反射系数、声速、散射等声学特性反演出底质的类型、粗糙度、硬度等属性, 具有效率高、费用低、获取内容丰富连续的特点, 为海底分类提供了快速可靠的方法^[1]。对于单波束、浅地层剖面仪等正入射仪器进行底质分类的主要方法有海底声学参数反演法^[2]、统计特征分析法等^[3]; 对于多波束、侧扫声呐等斜入射仪器进行底质分类的主要方法有反向散射强度分析法^[4-5]、图像纹理特征分析法等^[6-9]。目前, 对于回波信号的统计多基于回波信号的波形特征, 对声呐回波能量特征研

究较少。

浅地层剖面仪 (Sub-bottom Profiler), 简称“浅剖”, 是一种用于获知海床以下浅部地层信息的声学探测设备, 是进行海洋地质勘探常用的设备之一, 其工作时通过发射换能器垂直向海底发射声波信号, 声波穿透水层、海底地层, 下行过程中经各层介质的滤波, 遇声阻抗界面返回, 最终携带有界面层信息的反射声波被接收单元接收^[10]。相较于其它设备, 浅剖具有发射频率低、底质穿透力强的特点, 获得的回波信号可携带很丰富的沉积物特征。本文主要对不同底质下浅地层回波信号的能量序列进行分析, 提取海底浅层沉积物的回波能量序列进行地质分类。

1 浅表地层回波能量序列

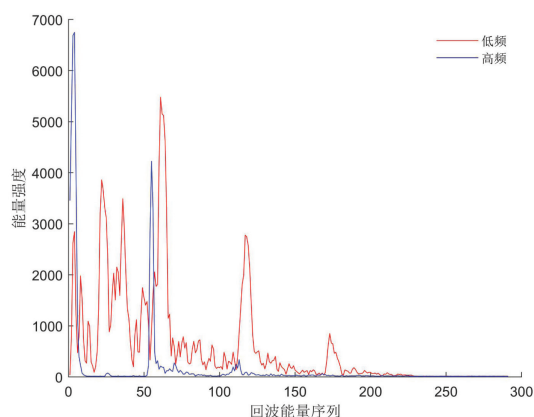
浅剖在工作中发射的脉冲强度是固定的, 接收的回波强度由于深度和海底不同沉积物的反射

收稿日期: 2020-06-30

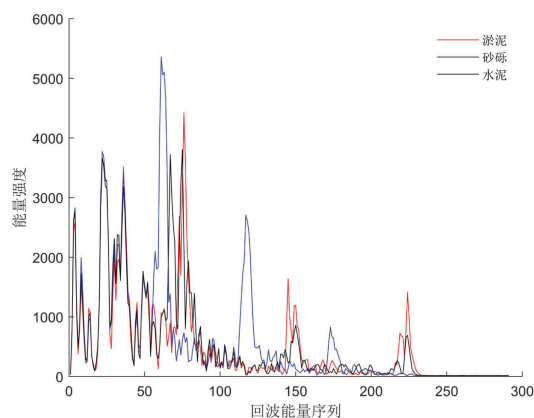
作者简介: 罗宇 (1974-), 男, 博士, 教授, 主要从事声学信号处理研究。E-mail: luoyu@sdust.edu.cn

通讯作者: 施剑 (1978-), 男, 博士, 副教授, 主要从事水声探测与定位研究。E-mail: shijian@sdust.edu.cn

系数、硬度、粗糙度等物理性质不同,其回波的能量序列特征也有所不同。具体表现为:(1)回波能量序列的总强度不同。不同底质之间回波能量序列的总强度存在较大差异,且与深度有关,深度较浅回波能量序列的总强度较强,较深回波能力较弱;(2)回波能量序列分布特征不同。浅剖发射的脉冲信号接触到海底沉积物,受到反射的同时,也会有一部分能量被吸收,利用反射的部分可以进行测深,而不同的沉积物对声波的吸收不同,其能量分布存在较大的差异^[11];(3)回波能量序列分布与换能器的工作频率有关。工作频率越低,其回波能量的平均强度越高,且出现波峰的序列越多,工作频率越高,其回波总能量越低,波峰数量越少。



