基于听觉感知机理的水下目标识别研究进展

徐新洲，罗昕炜，方世良，赵 力
(东南大学水声信号处理教育部重点实验室，南京 210096)

摘要：水下目标特征提取与识别技术在国防、生产等领域内有着迫切的需求，目前基于听觉感知机理的水下目标识别研究正处于起步阶段，有着较为广阔的研究与应用前景。在简要介绍听觉感知机理研究情况的基础上，对近年来国内外使用听觉感知机理进行水下目标识别的各类研究成果进行了总结与归纳，针对基于听觉感知机理的水下目标识别领域各方面的研究进展给出了相应的客观评价，并从特征提取与目标的自动识别方面提出了对未来研究方向的展望。

关键词：听觉感知；听觉模型；舰船辐射噪声；目标识别
中图分类号：TN929.3 **文献标识码**：A **文章编号**：1000-3630(2013)-02-0151-08
DOI 编码：10.3969/j.issn1000-3630.2013.02.016

Research progress of underwater target recognition based on auditory perception mechanism

XU Xin-zhou, LUO Xin-wei, FANG Shi-liang, ZHAO Li
(*Key Laboratory of Underwater Acoustic signal Processing of Ministry of Education*, *Southeast University, Nanjing* 210096, *China*)

Abstract: Feature extraction and recognition of underwater targets are of urgent needs in various fields. Currently, the research on the auditory perception based underwater target recognition is just on the initial stage. Based on description of the research condition of auditory perception mechanism, a summary and introduction of many kinds of researches on underwater target recognition in recent years both at home and abroad, as well as related objective evaluations in each aspect, are narrated. The outlook for future researches is also proposed from the view of feature extraction and automatic object recognition.

Key words: auditory perception; auditory model; ship radiated noise; target recognition

0 引 言

水下目标识别领域的研究已有较长时间了，其最主要的目的是针对军事用途的舰船目标，其次实现对海洋生物的监测。传统的水下目标识别方法主要靠声纳员及专门的仪器辅助下进行的，但很多目标识别算法还远远达不到实用的要求，需要进行深入的研究和改进。

由于军事用途上在探测目标的同时也要进行自身的隐蔽，所以很多情况下采用被动声纳。在使用被动声纳识别舰船目标方面[1,2]，针对的舰船辐射噪声[3,4]主要由机械噪声、螺旋桨噪声、水动力噪声等组成。但一般情况下舰船辐射噪声以机械噪声和螺旋桨噪声为主，从水下设备的类型看，潜艇、水面舰船、鱼雷等的噪声具有各自的特点，其中水面舰船的螺旋桨噪声在其辐射噪声中所占比例较小，潜艇、鱼雷的辐射噪声主要来源于螺旋桨叶片转动产生的空泡[3]。可见水下目标识别所针对的对象种类较多，加上水下环境的复杂性，这些都给识别工作带来了不小的困难。

传统的自动识别方法包括 LOFAR 图、DEMON 分析及对这些方法的改进算法等，但这些都相对忽略了声纳员的听觉感知能力在声纳目标识别中的作用。而且，培养一个经验丰富的声纳员相对困难。鉴于此，很多学者尝试使用听觉感知机理及听觉模型对水下目标识别进行深入的研究。在听觉机理下，可以较容易地分辨声信号的主观描述特征，但这与舰船目标类型以及声源特性间的联系仍有待研究，所使用的算法也需大量实验的支撑。

本文首先对听觉感知机理及听觉系统建模发展情况进行了简要的介绍，在此基础上介绍了听觉感知机理及听觉模型用于水下声信号目标识别的发展历程与研究现状，以及其他一些语音信号处理方法用于水下目标识别的研究情况，同时叙述了国

收稿日期：2012-03-09；**修回日期**：2012-06-23
基金项目：国家自然科学基金(60872073, 60975017, 51075068)，教育部博士点专项基金(20110092130004)资助项目
作者简介：徐新洲(1987－)，男，南京人，博士研究生，研究方向为水下目标识别、模式识别与机器学习。
通讯作者：徐新洲，E-mail: radar256@163.com

内在这方面的研究现状，对各类方法进行了归纳，最后给出了未来研究的展望。

1 听觉感知及听觉系统建模

人的听觉感知系统具有时频分析、听觉掩蔽等特性[5,6]，本文所述的听觉系统主要指听觉外周系统。听觉外周系统主要由外耳、中耳和内耳组成[5]。外耳起着声源定位和声音放大的作用；中耳的作用是进行声阻抗的变换；内耳主要由耳蜗构成，其作用是将接收到的声信号经过处理传输到听觉中枢。听觉中枢主要由耳蜗核、上橄榄复合体、下丘、内侧膝状体和听觉皮层等构成[1]。

人们对于听觉外周的研究[1]主要集中于耳蜗。耳蜗对声音的感知和处理可大致由三个模块表示，即：基底膜对声信号的频率分解、内毛细胞(IHC)及听神经的能量转换、听神经的发放作用。

