

听觉外周计算模型在水中目标 分类识别中的应用

王 磊, 彭 圆, 林正青, 蒋行海, 牟 林, 张凤珍

(水下测控技术重点实验室, 辽宁大连 116013)

摘 要: 听觉外周的理论 and 建模已取得长足的发展, 并已广泛应用于语音信号处理. 本文集成 Gammatone 听觉滤波器和 Meddis 内毛细胞模型来模拟耳蜗的处理机制, 并根据水中目标辐射噪声信号的特点对 Meddis 模型的参数进行了修正. 提出基于 Gammatone-Meddis 听觉外周计算模型的水中目标特征提取方法, 得到一个 23 维的特征向量. 对大量海上实测数据的分析表明该特征具有以下优点: (1) 分类识别效果好, 对测试集识别率达到 94% 以上; (2) 抗卷积噪声能力强, 对原始信号加入卷积噪声, 识别能力没有下降. 最后通过实验证明基底膜对频率的非线性选取和内毛细胞都能够很好地抑制噪声.

关键词: 水声信号处理; 听觉模型; 目标识别; 特征提取

中图分类号: TN911.7 **文献标识码:** A **文章编号:** 0372-2112 (2012) 01-0199-05

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2012.01.033

The Application of Computational Auditory Peripheral Model in Underwater Target Classification

WANG Lei, PEN Yuan, LIN Zheng-qing, JIANG Xing-hai, MU Lin, ZHANG Feng-zhen

(Science and Technology on Underwater Test and Control Laboratory, Dalian, Liaoning 116013, China)

Abstract: The theory and model of auditory peripheral has made great progress, and has been widely used in speech signal processing. The text combined Gammatone filter and Meddis model to simulate processing mechanism of cochlear, and modified parameters of Meddis model according to the characteristics of underwater acoustic signals. Based on the Gammatone-Meddis auditory peripheral model, an approach to extract features from underwater target signal was presented. Then a twenty-three dimension feature vectors were obtained. With the feature vectors, the training and recognition processes was performed. The real sea experiments show that the features can classify effectively and the recognition ratio may arrive 94% above, that the features are robust and the capability of recognition does not decline when convolution noises are added to original signals. Finally, the experiments prove that both nonlinear selecting frequency by basement membrane and inner hair cells can restrain noise.

Key words: underwater acoustic signal processing; auditory model; target recognition; feature extraction

1 引言

众所周知, 水中目标分类识别是水声领域的一个难题. 而人的听觉具有优异的声音与声源识别能力, 无论对于人的语音还是自然或人工的各种声源发声, 经过学习, 都能够进行高度准确的识别, 几乎不存在对声源种类和数目的限制. 经过多年研究, 人们对于听觉外周到听觉中枢的整个听觉通路的结构已经获得了比较详尽的知识, 有些甚至到达了细胞层次. 尤其是对于听觉外周, 人们在生理学实验及其应用方面取得了相当可观的成果. 目前, 基于听觉模型的声信号特征提取、分类和识

别主要针对语音信号处理方面^[1,2], 但也有不少研究者将其应用于水中目标分类识别. 马元峰等利用模拟人耳声信号处理过程的 CcGC 滤波器组模型来提取水声信号特征, 并研究了其相对于传统方法的优势^[3]. 王娜等通过心理声学参数对目标进行分类识别^[4]. 彭圆等提出了基于感知线性预测模型来提取水声信号鲁棒特征的方法^[5]. 阳雄等利用短时能量分析与人的主观听觉相结合的方法进行舰船辐射噪声特征提取^[6]. 这些研究成果都表明通过模拟人耳的听觉系统来提取水声信号特征, 都能够获得较好的分类识别效果, 正是因为如此, 本文首先集成可模拟基底膜频率分解特性的 Gammatone^[7,8]

听觉滤波器和模拟内毛细胞换能特性的 Meddis^[9,10]模型来表示人类耳蜗的处理机制,并根据水声信号的特点与实验结果对 Meddis 模型参数进行了修正.提取提出了基于 Gammatone-Meddis 听觉外周模型的水中目标特征提取方法.实验表明通过该方法提取的特征分类识别效果好、抗卷积噪声能力强,对于目标分类识别研究具有一定的参考价值.