(a)



(b)

图1 回波能量序列分布特征

对浅表层回波信号数据进行处理,可分离出回波能量序列,接收的1帧数据可以换算为垂直方向上有限的回波能量值,其数量与浅剖的发射频率、声速等有关^[12],可将回波能量序列映射为水深,计算公式为:

$$h = \frac{n(H_{\text{MAX}} - H_{\text{MIN}})}{S} \quad (1)$$

式中: S 为1帧中回波能量的数量; n 为回波能量的序列值; H_{MAX} 为声呐设置的最大量程,为 H_{MIN} 测量最小量程。

图1-a为淤泥底质在不同工作频率下的浅剖能量序列分布情况,高频(200 kHz)下其分布中只出现两个较为明显的波峰,第一个波峰为水面反射,第二个波峰为声波到达海底表面后的反射。低频(20 kHz)下则出现了多个波峰,其平均能量也更高、低频穿透性较强,其分布更能反映出不同底质的特征。图1-b为淤泥、砂砾、水泥3种不同底质的回波能量序列分布,可以观测到不同底质的回波能量序列分布具有差别。

2 底质分类方法

支持向量机(Support Vector Machine, SVM)是一种有监督的分类算法,其最初由 Vapnik 首先提出^[13],其主要思想是通过映射函数将向量映射到一个更高维的空间,求解一个分类超平面,使正例和反例之间的隔离边缘被最大化。支持向量机针对有限的样本数量下的分类问题,根据结构风险最小化原则在保证学习精度的前提下提高了泛化能力。考虑到在实际的勘测工作中,海底样本采集困难、样本数量较少等情况,使用支持向量机对海底底质进行分类识别研究。

2.1 支持向量机原理

支持向量机最初用于对线性可分问题进行分类,给定训练样本 $\{(x_1, y_1), (x_2, y_2), (x_3, y_3) \cdots (x_n, y_n)\}$, 得到期望输出 $y \in \{-1, 1\}$ 。即寻找一个超平面将训练样本分离,并使任意样本的点到超平面的距

离大于等于 1。用于分类的最优超平面的方程为：

$$\begin{aligned} w^T x + b &= 0 \\ y_i(w^T x_i + b) &\geq 1 \end{aligned} \quad (2)$$

式中： w 为法向量； b 为截距。满足上述条件的超平面实际构建了两个平行的超平面作为间隔边界以判别样本的分类：

$$\begin{aligned} (w^T x_i + b) &\geq +1 \Rightarrow y_i = +1 \\ (w^T x_i + b) &\leq -1 \Rightarrow y_i = -1 \end{aligned} \quad (3)$$

所有在上间隔边界上方的样本属于正类，在下间隔边界下方的样本属于负类，位于间隔边界上的正类和负类样本为支持向量（Support Vector）。对于非线性问题，需适当放松约束条件，引入松弛变量 $\xi_i > 0$ ，此时的分类条件变为：

$$\begin{aligned} y_i(w^T x_i + b) &\geq 1 - \xi_i \Rightarrow y_i = +1 \\ y_i(w^T x_i + b) &\leq -1 + \xi_i \Rightarrow y_i = -1 \end{aligned} \quad (4)$$

当 $\xi_i = 0$ 时，问题变为线性分类问题，当 $0 < \xi_i < 1$ 时，样本可落在超平面内并被超平面正确分类，当 $\xi_i > 1$ 时，允许错误分类。松弛因子越大表示对离散点的容忍程度越高。同时为松弛因子加入一个惩罚系数 c ，以免松弛因子过大。此时的目标函数为：

