

· 研究简报 ·

# 鲸类声信号的分类系统设计

叶 蓁<sup>1,2</sup>, 孙海信<sup>1\*</sup>, 颜佳泉<sup>1</sup>, 陈清峰<sup>1</sup>, 齐 洁<sup>1</sup>

(1. 厦门大学 信息科学与技术学院, 水声通信与海洋信息技术教育部重点实验室, 福建 厦门 361005;

2. 闽南师范大学物理与信息工程学院, 福建 漳州 363000)

**摘要:** 针对传统的水声信号分类技术处理方法复杂、特征提取时间长以及特征量多等问题, 提出了一种基于稀疏表示的分类系统, 先利用正交匹配追踪法(OMP)算法提取与水声信号最为匹配的少数原子作为目标特征, 再采用支持向量机(SVM)进行分类. 对鲸类声信号进行仿真实验, 实验结果表明, 不仅提高了压缩效率和运算速度, 而且识别率高, 在水声信号的实时处理中具有较高的实用价值.

**关键词:** 压缩感知; 水声信号; 稀疏表示; 特征提取; 分类识别

**中图分类号:** TP 274. 5

**文献标志码:** A

**文章编号:** 0438-0479(2017)01-0144-05

在水声信号处理中, 特征提取方法及分类器设计是制约目标识别准确率的两个关键环节. 特征提取的任务是选取能有效表征目标身份且稳定可靠的特征. 目前水声信号的特征提取方法主要包括: 时域特征提取、谱估计技术、时频分析技术等<sup>[1]</sup>. 此外, 还有基于语音信号处理的 Mel 频率倒谱系数(MFCC)特征提取, 以及利用分形、极限环、混沌、相空间重构理论<sup>[2]</sup>等进行的新方法探索. 但采用这些特征提取方法所得到的原始特征向量维数很多, 需要大量的存储空间, 也加大了分类判别的运算量, 因而需要进一步进行特征选择, 降低特征空间的维数. 2006 年, Donoho<sup>[3]</sup> 和 Candès 等<sup>[4]</sup> 提出了压缩感知(compressed sensing, CS)理论. 该理论表示: 只要信号是可压缩的或通过某种变换后是可稀疏表示的, 则可用一个与变换基不相关的观测矩阵将变换得到的高维信号投影到一个低维空间上, 再通过求解一个优化问题就能从这些少量的投影中精确或近似地重构信号. 因此, 稀疏表示的低维观测值中必然包含原始信号的大部分有用信息. 自然界的大多数信号都具有一定的稀疏性, 水声信号也不例外. 如果能够构建水声信号的稀疏基, 便可将 CS 理论应用于水声信号的处理, 从实质上降低信号处理的成本, 提高压缩效率, 并增强系统的抗噪声性能, 改

善鲁棒性. 分类器设计的任务是对训练和识别时的特征模式作准确的类别划分. 目前常用的分类决策方法有模板匹配法、统计概率模型法、人工神经网络法和支持向量机(SVM)法等<sup>[5]</sup>.

在不同的应用场合下, 不同的特征提取技术与不同的识别算法组合可以产生不同的识别效果. 李新欣<sup>[6]</sup> 以船舶和鲸类的水声信号作为研究对象, 提出了几种不同的特征提取算法和分类识别模型, 取得较好的识别效果. 其中, 利用概率神经网络设计的分类器对基于小波包分解不同频段的能量特征和基于希尔伯特-黄变换(HHT)的边际谱特征 2 种提取算法得到 3 类差异较大的鲸鱼测试样本的识别率在 90% 左右, 但基于小波包分解的能量特征提取算法难以分辨某些声音频率相近的种类, 且分解层数和小波函数的选取也会对识别结果造成影响, 而 HHT 边际谱的特征有很高的维数运算时间与空间复杂度; 基于 MFCC 和高斯混合模型(GMM)分类的鲸类识别系统在测试时间足够长(30 s 以上)的条件下识别率可以达到 95% 以上, 但在 GMM 阶数的选择上, 阶数较小时, 识别性能一般较差, 而当阶数逐渐增大时, 又会增加运算量, 使模型更为复杂. 本研究提出一种基于 CS 理论中稀疏表示的特征提取与 SVM 分类器相结合的分类系统,

收稿日期: 2016-03-31 录用日期: 2016-06-03

基金项目: 国家自然科学基金(61471309); 福建省自然科学基金(2013J01258)

\* 通信作者: hxsun@xmu.edu.cn

引文格式: 叶蓁, 孙海信, 颜佳泉, 等. 鲸类声信号的分类系统设计[J]. 厦门大学学报(自然科学版), 2017, 56(1): 144-148.