早期，Békésy 通过频闪观测仪，发现了耳蜗基底膜上的行波及基底膜的频率分解作用，并据此建立了最早的耳蜗一维传输模型[7]。Johnstone 等[8-10]采用 Mossbauer 技术对耳蜗中的基底膜振动进行了测量研究，得到了比 Békésy 相对更为精确的实验结果。而 Zwislocki[11]则建立了一维传输线模型来解释 Békésy 的实验结果。Rhode[12,13]发现了基底膜振动用一维模型无法解释的非线性特征，为耳蜗的研究开辟了一个新领域。

耳蜗模型研究的深入使得对模型的构造有了更复杂的要求，文献[14,15]中建立了耳蜗的二维模型。早期基底膜模型中[1]，Lyon 被动耳蜗模型[16,17]、Seneff 模型[18,19]得到了较多的关注，而 Yang 等[20]则提出了基于小波变换的基底膜模型。此后，Lyon 和 Mead[21]对 Lyon 模型进行了改进，Kates[22]建立了时域耳蜗模型。de Bore 等[23,24]建立的 GT (gammatone)滤波器模型对基底膜建模发展有着重要的意义，Carney 等[25]在此基础上做了改进。Irino 等[26,27]在 GT 滤波器基础上使用了 gammachirp 滤波器，文献[28,29]对此进行了进一步论述。

对内毛细胞的建模在听觉系统建模中同样起着重要的作用，常用的内毛细胞模型有 Meddis 等[30,31]提出的蓄水池模型、Westerman 等[32]的化学扩散模型等。Robert 等[33]对 Meddis 的模型进行了修正。Seneff[34]和 Martens[35]采用更为简单的半波整流、压缩非线性和低通滤波来对内毛细胞进行建模。内毛细胞模型的其他相关研究仅在 Meddis 等人的基础上进行的改进，近几年来研究成果较少[1]。

Yang、Shamma 等在文献[20,36-39]中提出主要由频率分解、内毛细胞能量转换、侧抑制神经网络三个模块所构成的模型成为目前较为常用的一种听觉系统模型，该模型在语音信号处理中得到了较为成功的应用[6]，模型的具体算法流程如图 1 所示。

Yang 等的模型主要由五个步骤[6,20,38-40]组成，分别模拟听觉系统中的耳蜗频率分解、内毛细胞能量转换、侧抑制神经网络的信号增强、半波整流、中枢神经短时积分：

$$\begin{cases} y_1(t,s)=h(t,s)*_t x(t) \\ y_2(t,s)=g(\dfrac{\partial y_1(t,s)}{\partial t})*_t w(t) \\ y_3(t,s)=\dfrac{\partial g(\dfrac{\partial y_1(t,s)}{\partial t})}{\partial s}*_t w(t)*_s v(s) \\ y_4(t,s)=\max(y_3(r,s),0) \\ y_5(t,s)=\dfrac{1}{T}\int_{t-T}^{t} y_4(r,s),\mathrm{d}r \end{cases} \tag{1}$$

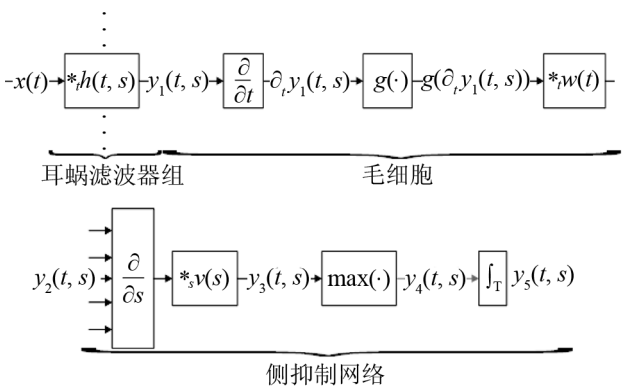

图 1 Yang 等提出的算法流程示意图

Fig.1 The diagram of the algorithm proposed by Yang et al.

式(1)中，$x(t)$ 为初始输入信号，$h(t,s)$ 表示基底膜 s 处的传输函数，$w(t)$ 表示过毛细胞膜的 LPF，$v(s)$ 表示频域平滑，$g(\cdot)$ 表示非线性映射。目前该模型较为广泛地应用于语音信号处理中。

2 基于听觉感知机理的水下目标识别研究进展

2.1 基于听觉感知的水下目标识别

听觉模型在语音识别等语音信号处理工作中已得到广泛的研究，而在水下目标识别中的研究还比较少。李朝晖等[1]对近 30 年听觉外周计算模型的研究进行了评述，并讨论了在被动声纳目标识别中应用听觉模型的可能性，同时结合声纳目标识别的特点，对听觉外周模型的适用性建模研究提出了若干意见。