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (5)$$

对于非线性分类问题，可使用映射函数将其转化为线性可分问题，此时作为决策边界的超平面表示为 $w^T \varphi(x) + b = 0$ 。公式中 $\varphi(x)$ 即为映射函数，可定义映射函数的内积为核函数 $k(x_1, x_2) = \varphi(x_1)^T \varphi(x_2)$ 。常见的核函数有多项式核、径向基（RBF）函数核、拉普拉斯核等。本次实验使用 RBF 核函数：

$$K(x, x_i) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0. \quad (6)$$

2.2 分类处理方法

在进行底质分类之前，首先对回波能量强度数据进行预处理，根据回波能量序列的特点，最大限度地消除非海底底质因素的影响，提取出代表底质特征部分的序列，对数据进行分类整理，

选取适当数量的数据作为 SVM 的训练集和测试集，对 SVM 进行训练，并用测试集进行测试，得到分类结果。流程如图 2 所示：

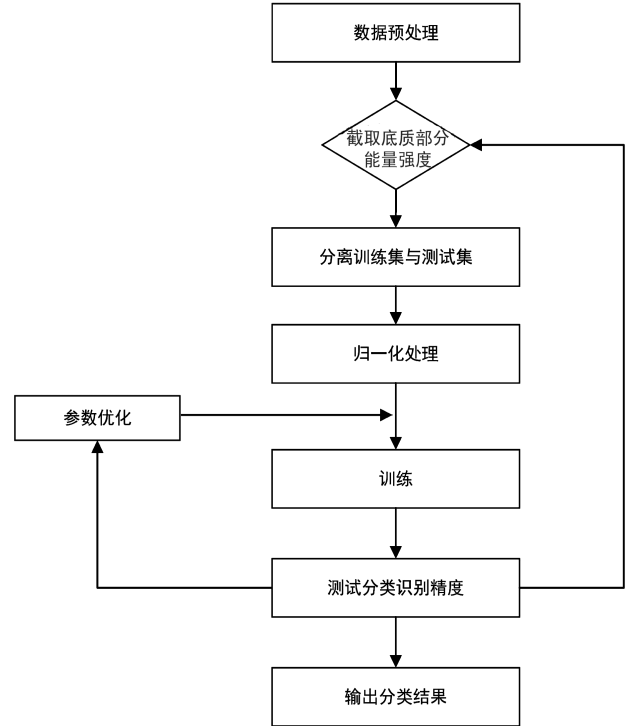


图 2 底质分类处理流程

回波能量序列分布与深度有关，进行数据选取时，需剔除深度的影响，利用阈值算法来获取回波能量序列中代表底质特征的部分。如图 3 所示设每帧回波能量序列为 $\{P_1(h_1, k_1), P_2(h_2, k_2), \dots, P_n(h_n, k_n)\}$ ，其中 h_i 为序列值， k_i 为能量强度。根据测量时的水深值可根据公式（1）计算出水深的序列 D ，根据水深序列 D 设定合适的范围 $\{D \in [h_{\min}, h_{\max}]\}$ ，即图 3 中的黑色虚线。在设定的阈值内，寻找能量强度的最大值得到水底表面的深度序列 h ，选取适当的厚度 Δh ，截取海底表面深度 h 后的 $[h, h + \Delta h]$ 范围内的回波能量序列作为底质分类的有效原始数据，即图 3 中的红色虚线。对数据进行筛选，这样可以有效地过滤掉水面反射回波、二次回波等其它异常的回波能量。对底质厚度的选取，可在一定的厚度内进行循环测试，根据分类成功的准确率选取合适的底质厚度。