Citation: YE Z, SUN H X, YAN J Q, et al. Design of Classification system of whales sound signal[J]. J Xiamen Univ Nat Sci, 2017, 56(1): 144-148. (in Chinese)



仿真结果表明该系统在获得较高识别率的同时大幅度地提高了运算速度。

## 1 稀疏表示理论

信号的稀疏表示是将信号在一过完备的原子库中进行分解,若原子库中的原子与信号的主要成分相似,则仅需要少数原子的线性组合就能比较精确地表示信号,分解结果也将是稀疏的。

### 1.1 信号的稀疏表示

信号具有稀疏性或可压缩性是 CS 理论的前提条件。假设一维离散信号  $\mathbf{x} \in \mathbf{R}^{M \times 1}$ , 稀疏度为  $K, K \leq M$ , 将  $M$  维信号投影到观测矩阵  $\Phi_{N \times M}$  上,  $N < M$ , 获得  $N$  维观测信号  $\mathbf{y} \in \mathbf{R}^{N \times 1}$ , 用公式表示为

$$\mathbf{y} = \Phi \mathbf{x}. \quad (1)$$

如果  $\mathbf{x}$  为非稀疏信号, 必须先采用某种正交变换使其稀疏化。稀疏后的信号表示为

$$\mathbf{x} = \Psi \boldsymbol{\alpha}, \quad (2)$$

其中:  $\Psi$  为  $M \times M$  维的正交变换矩阵; 稀疏表示系数  $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_K\} \in \mathbf{R}^K$ ,  $\boldsymbol{\alpha}$  为有限个 ( $K$  个) 非零元素, 因此称  $\boldsymbol{\alpha}$  是  $K$  稀疏的, 是信号  $\mathbf{x}$  在某个变换域的稀疏表示。

由式(1)和(2)可得:

$$\mathbf{y} = \Phi \mathbf{x} = \Phi \Psi \boldsymbol{\alpha} = \mathbf{D} \boldsymbol{\alpha}, \quad (3)$$

$\mathbf{D} = \Phi \Psi = \{d_1, d_2, \dots, d_N\}$  是一个  $N \times M$  维的传感矩阵, 又称过完备字典,  $d_i$  称为原子。当观测矩阵与稀疏信号  $\mathbf{x}$  相互独立时,  $N$  维观测信号  $\mathbf{y}$  保留了  $M$  维信号  $\mathbf{x}$  的主要信息。因此, 通过  $\mathbf{y}$  的  $N$  个观测值可以准确地恢复出信号  $\mathbf{x}$ <sup>[7]</sup>。

### 1.2 基于稀疏分解算法的特征量提取

获取信号在过完备字典下的最优稀疏表示或稀疏逼近的过程叫信号稀疏分解算法。稀疏问题可表示为求解

$$\min \|\boldsymbol{\alpha}\|_0, \text{ s.t. } \mathbf{y} = \mathbf{D} \boldsymbol{\alpha}. \quad (4)$$

式中,  $\|\cdot\|_F$  为 Frobenius 范数,  $F=0$  时,  $\|\cdot\|_0$  为非零元素的个数。在实际中, 并不直接对式(4)求解, 而是转化为以下的稀疏逼近问题来求得近似解:

$$\min \|\boldsymbol{\alpha}\|_0, \text{ s.t. } \|\mathbf{y} - \mathbf{D} \boldsymbol{\alpha}\| \leq \epsilon. \quad (5)$$

式(5)的求解属于典型的 NP-hard 问题, 常用的求解算法有基追踪(OP)<sup>[8]</sup>、匹配追踪法(MP)<sup>[9]</sup>和正交匹配追踪法(OMP)<sup>[10]</sup>。本文所采用的 OMP 算法具有收敛速度快, 迭代结果最优的特点。

OMP 算法步骤如下:

- 1) 初始化残差  $\mathbf{r}_0 = \mathbf{y}$ , 索引集  $\Lambda_0 = \emptyset$ , 计数器  $t=1$ ;
- 2) 找出残差  $\mathbf{r}_t$  和字典列  $\mathbf{d}_j$  内积中最大值  $d_j$ , 记录其下标, 即第  $t$  次迭代找到的索引(列序号):  $\lambda_t = \operatorname{argmax}_{j=1, \dots, M} |\langle \mathbf{r}_{t-1}, \mathbf{d}_j \rangle|$ ;
- 3) 更新索引集  $\Lambda_t = \Lambda_{t-1} \cup \{\lambda_t\}$ , 及按索引  $\Lambda_t$  选出的矩阵  $\mathbf{D}$  的集合:  $\mathbf{D}_t = \mathbf{D}_{t-1} \cup \{\mathbf{d}_{\lambda_t}\}$ ;
- 4) 求  $\mathbf{y} = \mathbf{D}_t \boldsymbol{\alpha}_t$  的最小二乘法解, 得到:  $\hat{\boldsymbol{\alpha}}_t = \operatorname{argmin} \|\mathbf{y} - \mathbf{D}_t \boldsymbol{\alpha}_t\|_2$ ;
- 5) 更新信号残差  $\mathbf{r}_t = \mathbf{y} - \mathbf{D}_t \hat{\boldsymbol{\alpha}}_t$ ;
- 6)  $t=t+1$ , 判断是否满足  $t > K$ , 若满足, 停止迭代进入 7), 否则跳转到步骤 2)。
- 7) 重构所得  $\hat{\boldsymbol{\alpha}}$  在  $\Lambda_t$  处的所有非零项, 其值分别为最后一次迭代所得  $\hat{\boldsymbol{\alpha}}_t$ 。

经过  $K$  次迭代后, 即可得到  $\boldsymbol{\alpha}$  为所求的稀疏表示系数。

稀疏分解就是不断跟踪并提取最能匹配原始信号及其残差信号的原子向量的过程, 这些被提取的原子向量可看作是原始信号的特征向量, 而剩余的信号残差则可认为是噪声。

## 2 基于 SVM 的分类技术

SVM 是一种基于统计学习的算法, 是 Cortes 和 Vapnik 等<sup>[11-12]</sup>提出的以结构风险最小化原理为基础的一种分类方法。由于水下环境复杂, 水声信号获取不易且类型多样, 而 SVM 可自动寻找出对分类有较好区分能力的支持向量, 由此构造出的分类器可以最大化类与类的间隔<sup>[13]</sup>, 在解决类似水声信号这样的小样本、非线性及高维模式识别问题中表现出许多的优势。因此, 本文中选择 SVM 作为水声信号的分类识别器。

对训练数据集:

$$(\mathbf{x}_i, y_i), i = 1, 2, \dots, l, \mathbf{x}_i \in \mathbf{R}^M, y_i \in \{+1, -1\},$$

式中  $\mathbf{x}_i$  和  $y_i$  分别表示训练向量和类别标识,  $l$  为训练向量的个数。对线性不可分问题, 可通过非线性变换转化为另一个空间的线性可分问题, 并在变换空间求最优分类面, 其数学形式为:

$$\min \phi(\mathbf{w}, \boldsymbol{\xi}) = \frac{1}{2} (\mathbf{W}^T \cdot \mathbf{W}) + C \sum_{i=1}^l \xi_i. \quad (6)$$

$$\text{约束条件: } y_i (\mathbf{W}^T \cdot \mathbf{x}_i + b) - 1 + \xi_i \geq 0,$$

$$i = 1, 2, \dots, l. \quad (7)$$

式中:  $\mathbf{w}$  为判别函数的权向量;  $\boldsymbol{\xi}$  为松弛变量;  $\mathbf{W}$  为参量, 值越大边界越明显; 常数  $C$  为惩罚参数;  $b$  是阈值。

引入核函数将输入空间映射到高维的特征空间,求最优分类超平面.对  $a_i$  求解下式的最大值:

$$\max W(a_i) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j K(x_i, x_j). \tag{8}$$

式中,  $a_i$  为 Lagrange 乘子, 本文选择线性核函数  $K(x_i, x_j) = x_i \cdot x_j$ . 则最优分类函数表示式为

$$f(x) = \text{sgn}\left(\sum_{i=1}^l a_i^* y_i K(x_i, \alpha) + b_i^*\right). \tag{9}$$

式中:  $a_i^*$  为最优解,  $b_i^*$  为最优偏置; 组中每一个值代表一个特征; 分类输出为 A 类或 B 类(正类或负类).

### 3 基于稀疏表示的鲸类声信号分类识别

海豚和鲸类发出的声信号是生物声源中很重要且较容易观察和记录的一类水声信号. 对此类声信号的研究不仅在海洋生态保护中有重要意义, 而且能为水下信息战提供有效的海洋声信号库, 为水下设备研制及国防安全提供依据. 例如, 将海豚的 whistles 信号用于仿生水声通信的技术研究<sup>[14]</sup>. 本文将基于稀疏表示的分类方法应用于鲸鱼的声音识别.