很多学者对于一般的声信号如语音、音乐等进行了分析研究，这些方法在声信号共性问题上给水声信号的研究提供了一定程度的参考。Todd 等[41,42]发现 rhythmogram 方法可以识别出音乐、语音信号中的层次结构，并利用基于听觉的、多尺度的 rhythmogram 方法分析声信号时域；McAdams[43]对音乐音质特征的提取进行了研究；Gaumond 等[44]通过人耳实验研究了主观听觉对脉冲声信号的分类效果；McLaughlin 等[45]同样分析了听觉对声信号分类的有效性；Lotto 等[46]在听觉分类方面叙述了听觉的特殊性质，指出学习和注意力在听觉中起到了较为重要的作用。

使用听觉模型对声信号，尤其是水声信号进行处理始于 Teolis、Shamma 等的研究，Teolis 等在文献[47]中使用文献[20,38,39]提出的听觉模型对 DARPA/ONR 提供的标准瞬态信号数据进行分类，发现由听觉模型提取出的特征相对于 STFT 所得特征更有利于分类，这为后续的研究工作提供了很大帮助。Parks 等[48]则研究了使用 Lyon 耳蜗模型对瞬态声信号进行分类，实验样本为鲸鱼和海冰的声信号。相比于 STFT、小波变换等，耳蜗模型更具有优势。

在前人使用听觉模型及听觉感知机理处理声信号，以及大量主观实验、算法研究的基础上，Tucker 等提出了一种全新的基于听觉感知机理的水下目标识别算法。

Tucker、Brown 等[49-54]在这一方面做了较多的研究工作，评估了用听觉方法分析水声信号的可行性，并根据研究成果尝试进行水下目标识别。文献[49,50]研究了听觉机制对声源大小、形状、材质等物理特性的感知。文献[51]研究了音质对于使用听觉区分声纳信号的重要性，在已经证明听觉可用于识别瞬态信号的基础上进一步用实验证实了听觉机理用于识别声纳信号的可能性。文献[52,53]在之前研究基础上提出了综合音质、材质、时域多尺度特征的瞬态声纳信号特征提取、分类方法，利用了 Parcel 等算法进行进一步的特征降维，算法总体流程如图 2 所示，图 2(a)表示特征提取及训练过程，图 2(b)表示识别分类过程。

有些情况，如当舰船目标辐射强度过小或采取了消噪措施时，被动声纳检测效果并不理想，所以利用主动声纳进行目标识别具有一定的研究前景。Young、Allen、Myers、Hines、Murphy 等[55-61]研究了听觉模型对声纳目标，尤其是使用主动声纳时的分类识别。文献[55]测试了人类听觉系统对于主动

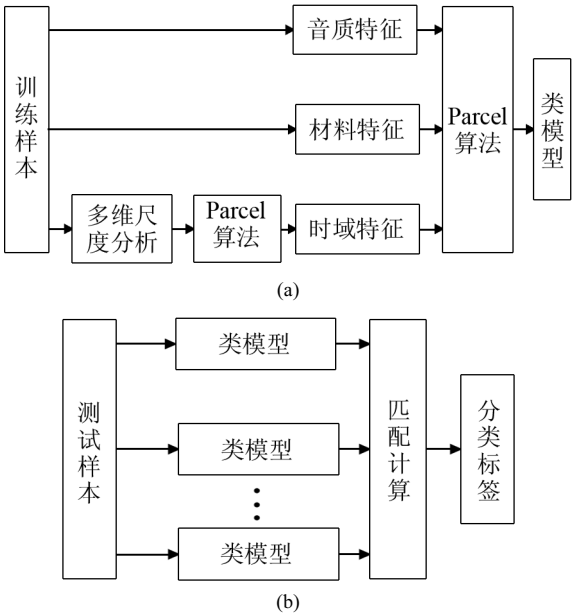

图 2　Tucker 等提出的水下目标识别算法框架[60]

Fig.2　The framework of the underwater object-recognition algorithm proposed by Tucker et al[60]

声纳回声的分类效果。文献[56]使用听觉模型对脉冲源主动声纳回声进行分类，由时频域的包络特征、响度特征等音色特征区分目标回声和混响。文献[57]研究了相干源回声的分类。文献[58]和文献[56,57]在分别研究了脉冲源、相干源声纳信号用听觉模型进行分类识别的基础上，进一步研究了声纳带宽对于特征的影响。文献[59]使用文献[56]的方法提取音色特征，并进行进一步的特征选择提取。文献[60,61]同样研究了人类听觉系统区分声纳目标回声和混响的能力，研究表明听觉系统足以区分这些信号，但当信号的低频信息移除时性能会显著下降，而这时机器听觉分类器相对于人耳性能有明显的回升，这证明了文献[56]方法的有效性。