2 听觉外周计算模型

听觉外周计算模型,主要是耳蜗模型.它是在耳蜗的解剖学、生理学以及听觉心理学实验研究的基础上,建立适当的模块,用数学的方法实现一套具有类似耳蜗处理功能的算法.与耳蜗的实际工作情况类似,本文采用基底膜和内毛细胞两个模块的顺序结构来模拟耳蜗的处理功能.

2.1 内耳基底膜模型

基底膜是听觉系统对声音信号处理的关键部分.人类的基底膜展开长约 35mm,基底膜的刚度是不均匀的,不同的部位对应不同的共振频率,因而不同频率的声音能够激起基底膜不同部位的振动,从而完成对声音信号的频率分解功能.这里用 Gammatone^[7,8]滤波器组来模拟基底膜的频率分解特性.其冲击响应为:

$$G(t) = t^{n-1} \exp(-2\pi b_m t) \cos(2\pi f_m t + \varphi) U(t) \quad (1)$$

$$1 \leq m \leq M$$

其中: $U(t)$ 为单位阶跃函数, φ 为相位, M 为滤波器通道数, n 为 $G(t)$ 滤波器的阶数, f_m 为滤波器的中心频率,参数 $b_m = b_1 \cdot (24.7 + 0.108 f_m)$, 滤波器中心频率 f_m 在 Bark 域等间距分布,其与 Bark 域频率 Ω 的转换公式^[4]为:

$$\Omega = \begin{cases} f_m/100, & f_m < 500\text{Hz} \\ 9 + 4 \times \lg(f_m/1000), & f_m > 500\text{Hz} \end{cases} \quad (2)$$

Gammatone 滤波器组的通道数与中心频率要考虑到以下几个因素:(1)人耳可听声音的频率范围为 20Hz ~ 20kHz;(2)舰船辐射噪声不同于语音信号,其频率范围从低频到高频;(3)实验室数据采样频率为 25kHz;(4)虽然基底膜的带通滤波器组采用越多模型越精细,但一方面,模型越精细,效果反而不一定好,另一方面,通道数越多,计算的复杂性也会增大.综合以上各方面因素,本文设计的滤波器组中心频率从 50Hz ~ 10500Hz,对应的 Bark 带编号从 1 ~ 23,间隔为 1.

图 1 给出了这 23 个 Gammatone 滤波器在频率上的分布情况.由图可以看出滤波器组中每

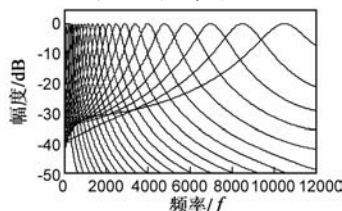


图1 $G(t)$ 滤波器的频率响应

一个滤波器的带宽不同,其带宽取决于滤波器的中心频率.

2.2 修正内毛细胞模型

内毛细胞主要作半波整流、非线性饱和和抑制、短时自适应等反应,把基底膜振动的机械能转化为听神经发放率.听神经发放率是人的听觉外周对噪声和信道干扰进行了抑制,对环境进行了自适应的一种有效的信号表示.它包含了声音信号的多种信息,大脑通过对这些信息进行高级层次的处理获得声音识别.内毛细胞模型有多种,本文采用 Meddis^[9,10]模型来提取听神经发放率:

$$k(t) = \begin{cases} g dt [x(t) + A] / [x(t) + A + B], & x(t) + A \geq 0 \\ 0, & x(t) + A < 0 \end{cases} \quad (3)$$

$$\frac{dq(t)}{dt} = y(1 - q(t)) - k(t)q(t) + xv(t) \quad (4)$$

$$\frac{dc(t)}{dt} = k(t)q(t) - lc(t) - rc(t) \quad (5)$$

$$\frac{dw(t)}{dt} = rc(t) - xv(t) \quad (6)$$

听神经发放率可以表示为:

$$p(t) = hc(t)dt \quad (7)$$

其中 $g, dt, y, x, l, r, h, A, B$ 为常量参数, $x(t)$ 为基底膜的输出.