#### 3.1 水声信号的分类识别步骤

水声信号识别系统由生成样本集、稀疏变换、特征提取和分类识别 4 部分组成.

具体步骤如下:

1) 生成样本集: 首先对采集的水声数据进行截断和归一化处理. 由于采集的信号在比较广的时域和频域范围内波动, 信号时断时续, 因此要先把空白的部分截去, 再把保留下的信号截成长度一致的信号段. 由于各个信号值之间的大小差异很大, 为了避免弱信号被强信号“湮灭”, 降低系统的正确性, 有必要对输入样本进行统一的归一化处理, 将数值幅度限制在  $[-1, 1]$ , 使变量基本处于相同的变化范围内.

2) 稀疏表示: 由于时域内的自然信号基本都是非稀疏的, 在本文中采用快速傅里叶(FFT)正交变换基对样本集中的信号进行变换.

3) 特征提取: 测量矩阵采用高斯随机矩阵, 变换后的信号用 OMP 算法提取字典中最匹配的原子向量作为水声信号的特征向量.

4) 分类识别: 将特征向量送到 SVM 分类器中进行分类识别.

#### 3.2 实验

##### 3.2.1 生成样本集

本文中以鲸类声信号作为样本声音进行实验. 实

验的音频文件来源为网上下载的采集自世界各地的鲸类声信号录音<sup>[15]</sup>, 参与辨识的鲸类共有 5 种. 测试信号的长度为 1 024 帧(约 32 ms), 采样频率  $f_s$  为 32 kHz. 根据上述生成样本集的方法, 选择 5 类声音信号中特征较为完整的 10 058 组信号作为总样本, 并从中随机抽选出一定比例的样本作为 SVM 的训练样本, 其余的用作测试样本.

图 1 所示为采集的蓝鲸声音信号及截断和归一化处理后的声音样本时域波形.

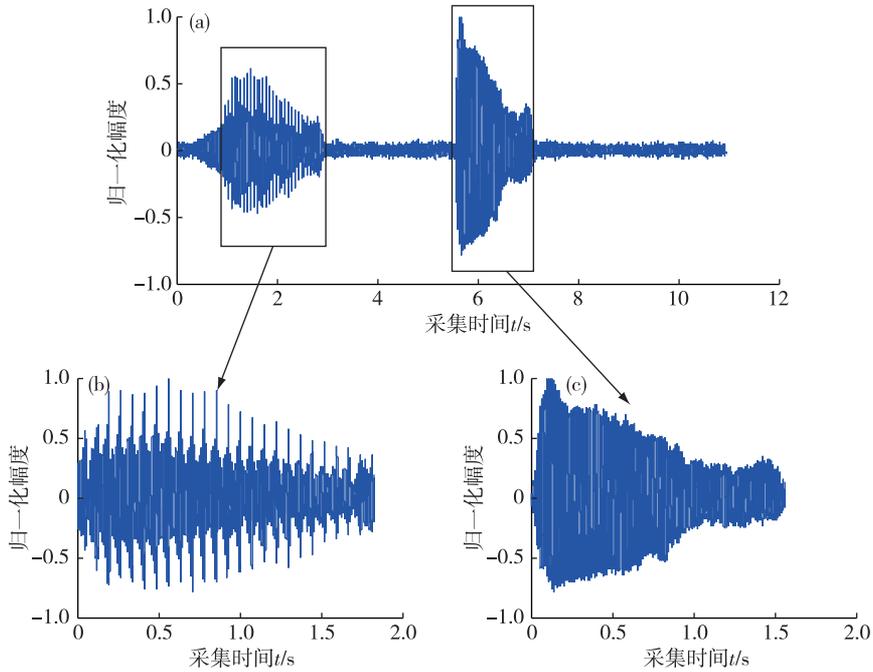
##### 3.2.2 结果与分析

当  $K=10, N=512(N \geq K \times \log(M/K))$ , 迭代次数  $d=20(d \geq K)$  时, 从每种鲸鱼的信号样本中随机抽取相邻的 2 个样本, 采用 OMP 算法计算出每个样本的稀疏表示系数, 得到 20 个稀疏表示系数(即特征量), 稀疏表示系数分布如图 2 所示.