与此同时，其他学者也对使用听觉模型进行水下目标特征提取和识别进行了研究工作，取得了一些较为重要的成果。Mill 等[62]通过听觉特性提取声纳信号的音调特征，采用了内毛细胞及侧抑制神经网络的模型架构。Lutfi 等[63]研究了人耳对撞击声的识别，并研究了当撞击声生成模型参数变化时不同类型测试者的情况。Sildam[64]则是通过计算熵及熵估计来掩蔽混响，提取出感兴趣的目标信号。

此外，一些学者使用语音信号处理中的一些方法[5,6]解决水下目标特征提取与识别问题。Lim 等[65]使用 MFCC 逐帧提取特征，并进行分类识别。Tolba 等[66]比较了水下目标识别中的 CHMM 和 ANN 方法的识别效果差异，在具体的算法中，提出了信号预处理、基于 MFCC 的特征提取、基于 CHMM 的模

式匹配的流程，取得了较好的效果。Xie 等[60]用 STFT 对海洋生物进行识别，其中使用了多种模型来提取参数。

2.2 国内研究现状

国内对听觉感知机理及听觉模型在水下目标特征提取和识别中应用的研究还处于起步阶段，当前已经取得了一些成果。

在将听觉模型用于一般声信号处理识别方面，吴玺宏、迟惠生[68-70]着重研究了听觉计算模型在语音信号处理识别中的应用。张焱等[71]在国内较早地使用了 Yang 和 Shamma 的听觉模型[20,38,39]解决语音识别问题。

在使用常用的声信号处理方法进行水声目标识别，尤其是舰船目标识别方面，蔡悦斌等[72]分析了舰船噪声信号的时域信息，将波形分类表达为 11 维的分类特征向量，并用实验数据证明了其分类的有效性。陆振波等[73]将 MFCC 特征参数应用于舰船辐射噪声的特征提取，分析了 MFCC 在水声信号特征提取中的有效性。

对于听觉感知机理与水下目标识别结合的研究也得到了越来越多的重视。陆振波等[74]在临界带分析、等响度预加重-响度变换、用 AR 模型及 DCT 进行特征参数表示这三个层次上模拟人耳处理机制解决水中目标辐射噪声的特征提取问题。阳雄等[75]提出利用短时能量和节奏提取舰船辐射噪声的特征。吴岳松等[40]利用人耳的听觉模型，提取了舰船辐射噪声的听觉特征。彭圆等[76]提出了一种基于感知线性预测、模拟人耳听觉特性来提取水声信号特征的方法。杨宏晖等[77]利用耳蜗分频、听觉掩蔽等特性，提出了步骤为计算功率谱、临界带分析、等响度级预处理、等响度转换、集成 SVM 分类的算法。文献[78]将听觉耳蜗模型应用于舰船噪声信号的分析中，分析了舰船噪声的时频幅特征。

在水声条件下听觉模型的具体模块设计方面，马元锋、陈克安等[79-81]对使用听觉模型进行声目标识别(包括了水声目标的识别)做了研究。文献[79]使用 CcGC 滤波器组模型，模拟耳蜗的抑噪机理，指出听觉系统的临界带划分、非线性压缩在抑制噪声中起着关键的作用。文献[80]提出了自适应的 EWD 模型作为听觉时频分析模型。文献[81]中使用 Moore 模型来模拟听觉的掩蔽特性，并进行了海洋生物的声信号识别。

在主观听觉特性在水声信号感知中的适用性方面，王娜、陈克安等[82-88]用和 Tucker 等相似的方法重点研究了主观听觉特性对水声信号的感知特性，发现声音的沉杂性、起伏性、快变性、尖锐性、规律性在水下噪声主观识别中起着较为重要的作用，并同样对音色等特征进行提取，进而实现水下目标识别。

2.3 各类算法研究现状

2.3.1 基于听觉模型的算法

听觉模型是对听觉系统的近似逼近，建立合适的听觉模型对于声信号的处理有着重要的意义，以更好地模拟人耳的功能。听觉模型的整体结构在一般声信号处理中通常采用 Yang、Shamma[20,38,39]的结构，或与其类似的模块结构，目前很多研究也是基于此。

Teolis 等[47]提出了使用听觉模型解决瞬态信号的分类，证明了听觉模型应用在瞬态信号处理中能够取得较好的效果，但模型中很多参数的确定仍有待研究。

Mill[62]、吴岳松[40]较为详细地给出了使用听觉模型解决水下目标识别问题的方案，并进行了实验验证，但缺乏与其他算法性能的比较。

一些研究集中于听觉模型的各模块，目前对于基底膜的仿真研究相对较多，而其他的模块并没有得到太多的关注。

Parks 等[48]的早期研究工作主要针对耳蜗滤波器模型与其他频率分解滤波器组的比较。

马元锋、陈克安等[79-81]的研究侧重于对基底膜滤波器组的设计，以实现听觉系统的非线性、掩蔽特性等。杨宏晖[77]、谢骏[78]等也对适用于水声信号的听觉模型基底膜进行了设计。