Meddis 模型的参数是根据已得到的生理实验数据计算出来的,因此直接用其处理舰船辐射噪声信号不符合实际情况.本文通过实验,利用舰船辐射噪声数据来修正 Meddis 模型参数.由于内毛细胞具有半波整流和低通滤波的特性,可以理解在这里发生了包络提取的过程.因此可以对 Meddis 模型的输出即听神经发放率做 DEMON 谱分析.但是,内毛细胞的低通截止频率通常在 1kHz 左右.所以将听神经发放率通过一个截止频率为 30Hz 的低通滤波器,提取其时域包络,在对包络求 FFT 变换到频域,即可获得其 DEMON.最后通过 DEMON 谱的效果来调整该模型的参数.

在 Meddis 模型参数修正实验中,首先用常规调制谱提取方法选取大量不同类别、海况、工况下具有明显调制谱特征的舰船辐射噪声信号.然后利用 Meddis 模型来提取这些信号的调制谱,并通过改变模型参数来调整 DEMON 谱的调制效果.通过多次实验分析发现影响 DEMON 谱质量的主要参数是 A 的值,因此我们主要对 A 值进行了修正,其它参数的取值可参考文献[9, 10].在原 Meddis 模型中 A 取值为 5,但经对大量具有调制谱特征信号的分析发现 A 取值为 5 时信号的 DEMON 谱不够明显,而在 A 取值为 0.01 时,均能得到明显的 DEMON 谱,因此在本文中参数 A 取值为 0.01.

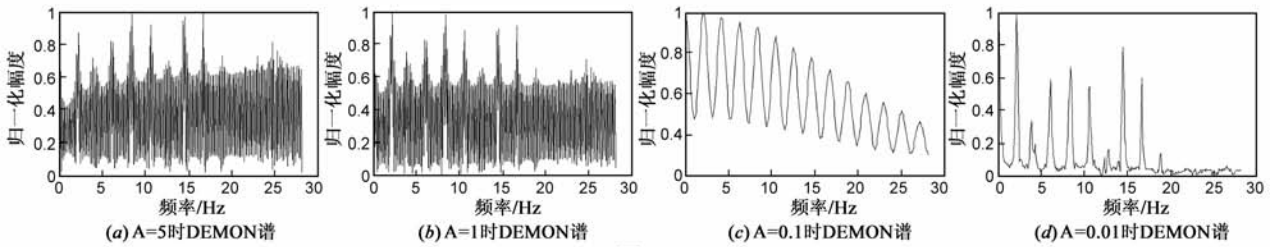


图2

图 2 为实验中一艘船辐射噪声信号分别取 $A = 5$ 、 1 、 0.1 、 0.01 的 DEMON 谱分析结果.从结果可以看出,

当 A 取 0.01 时 DEMON 谱效果最好.图 3 为同一信号

通过传统方法提取的 DEMON 谱,将其与图 2(d)

对比可以看出,通过 Meddis 模型提取的 DEMON 不仅调制幅度得到了加强,

而且相对于传统方法,其调制谱更加丰富,包含了一些

通过传统方法无法检测到的信息.

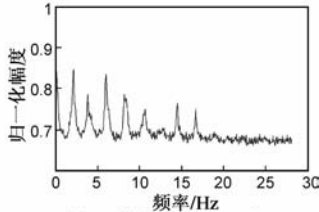


图3 传统DEMON谱

3 特征提取及实验

3.1 特征提取及实验验证

特征提取流程图如图 4 所示,信号进行零均值归一化处理,通过听觉外周计算模型,求每个通道听神经发放率的能量,得到一组 23 维特征向量.