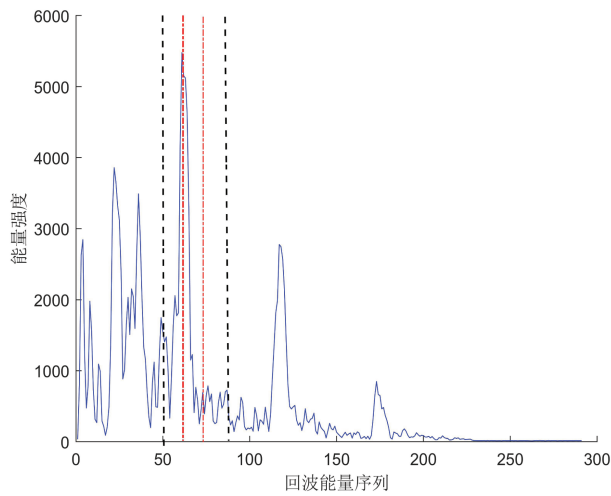


图3 底质特征数据选取

支持向量机在实际应用中,惩罚参数 c 与核函数参数 g 的选择对于最后的分类效果具有十分重要的影响,选择恰当的参数十分必要。在原始数据底质已知的情况下,可将参数 c 与 g 在一定范围内取离散值,将原始数据一部分作为训练集,一部分作为测试集,并对参数进行归一化处理,归一化处理可将不同来源的特征数据统一到同一数量级下,消除不同特征间量纲的影响。使用离差标准化,使结果值映射到 $[0, 1]$ 之间,转换函数如下:

$$k_i' = \frac{\max\{k_i\} - k_i}{\max\{k_i\} - \min\{k_i\}} \quad (7)$$

式中: k_i 表示序列 i 的能量强度; $\max\{k_i\}$ 表示所有样本数据中回波能量强度最大值; $\min\{k_i\}$ 表示所有样本数据中回波能量强度最小值。训练完成后利用测试集对分类准确率进行测试,如图2所示。以测试集的分类准确率作为评价标准,不断对参数进行优化测试,最终得到一定条件下最佳的惩罚参数 c 与核函数参数 g 。

3 实验数据分析

本次实验数据来源于山东科技大学海洋测绘

重点实验室消声水池,该水池长 40 m,宽 6 m,水深 3 ~ 8 m。水池池壁放置了消声材料,水池底部有规律阶梯状布设有卵石、砂砾、淤泥、水泥等不同底质,模拟真实的海洋环境。采集数据使用到的仪器为某型号浅地层剖面仪、GPS 定位系统、处理软件南方自由行 2016 等。设置浅地层剖面仪的最大工作深度为 20 m,声速为 1 500 m/s,保持参数不变,分别对淤泥、砂砾、水泥 3 种底质进行数据采集,每种底质选取不同的点进行 5 次采集,每个采集点采集数据 5 分钟。对数据进行回放,选取其中 4 组回波较清晰的信号。

每组选取 50 帧数据,其中 100 帧作为训练集,100 帧作为测试集,根据上述方法,首先对数据进行处理,利用阈值算法截取底质能量序列。如图4所示,为了剔除水深的影响,截取的序列为表面波峰后的 13 列。可以观测到不同底质之间的回波能量序列具有较明显的差异。水泥底质回波序列中会有一个明显的波峰,砂砾底质与水泥底质相比波峰较小,且波峰序列靠前,淤泥底质回波序列能量强度缓慢下降。

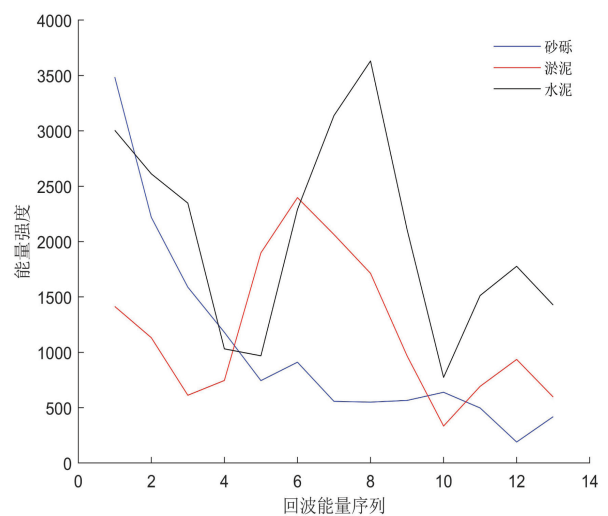


图4 不同底质回波能量序列

选取能量回波序列的多少即选取的底质厚

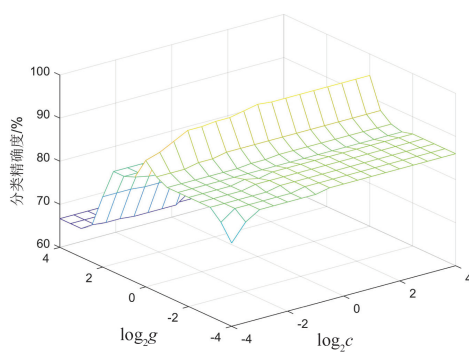
度不同, 对分类结果影响很大, 默认 SVM 参数 $c=1$, $g=1$, 不同厚度下的分类结果如表 1 所示。