用于水下目标识别的听觉模型设计目前都是遵循于对几个常用模块进行研究，并将各模块组合到一起，但对于模块间的耦合作用，以及新模块的设计并没有太多的涉及。

2.3.2 用于水声环境的听觉主观参数研究

信号主观特性及其描述、参数的研究可以为后续的算法设计等工作提供主观条件下的依据，也可以反映听觉机理对于信号某种特点上的关注。

Tucker、王娜等对使用听觉主观特性进行水下目标识别做了大量的研究，这方面的研究着重于发现对于不同类型、不同特性的目标识别有帮助的主观心理学特征。

Tucker、Brown 等的研究[49-54]表明，表征声源的物理材料特性、信号的音质等参数在水下目标识别中同样适用，并且能够较好地反映主观特性。

王娜、陈克安等[82-88]在水声信号主观特性的描

述方面进行了大量的研究，给出了在水声信号识别中起到重要作用的主观特性，通过这些主观特性确定一系列参数用于信号识别。

Young、Allen、Myers、Hines、Murphy 等[55-61]的研究工作则主要集中于对声纳接收信号音质音色方面的分析。

虽然目前的研究已经得到了很多有利于分类的特性参数，但很多情况下仍不足以对成分更为复杂的水下目标样本进行识别。所以，针对不同的样本找出更多的分类特性及将各类特性联系起来进行研究是必要的。

2.3.3 基于传统语音信号处理方法的算法

传统的语音识别，尤其是对说话人识别中，常用的算法[65]包括 MFCC、VQ、HMM、GMM，以及一些后续的分类识别方法等，从另一个角度看，很多算法也可以视作是对听觉系统在某种程度上的近似模仿。而且，声纳员根据长期经验，能够对水下目标信号进行有效的识别，也为此提供了充分的研究依据。

Lim[65]、Tolba[66]、陆振波等[73]采用模拟人耳听觉的 MFCC 方法处理声纳信号，提取出有效的特征参数进行潜艇等水下目标的识别。

Tolba 等[66]还使用了 CHMM 算法，对 MFCC 得到的各样本特征量进行后续的样本分类识别工作，通过实验仿真说明了其有效性。

语音信号常用的处理方法已被证明能够用于水下目标识别，但声纳信号所针对的目标特性与语音信号相比有着一定程度的区别，这有赖于主观参数方面的研究。

2.3.4 基于其他一些方法的算法

很多学者并不满足于已有的算法，而是提出新的算法，或将已有算法进行融合，以得到更适用于水声信号的方法。利用听觉感知机理解决水下目标识别方面的研究工作目前还相对较少，所以很多的算法还有待于实验的验证。

Tucker 等提出的算法[52,53]分别通过计算提取信号的音质特征、材料特征和时域特征，将多个算法集成起来，并使用模糊 KNN 对测试样本集进行分类判决。其中，通过计算统计参数等获取音质特征。将信号分帧通过耳蜗频率选择滤波器，时间平均得到材料特征。通过分级平滑、峰值间隔检测等得到 Todd 等[41,42]所使用的 rhythmogram 特征，处理后得到时域特征。Xie[67]、蔡悦斌[72]、陆振波[74]、杨宏晖等[77]也分别提出了利用听觉感知特性解决水下目标识别的方法。

目前对于基于听觉感知机理的水下目标特征提取及识别的研究还处于探索阶段，主要借鉴于其他声信号识别的研究成果。现有的各种方法都有各自的特点，还没有一种公认的主流方法，这无疑给研究工作的开展增加了难度，但也提供了更多提出合理假设猜想的空间，可以对各种不同方案的合理性进行论证。

3　未来研究方向展望

水下目标识别的研究目前主要集中于几种经典方法，而听觉感知机理以及听觉模型的引入为水下目标识别研究提供了一种全新的思路。

在 Yang、Shamma 等提出的听觉模型[20,38,39]基础上发展起来的由耳蜗频率分解模型、内毛细胞能量转换模型、侧抑制神经模型听觉三个主要模块以及后续处理所组成的模型结构目前得到了较为广泛的使用，尤其是在语音信号处理领域内。模型在模拟人耳听觉机理的同时用解析的方法将信号处理过程表示出来，能够较好地实现人耳听觉系统的仿真。但由于水下声场与空气中语音信号的声场具有差异性，所以适用于水声信号处理的耳蜗频率分解模型、能量转换模型与侧抑制神经网络模型的设计等还有待更深一步的研究探索，包括非线性特性、自适应性等。

Tucker 等[49-54]提出的利用听觉感知参数进行水下目标识别具有较好的性能，但主观参数的选取方式会在很大程度上影响识别的性能。同时，大量算法的集成对理论方面的分析提出了更高的要求。