特征有效性实验验证:

本文实验数据均来自真实的水中目标辐射噪声海上实测信号.共 5 类目标,每种类型包括十几种、甚至几十种目标,每个目标有多种工况.每段数据长度为 1s,采样频率为 25kHz,信噪比大于 6dB.当然如果每个样本的时间长度取得更长,则时间处理增益越大,特征的识别效果会更好.图 5 为 B、C 两类目标的特征向量图,可以看出这两类目标的特征向量图明显不同,区分度很好.

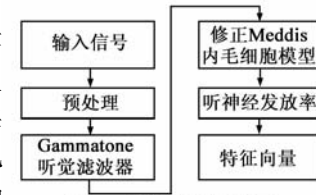


图4 特征提取流程图

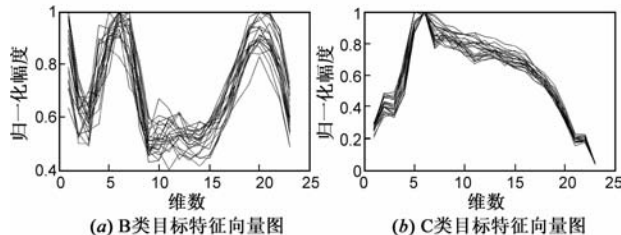


图5

最后使用模糊自适应共振分类器(FART)^[11,12]作为训练和测试网络,将提取的特征输入网络进行验证,总

共 4187 个数据样本,目标信号数据分为训练集和测试集两部分,训练样本与测试样本数量比为 1:6,测试结果如表 1 所示,AVE 表示平均识别率.

表 1 5 类目标识别率

目标类别	测试样本总数	识别结果					识别率
		A	B	C	D	E	
A	300	290	7	3	0	0	96.67%
B	1059	1	1043	14	0	1	98.49%
C	2251	3	24	2201	1	22	97.78%
D	80	0	0	0	77	3	96.25%
E	497	0	1	25	0	471	94.77%
AVE		97.49%					

表 1 反映出对 A、B、C、D 和 E 共 5 类目标,该特征的识别率均到达 94% 以上,基本上不存在特征识别率偏重问题,其平均识别率可以达到 97%.由此可见本文提取的基于听觉外周计算模型的特征是一种可以表征舰船辐射噪声信号的好特征.

3.2 特征抗卷积噪声实验验证

为了验证本文所提特征的抗卷积噪声的能力,在信号中加入了卷积噪声.添加方法如下:将长度为 500 点的白噪声与原始信号做卷积,即可得到加噪信号.图 6 为某一类目标加噪前和加噪后的频谱对比.显然其频谱发生了明显的变化.

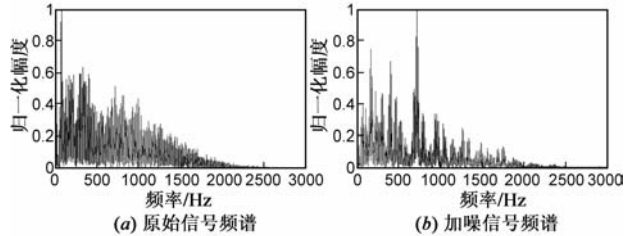


图6

图 7 为加噪信号 B 类目标特征向量图,将其与图 5(a)对比,可以看出加入卷积噪声后,B 类目标的曲线变化趋势基本与未加噪声时保持一致.