表 1 不同序列长度下分类准确率

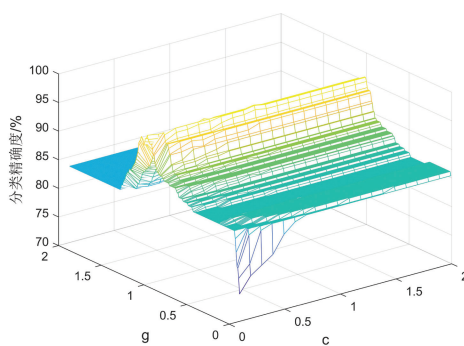
序列长度	底质厚度 /m	分类准确率 /%
5	0.33	80.67
6	0.40	72.67
7	0.47	67.33
8	0.54	73.00
9	0.60	76.67
10	0.67	75.33
11	0.74	80.00
12	0.80	88.33
13	0.87	95.33
14	0.94	93.67

通过表 1 观察到, 在 SVM 参数不变的情况下, 截取能量序列长度为 13 时分类准确率为 95.33%, 而选取序列长度为 7 时分类准确率仅为 67.33%, 本次所取的能量强度序列为 13, 其对应的底质厚度为 0.87 m。

本次测试集的底质类型已知, 设定不同的参数 c , g 利用训练集进行训练, 训练完成后利用该支持向量机对测试集进行分类, 预测结果与实际底质类型结果对比可选出最优的参数 c 与 g , 如图 5 所示。先在较大范围内粗略的选取 c 与 g , 将 c 与 g 范围设为 $[2^{-4}, 2^4]$, 递进步长取 2 为底的幂指数 0.5, 结果如图 5-a 所示, 在 $\log_2 g=0$ 处, 即 $g=1$ 处分类准确率较高, 且参数 g 对分类准确率影响较大, 进一步缩小范围, 将 c 设为 $[0, 2]$ 步进间隔为 0.1, 将 g 设为 $[0, 2]$ 步进间隔为 0.01, 得到结果如图 5-b 所示。在这种条件下的最优参数为 $c=1$, $g=1.01$ 。



(a)



(b)

图 5 SVC 参数优化示意图

利用优化参数后的支持向量机, 对测试集 3 类底质共 300 个样本数据进行自动分类, 分类结果如图 6 所示, 预测底质类型与测试集已知的底质类型相同, 分类准确率为 96%。其中淤泥有 2 个样本被错分为水泥, 水泥有 10 个样本被错分为淤泥。

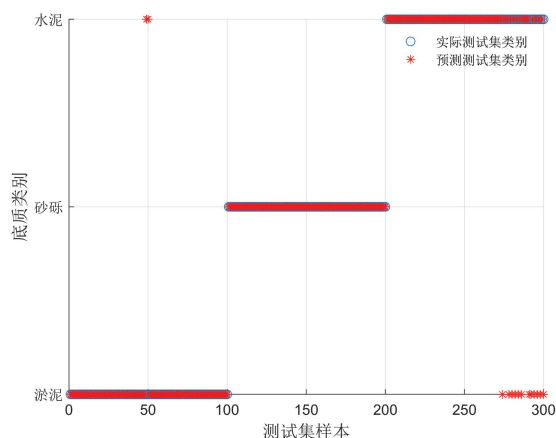


图 6 底质分类测试结果

4 结 论

从浅地层剖面仪回波信号中提取的回波能量序列,可有效地应用于海底底质分类。本文总结了浅地层回波信号能量序列的特征,提出了一种简单有效的截取含有底质特征的回波能量序列的方法。在分类方法上,利用 SVM 算法进行分类,优化了 SVM 的参数,使其在一定意义上达到了最优,在实验水池进行数据采集,浅剖工作状态与