听觉感知机理及听觉模型应用的现状并不是很乐观。针对不同应用背景，模型利用的数据和反映的听觉特性也各有侧重，很少出现完整的听觉外周模型。另外，在实际应用中，真正被识别系统使用或者在一些系统上被测试过的模型更少。一些采用听觉模型前端的语音识别系统，获得的性能改善十分有限，且适应性较差。因此，理想的听觉模型还没有实现。而且现有的语音信号处理理论已经提出了很多有效的方法[5,6]，这都对听觉模型的发展起到了不利的影响，但这并不影响将听觉模型与其他方法结合起来，相信可以取得相辅相成的效果。

对于水下信号处理识别，由于水下环境的复杂性，相对于语音信号，声纳信号呈现出较多的不确定性，现有的听觉模型不一定适用，需要在模型的前、后端进行附加的适应性处理，或改进现有听觉模型、提出新的模型等。此外，将听觉模型同其他

模式识别方法[89,90]相结合也为水下目标识别提供了一个新的解决途径。

为了实现听觉模型及听觉感知机理在水下目标识别，尤其是舰船目标辐射噪声识别中的应用，还需在以下几方面进行深入的研究：

(1) 耳蜗基底膜滤波器组的非线性[1]、自适应特性设计的研究工作，以及对侧抑制神经网络的复杂性建模在听觉模型的设计中较为关键。

(2) 对于水下目标的听觉主观描述特性的选择以及建模，将是完善使用听觉感知机理提取水下目标信号特征的重要途径。各类舰船辐射噪声信号以及水下各类干扰噪声、混响等都具有各自的声学特性，对于如何将各自的特性用听觉模型展现出来，在水声信号特性的研究方面有着较为重要的意义，同时也为识别工作提供了依据。

(3) 在使用听觉模型提取特征及识别之前需要进行必要的水声信号的预处理工作。

(4) 针对声纳信号的特点，结合合理的分类器，在建立的听觉模型上进行目标识别评估。其中，可以引入 PCA、LDA、ANN、SVM 等经典的模式识别和机器学习方法，也可以使用最新提出的一些方法，而语音信号处理中常用的 HMM、MFCC 等方法也同样可以使用，应展开深入的研究。

4 结 论

本文简要叙述了听觉系统及其建模的一些研究成果，并对使用听觉感知机理及听觉模型进行水下目标识别的研究进展进行了总结归纳。在此基础上，对基于听觉感知机理及听觉模型的水下目标识别，尤其是对采用听觉感知机理进行舰船目标辐射噪声信号的特征提取与舰船目标识别进行了展望。

参 考 文 献

[1] 李朝晖, 迟惠生. 听觉外周计算模型研究进展[J]. 声学学报, 2006; **31**(5): 449-465.
LI Zhaohui, CHI Huisheng. Progress in computational modeling of auditory periphery[J]. Acta Acustica, 2006, **31**(5): 449-465.

[2] 丁玉薇. 被动声纳目标识别技术的现状与发展[J]. 声学技术，2004，**23**(4)：253-256.
DING Yuwei. Review on passive sonar target recognition[J]. Technical Acoustics, 2004, **23**(4): 253-256.

[3] 李启虎. 声纳信号处理引论[M]. 北京: 海洋出版社, 1985.
LI Qihu. Sonar signal processing introduction[M]. Beijing: Ocean Press, 1985.

[4] Waite A. Sonar for practising engineers[M]. 3rd Edition. John Wiley&Sons, Ltd., 2002.

[5] 赵力. 语音信号处理[M]. 第二版. 北京: 机械工业出版社, 2009.
ZHAO Li. Speech signal processing[M]. 2nd Edition. Beijing: China Machine Press, 2009.

[6] Quatieri T. Discrete-time speech signal processing: principles and practice[M]. Prentice Hall PTR, 2002.

[7] Békésy G. On the elasticity of the cochlear partition[J]. J. Acoust. Soc. Am., 1948, **22**(3): 227-241.

[8] Johnstone B, Taylor K, Boryle A. Mechanics of the guinea pig cochlea[J]. J Acoust Soc Am, 1970, **47**(2B): 504-509.

[9] Johnstone B, Boryle A. Basilar membrane vibrations examined with the Mossbauer technique[J]. Science, 1987, **158**(3799): 390-391.

[10] Johnstone B, Yates G. Basilar membrane tuning curves in the guinea pig[J]. J Acoust Soc Am, 1974, **55**(3): 584-587.

[11] Zwislocki J. Five decades of research on cochlear mechanics[J]. J. Acoust. Soc. Am., 1980, **67**(5): 1679-1685.

[12] Rhode W, Robles L. Evidence from Mössbauer experiments for nonlinear vibration in the cochlea[J]. J. Acoust. Soc. Am., 1974, **55**(3): 588-596.

[13] Robles L, Rhode W, Geisler C. Transient response of the basilar membrane measured in squirrel monkeys using the Mössbauer effect[J]. J. Acoust. Soc. Am., 1976, **59**(4): 926-939.