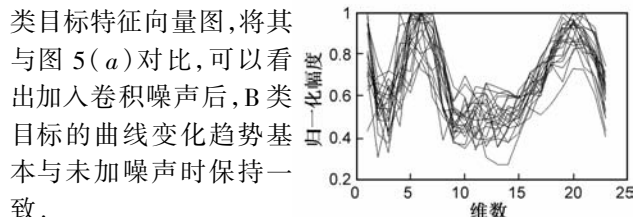


图7 加噪数据B类目标特征向量图

用原始数据提取的

特征作为训练集,加噪数据提取的特征作为测试集,训练样本与测试样本数量比为 1:6,测试结果如表 2 所示.

表 2 5 类目标识别率

目标类别	测试样本总数	识别结果					识别率
		A	B	C	D	E	
A	300	260	25	8	2	5	86.67%
B	1059	31	909	87	8	24	85.84%
C	2251	57	121	2016	10	47	89.56%
D	80	1	1	3	70	5	87.50%
E	497	5	16	77	0	399	80.28%
AVE		87.27%					

从表 2 中可以看出,加入卷积噪声以后,对 5 类目标的识别率仍能够达到 80% 以上,平均识别率为 87%,具有很好的分类识别效果.这说明本文提取的特征能够很好的抑制卷积噪声.

4 特征抑制噪声分析

前面的实验分析结果表明,本文所提取的特征不仅有很好的分类效果,而且具有很强的抗卷积噪声的能力.现在还没法证明听觉模型对噪声的抑制机理,但可以通过实验来分析模型的各个部分对噪声的抑制效果.具体步骤为:对信号提取下面两组特征,(1)将原始数据分别通过 23 个 Gammatone 滤波器,其中心频率在 50 ~ 10500Hz 均匀分布,求每个通道滤波器输出的能量,组成一个 23 维的特征向量,记为 1# 特征.(2)与 1# 特征的提取方法相似,只不过滤波器中心频率在 Brak 域等间距分布,提取的特征记为 2# 特征.本文所提取的基于听觉外周计算模型的特征记为 3# 特征.

分别测试这三组特征对原始数据和加噪数据的识别率,‘○’表示 1# 特征的识别率,‘☆’和‘◇’分别表示 2# 特征和 3# 特征的识别率.结果如 8、9 图所示.

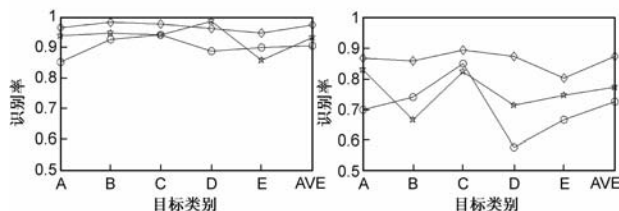


图 8 原始数据 5 类目标识别率对比

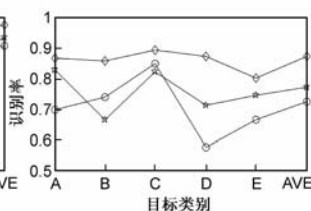


图 9 加噪数据 5 类目标识别率对比

由图 8、9 可以看出,对于原始数据,1# 特征的平均识别率为 90.55%, 2# 特征为 93.38%, 3# 特征为 97.49%,这说明 Gammatone 滤波器对频率的非线性分解与 Meddis 内毛细胞均能够抑制海洋环境噪声.对于加噪数据,1# 特征的平均识别率为 72.57%, 2# 特征为 77.26%, 3# 特征为 87.27%,可以看出,Meddis 内毛细胞模型能够很好的抑制卷积噪声,使分类识别效果得

到很大的提高,基底膜对卷积噪声也有一定的抑制效果.综合上面的分析可以得出如下结论:基底膜对频率的非线性选取和内毛细胞都能够很好的抑制噪声.

5 结论

基于听觉特征的声信号处理是近年来语音信号处理界新兴的一个研究内容,并已取得了较好的结果^[1,2,13].本文通过建立能够模拟人耳听觉外周的计算模型,并根据水声信号的特点与实验结果对该模型参数进行了修正.提出了基于 Gammatone-Meddis 听觉模型的水中目标特征提取方法,提取了一个 23 维的特征向量.实验结果表明,该特征不仅分类识别效果明显,而且对卷积噪声具有较强的鲁棒性.最后通过实验分析得出结论:基底膜对频率的非线性选取和内毛细胞都能够抑制噪声.