工作环境较为理想,且采集的底质数据样本数量较少,每个底质类型用于训练的数据相似度较高,所以取得了很高的分类精度。浅地层剖面仪在实际的测量工作中面临船运动、环境噪声等多种因素的影响,且海底底质类型复杂多变,需要进一步利用浅剖采集真实的海底底质数据,并在数据预处理、底质特征选取以及多底质类型分类等方面进行更深入的研究。

参考文献:

- [1] 陶春辉,金翔龙,许枫,等.海底声学底质分类技术的研究现状与前景[J].东海海洋,2004,22(3):28-33.
- [2] 潘国富.声学方法进行海底沉积物遥测分类[J].海洋技术,1997,16(1):14-19.
- [3] 谈振藩.从声纳反射信号判断海底沉积物类型[J].哈尔滨工程大学学报,1998,17(2):26-33.
- [4] 唐秋华,周兴华,丁继胜,等.多波束反向散射强度数据处理研究[J].海洋学报(中文版),2006,28(2):51-55.
- [5] 金绍华,肖付民,边刚,等.利用多波束反向散射强度角度响应曲线的底质特征参数提取算法[J].武汉大学学报(信息科学版),2014,39(12):1493-1498.
- [6] 熊明宽,吴自银,李守军,等.基于 SVM 的海底声纳图像底质识别[J].海洋通报,2012,31(4):409-414.
- [7] Pace N G, Gao H. Swathe seabed classification[J].IEEE Journal of Oceanic Engineering, 1988,13(2):83-90
- [8] 徐超,李海森,王川,等.基于合成核 SVM 的多波束海底声图像底质分类研究[J].地球物理学进展,2014,29(5):2437-2442.
- [9] 杨词银,许枫.基于二次反锐化掩模的多特征侧扫声纳成像海底底质分类[J].电子学报,2005(10):115-118.
- [10] 李一保,张玉芬,刘玉兰,等.浅地层剖面仪在海洋工程中的应用[J].工程地球物理学报,2007(1):4-8.
- [11] 徐铭,王宁,高大治.窄带高频声信号海底底质分类研究[J].中国海洋大学学报(自然科学版),2006,36(S2):127-130,38.
- [12] 李吉祥,周新聪,初秀民,等.基于单波束仰扫回波信号强度处理的船舶吃水检测技术[J].船海工程,2017,46(6):12-16,22.
- [13] CortesC,VapnikVN. Supportvector networks[J].Machine Learning,1995,20:273-297.

Study on the Seabed Classification Approach Based on Echo Signal of Sub-bottom Profiler

LUO Yu¹, ZHENG Xu¹, SHI Jian¹, WU Yi-fan², LI Bin¹

1.College of Geomatics, Shandong University of Science and Technology, Qingdao 266590, China;

2.Second Institute of Oceanography, MNR, Hangzhou 310012, China

Abstract: Research on seabed classification and recognition is of great significance. Based on the characteristics of low transmission frequency and strong penetration of sub-bottom profiler, the original data of sub-bottom profile are processed, and the energy sequence distribution characteristics of echo signals from different substrates are analyzed. According to the characteristics of the echo energy sequence, the threshold detection method is used to extract the echo energy sequence with rich characteristics of seabed sediments which can reduce the influence of depth on the classification accuracy, with the support vector machine adopted to classify seabed sediment in an automatic way. With this method, the support vector machine (SVM) is trained for the data of silt, gravel and cement collected from the laboratory flume, and the key parameters g and penalty coefficient c of the classification model are optimized. The classification accuracy of the optimized model reaches 96%. It is proved that this method can be effectively used in the classification of seabed sediments.

Key words: seabed classification; support vector machine; sub-bottom profiler; echo energy sequence