从图 2 中可以看出, 在鲸发出的某个声音片段中, 同一类鲸的稀疏表示系数的变化是相似的, 而不同类型的鲸之间稀疏表示系数是有差别的, 因此稀疏表示系数可以用来作为鲸识别的特征量. 此外, 随着算法迭代次数的增加, 稀疏表示系数的值急剧衰减, 最后趋近于 0. 当迭代次数为 4 时, 稀疏表示系数的值已下降到最大值的 1/3 以下, 由此可见, 信号的大部分信息集中在稀疏分解后提取的少量特征值中, 特别是最前面的这几个系数基本包含了信号的最重要特征, 这些特征便是作为分类的主要依据, 即使把后面的系数去掉也不会对识别的准确率造成太大影响.

但即使是同一类鲸, 其稀疏表示系数也不是唯一的, 不同的稀疏表示系数代表了同类鲸鱼的多种特征, SVM 要通过学习多个不同的采集样本, 才能获取各个种类的鲸较为完整的特征, 实现准确的分类. 当训练样本比例取为 1:25(每 25 个样本中抽取 1 个进行训练)时的分类识别率如表 1 所示. 由表 1 可得到蓝鲸、大白鲸、座头鲸、虎鲸和抹香鲸的平均识别率分别为 95.37%、91.24%、88.65%、90.23%和 95.07%.

表 1 中, 除了座头鲸与虎鲸的识别率较低(83.10%)以外, 其他鲸类之间的识别率都达到 90% 左右, 说明本文中提出的识别方法具有很好的适用性和较好的分类识别效果. 而座头鲸与虎鲸的识别率较低的原因, 可能是由于下载的这 2 种鲸的声音样本之间的类别特征差异较小造成的. 在鲸类中, 虎鲸声音的频谱范围较广, 而且能模仿其他鲸的声音, 当频谱与其他类别的鲸重合度较高时, 识别率也将下降. 本文中下载的这 2 种鲸的声音来自同一海域, 噪声背景较为相似, 这可能也是造成识别率较低的原因之一.



(a)原始波形;(b)、(c)对原始波形作截断和归一化处理后的波形。

图 1 蓝鲸声音信号时域波形

Fig. 1 Time-domain waveform of the blue-whale sound signal

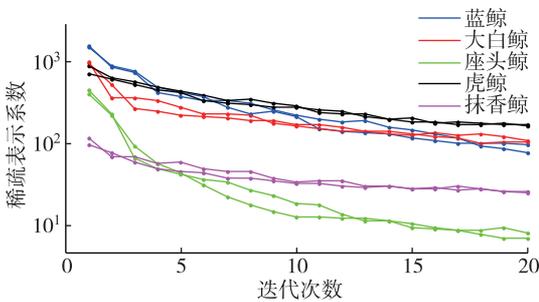


图 2 鲸声音样本的稀疏表示系数

Fig. 2 Coefficient of sparse representation of whale sounds samples

表 1 分类识别率比较( $d=20$ )

Tab. 1 Comparison of classification rate( $d=20$ )

对比类别	识别率/%	对比类别	识别率/%
蓝鲸-大白鲸	97.38	大白鲸-虎鲸	88.21
蓝鲸-座头鲸	88.17	大白鲸-抹香鲸	92.22
蓝鲸-虎鲸	96.84	座头鲸-虎鲸	83.10
蓝鲸-抹香鲸	99.10	座头鲸-抹香鲸	96.20
大白鲸-座头鲸	87.14	虎鲸-抹香鲸	92.77

$d=4$  时,得到的分类识别率如表 2 所示.由表 2 可得蓝鲸、大白鲸、座头鲸、虎鲸和抹香鲸的平均识别率分别为 92.02%、91.60%、86.33%、88.47%和 96.40%。

表 2 分类识别性能对比分析( $d=4$ )

Tab. 2 Comparison of classification rate( $d=4$ )

对比类别	识别率/%	对比类别	识别率/%
蓝鲸-大白鲸	95.34	大白鲸-虎鲸	88.34
蓝鲸-座头鲸	81.34	大白鲸-抹香鲸	95.19
蓝鲸-虎鲸	92.36	座头鲸-虎鲸	79.22
蓝鲸-抹香鲸	99.02	座头鲸-抹香鲸	97.41
大白鲸-座头鲸	87.36	虎鲸-抹香鲸	93.97