参考文献

- [1] 王●,钱志鸿.基于伽马通滤波器组的听觉特征提取算法研究[J].电子学报,2010,38(3):525-528.
WANG Yue, QIAN Zhi-hong. An auditory feature extraction algorithm based on γ -tone filter-banks [J]. Acta Electronica Sinica, 2010, 38(3): 525-528. (in Chinese)
- [2] 付强,易克初.语音信号的 Bark 子波变换及其在语音识别中的应用[J].电子学报,2000,28(10):102-105.
FU Qiang, YI Ke-chu. Bark wavelet transform of speech and its application in speech recognition [J]. Acta Electronica Sinica, 2000, 28(10): 102-105. (in Chinese)
- [3] 马元峰,陈克安.听觉模型输出谱特征在声目标识别中的应用[J].声学学报,2009,34(2):142-150.
MA Yuan-feng, CHEN Ke-an. Application of auditory spectrum-based features into acoustic target recognition [J]. Acta Acustica, 2009, 34(2): 142-150. (in Chinese)
- [4] 王娜.基于人耳主观反应的听觉特征量及其在目标识别中的应用[D].陕西西安:西北工业大学,2006.
- [5] 彭圆,王晟.感知线性预测在水下目标分类中的应用研究[J].声学学报,2006,31(2):146-150.
PENG Yuan, WANG Sheng. A study on underwater target classification applying perception linear prediction method [J]. Acta Acustica, 2006, 31(2): 146-150. (in Chinese)
- [6] 阳雄,程玉胜.短时能量分析及人耳的主观听觉在船舶辐射噪声特征提取中的研究[J].声学技术,2004,23(1):11-13.
YANG Xiong, CHENG Yu-sheng. Extraction of ship-radiated noise characteristics based on analysis of short-term energy and subjective hearing [J]. Technical Acustica, 2004, 23(1): 11-13. (in Chinese)
- [7] Patterson R, Holdsworth J. An efficient auditory filterbank based on the gammatone functions [R]. Annex B of the Svos

Final Report: The auditory filter bank, APU Report No. 2341. 1988.

- [8] Slaney M. An efficient implementation of the Patterson-Holdsworth auditory filterbank[R]. Apple Computer Technical Report # 35, 1993.
- [9] Meddis R. Simulations of mechanical to neural transduction in the auditory receptor[J]. Journal of the Acoustical Society of America, 1986, 79(3): 702 – 711.
- [10] Meddis R. Implementation details of a computer model of the inner hair-cell/auditory-nerve synapse[J]. Journal of the Acoustical Society of America, 1990, 87(4): 1813 – 1818.
- [11] Filip Mulier. Vapnik-Chervonenkis (VC) learning theory and its application[J]. IEEE Transactions on Neural Networks, 1999, 10(5): 340 – 356.
- [12] Vapnik V. The Nature of Statistical Learning Theory[M]. New York: Springer-Verlag, 1995
- [13] 赵鹤鸣, 葛良. 基于声音定位和听觉掩蔽效应的语音分离研究[J]. 电子学报, 2005, 33(1): 158 – 160.
ZHAO He-ming, GE Liang. Speech separation based on sound

localization and auditory masking effect[J]. Acta Electronica Sinica, 2005, 33(1): 158 – 160. (in Chinese)

作者简介



王 磊 男, 1984 年出生于陕西渭南, 硕士, 从事水声信号与信息处理方向研究.
E-mail: wanglei721127@yahoo.com.cn



彭 圆 女, 1972 年出生于黑龙江, 高级工程师, 博士, 长期从事模式识别、信号与信息处理方向研究. E-mail: Pengyuan067@tom.com