[14] Allen J. Two-dimensional cochlea fluid model: New results[J]. J. Acoust. Soc. Am., 1977, **61**(1): 110-119.

[15] Lien M. A mathematical model of the mechanics of the cochlea[D]. Ph. D. dissertation. Sever Institute of Technology. Washington University: St. Louis, Mo., 1973.

[16] Lyon R. A computational model of filtering, detection, and compression in the cochlea[C]// ICASSP, 1982: 1282-1285.

[17] Lyon R. Computational models of neural auditory processing[C]// ICASSP, 1984: 36.1.1-36.1.4.

[18] Seneff S. A computational model for the peripheral auditory system: Application to speech recognition research[C]// ICASSP, 1986: 1983-1986.

[19] Seneff S. A joint synchrony/mean-rate model of auditory speech processing[J]. Journal of Phonetics, 1988, **16**(1): 55-76.

[20] Yang X, Wang K, Shamma S. Auditory representations of acoustic signals[J]. IEEE Trans. Inform. Theory, 1992, **38**(2): 824-839.

[21] Lyon R, Mead C. An analog electronic cochlea[J]. IEEE Trans. on Acoustic, Speech and Signal Processing, 1988, **36**(7): 1119-1134.

[22] Kates J. A time-domain digital cochlear model[J]. IEEE Trans. on Signal Processing, 1991, **39**(12): 2573-2592.

[23] E. de Bore. Synthetic whole-nerve action potentials for the cat[J]. J. Acoust. Soc. Am., 1975, **58**(5): 1030-1045.

[24] E. de Bore, H. de Jongh. On cochlear encoding: Potentialities and limitations of the reverse correlation technique[J]. J Acoust Soc Am, 1978, **63**(1): 115-135.

[25] Carney L. A model for the responses of low-frequency auditory-nerve fibers in cat[J]. J. Acoust. Soc. Am., 1993, **93**(1): 401-417.

[26] Irino T, Patterson R. A time-domain, level-dependent auditory filter: The gammachirp[J]. J. Acoust. Soc. Am., 1997, **101**(1): 412-419.

[27] Irino T, Patterson R. A compressive gammachirp auditory filter for both physiological and psychophysical data[J]. J. Acoust. Soc. Am., 2001, **109**(5): 2008-2022.

[28] Patterson R, Unoki M, Irino T. Extending the domain of center frequencies for the compressive gammachirp auditory filter[J]. J. Acoust. Soc. Am., 2003, **114**(3): 1529-1542.

[29] Unoki M, Irino T, Glasberg B, et al. Comparison of the roex and gammachirp filters as representations of the auditory filter[J]. J. Acoust. Soc. Am., 2006, **120**(3): 1474-1492.

[30] Meddis R. Simulations of mechanical to neural transduction in the auditory receptor[J]. J. Acoust. Soc. Am., 1986, **79**(3): 702-711.

[31] Meddis R. Simulations of auditory-neural transductions: Further studies[J]. J. Acoust. Soc. Am., 1988, **83**(3): 1056-1063.

[32] Westerman L, Smith R. A diffusion model of the transient response of the cochlear inner hair cell synapse[J]. J. Acoust. Soc. Am., 1988, **83**(6): 2266-2276.

[33] Robert A, Eriksson J. A composite model of the auditory periphery for simulating response to complex sounds[J]. J. Acoust. Soc. Am., 1999, **106**(4): 1852-1864.

[34] Seneff S. A computational model for the peripheral auditory system: Application to speech recognition research[C]// ICASSP, 1986: 1983-1986.

[35] Martens J, Immerseel L. An auditory model based on the analysis of envelope patterns[C]// ICASSP, 1990: 401-404.

[36] Shamma S. Speech processing in the auditory system. Part I: The Representation of Speech Sounds in the Responses of the Auditory-Nerve[J]. J. Acoust. Soc. Am., 1985, **78**(5): 1612-1621.

[37] Shamma S. Speech processing in the auditory system. Part II: Lateral Inhibition and the Central Processing of Speech Evoked Activity in the Auditory-Nerve[J]. J. Acoust. Soc. Am., 1985, **78**(5): 1622-1632.

[38] Shamma S, Chadwick R, Wilbur J. A biophysical model of cochlea processing: Intensity dependence of pure tone responses[J]. J. Acoust. Soc. Am., 1986, **80**(1): 133-145.

[39] Yang X, Wang K, Shamma S. Auditory representations of acoustic signals[R]. Tech. Rep., University of MD, Systems Research Center, 1991.

[40] 吴岳松. 基于听觉模型的水下目标识别研究[D]. 西北工业大学硕士学位论文, 2005.
WU Yuesong. The researches on underwater targets recognition based on the auditory model[D]. MA. D. dissertation. Northwestern Polytechnical University, 2005.

[41] Todd N, Brown G. Visualization of rhythm, time and metre[J]. Artificial Intelligence Review, 1996, **10**(3-4): 253-273.