从表 1 和表 2 中的数据对比可以看出:特征量减少后,虽然损失了原信号的一部分信息,但用来识别的数据量压缩为原来的 1/5,大大提高了分类器的识别速度,而且识别率并没有明显地降低,甚至部分识别率还有所上升,如抹香鲸的平均识别率提高了 1.33 个百分点.这是由于数值越小的系数包含的特征信息越少,也可能是一些噪声信息,去除后反而使得不同类别的特征区分更加明显.综上所述,使用稀疏表示提取的鲸类声音信号特征向量,多数种类的识别率可达到

为了进一步压缩数据,在进行分类时可选取最前面的几个关键系数作为信号的类别特征.如当迭代次数

90%左右,部分鲸的类型识别率甚至可达到99%以上,但也有个别识别率较低,主要是受采集的声音样本信噪比和鲸类别之间差异不同的影响.总体来说,该分类系统在识别模型、数据压缩、运算时间和识别率等综合考评上具有一定的优势.

## 4 结 论

本研究提出了一种基于稀疏表示理论的鲸类声信号分类识别方法.虽然只是采用鲸类声信号进行了仿真实验验证,但考虑到许多的水声信号都具有稀疏性,该方法也可适用于其他一些水声信号处理的领域,如:舰船噪声的识别等.仿真实验结果表明,本方法用少量的特征向量取代大量的原始数据信号,提高了压缩效率,减少了分类识别的运算时间,且辨识能力强,在水声信号的实时处理中具有较高的实用价值.

### 参考文献:

- [1] 刘深,张小蓓,牛奕龙,等.基于IMF能量谱的水声信号特征提取与分类[J].计算机工程与应用,2014,50(3):203-206.
- [2] 行鸿彦,祁峥东,徐伟.基于选择性支持向量机集成的海杂波背景中的微弱信号检测[J].物理学报,2012,61(24):240504.
- [3] DONOHO D L. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [4] CANDÈS E J, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on In-

- formation Theory, 2006, 52(2): 489-509.
- [5] VAPNIK V. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995: 76-79.
- [6] 李欣新. 船舶及鲸类声信号特征提取和分类识别研究[D]. 哈尔滨: 哈尔滨工程大学, 2012: 72-74, 77-78.
- [7] 乔钢, 王巍, 王玥, 等. 基于压缩感知的 OFDM 水声通信信道二次估计算法[J]. 声学技术, 2013, 32(5): 357-361.
- [8] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit[J]. SIAM Journal on Scientific Computing, 1999, 20(1): 33-61.
- [9] NEEDELL D, VERSHYNIN R. Greedy signal recovery and uncertainty principles[C]// Computational Imaging VI. San Jose: SPIE, 2008: 68140J. doi: 10.1117/12.776996.
- [10] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655-4666.
- [11] VAPNIK V N. Statistical learning theory [M]. New York: John Wiley and Sons, 1998: 81-85.
- [12] CORTES C, VAPNIK V N. Support vector networks [J]. Machine Learning, 1995, 20(3): 273-297.
- [13] 江伟华, 曹秀岭, 童峰. 采用支持向量机的水声通信信号调制识别方法[J]. 厦门大学学报(自然科学版), 2015, 54(4): 534-539.
- [14] 韩笑, 殷敬伟, 郭龙祥, 等. 基于差分 Pattern 时延差编码和海豚 whistles 信号的仿生水声通信技术研究[J]. 物理学报, 2013, 62(22): 244-249.
- [15] MARK A M. Whale acoustics [EB/OL]. [2016-03-01]. <http://www.whaleacoustics.com/audiobaleenwhales.html>.

## Design of Classification System of Whales Sound Signal

YE Zhen<sup>1,2</sup>, SUN Haixin<sup>1\*</sup>, YAN Jiaquan<sup>1</sup>, CHEN Qingfeng<sup>1</sup>, QI Jie<sup>1</sup>

(1. Key Laboratory of Underwater Acoustic Communication and Marine Information Technology, College of Information Science and Engineering, Xiamen University, Xiamen 361005, China; 2. School of Physics and Information Engineering, Minnan Normal University, Zhangzhou 363000, China)

**Abstract:** Traditional classification technologies of underwater acoustic signals are involved with issues such as the complicated processing method, the prolonged feature extraction, vast features and other problems. In this paper, we propose a novel method based on sparse representation classification. First, we extract a spot of atoms matched best with underwater acoustic signals as signal features utilizing the OMP algorithm. Second, we adopt SVM as our classifier. Through experimental evaluations, the effect of this method is shown to provide a significant improvement in compression efficiencies, computing speeds and recognition rates.

**Key words:** compressed sensing; underwater acoustic signal; sparse representation; feature extraction; classification