[42] Todd N. The auditory"Primal Sketch": A multiscale model of rhythmic grouping[J]. Journal of New Music Research, 1994, **23**(1): 25-70.

[43] McAdams S. Perspectives on the contribution of timbre to musical structure[J]. Computer Music Journal, 1999, **23**(3): 85-102.

[44] Gaumond C, Brock D. Aural classification of impulsive sounds[C]// Proceedings of SYMPOL, 2009.

[45] McLaughlin J, Philips S, Pitton J. Perceptually-driven signal analysis for acoustic event classification[R]. Final Report, OMB No. 0704-0188, 2007.

[46] Lotto A, Holt L. Psychology of auditory perception[J]. Wiley Interdisciplinary Reviews: Cognitive Science, 2011, **2**(5): 479-489.

[47] Teolis A, S Shamma. Classification of the transient signals via auditory representations[R]. Tech. Rep. TR91-99, 1991.

[48] Parks T, Weisburn B. Classification of whale and ice sounds with a cochlear model[C]// ICASSP, San Francisco, CA, 1992, **2**: 481-484.

[49] Tucker S, Brown G. Modelling the auditory perception of size, shape and material: Applications to the classification of transient sonar sounds[C]// Proc. 114th Conv. Audio Engineering Society, Amsterdam, Netherlands, 2003.

[50] Tucker S, Brown G. Investigating the perception of the size, shape and material of damped and free vibrating plates[R]. Dept. Comput. Sci., Univ. Sheffield, Sheffield, U.K., Tech. Rep. CS-02-10, 2002.

[51] Tucker S, Brown G. Timbral segregation of sonar transients using auditory processing techniques[C]// Underwater Defence Technology Europe, La Spezia, Italy, 2002.

[52] Tucker S, Brown G. Classification of transient sonar sounds using perceptually motivated features[J]. IEEE Journal of Oceanic Engineering, 2005, **30**(3): 588-600.

[53] Tucker S. An ecological approach to the classification of transient underwater acoustic events: Perceptual experiments and auditory models[D]. Univ. of Sheffield, 2003.

[54] Tucker S. Auditory analysis of sonar signals[R]. Ph. D. Transfer Report, 2001.

[55] Hines P, Allen N, Young V. Aural classification for active sonar[R]. Tech. Rep., DRDC Atlantic TR 2007-361, 2008.

[56] Young V, Hines P. Perception-based automatic classification of impulsive-source active sonar echoes[J]. J. Acoust. Soc. Am., 2007, **122**(3): 1502-1517.

[57] Hines P, Young V, Scrutton J. Aural classification of coherent-source sonar echoes[C]// Proc. Int. Symp. On Underwater Reverberation and Clutter, Lerici, Italy, 2008.

[58] Myers V, Fawcett J, Hines P, et al. Reconstruction and fusion of perceptual features for automatic classification of sonar echoes[C]// Oceans 2008, Quebec City, Canada, 2008.

[59] Murphy S, Hines P. Aural classification and temporal robustness[R]. Tech. Rep., DRDC Atlantic TR 2010-136, 2010.

[60] Allen N, Hines P, Young V. Performances of human listeners and an automatic aural classifier in discriminating between sonar target echoes and clutter[J]. J. Acoust. Soc. Am., 2011, **130**(3): 1287-1298.

[61] Allen N. Receiver-operating-characteristic (ROC) analysis applied to listening-test data[R]. Tech. Rep., DRDC Atlantic TR 2007-353, 2008.

[62] Mill R, Brown G. Auditory-based time-frequency representations and feature extraction techniques for sonar processing[R]. Tech. Rep., CS-05-12, 2005.

[63] Lutfi R, Liu C, Stoelinga C. Auditory discrimination of force of impact[J]. J. Acoust. Soc. Am., 2011, **129**(4): 2104-2111.

[64] Sildam J. Masking of time-frequency patterns in applications of passive underwater target detection[J]. EURASIP Journal on Advances in Signal Processing, 2010, 298038.

[65] Lim T, Bae K, Hwang C, et al. Classification of underwater transient signals using MFCC feature vector[C]// International Conference on Signal Processing and Its Applications, 2007.

[66] Tolba H, Elgerzawy A. Comparative experiments to evaluate a CHMM-based identification approach to naval targets[C]// International Conference on Systems, Signals and Image Processing, Chalkida, Greece, 2009.

[67] Xie Z, Zhou Y. The study on classification for marine mammal based on time-frequency perception[C]// International Conference on Bioinformatics and Biomedical Engineering, Chengdu, China, 2010.

[68] 吴玺宏, 迟惠生, 王楚. 基于听觉外周模型的语音信号听觉神经表示[J]. 生物物理学报, 1997, **13**(2): 213-220.
WU Xihong, CHI Huisheng, WANG Chu. Auditory model based neural representation of speech signal[J]. Acta Biophysica Sinica, **13**(2): 213-220.

[69] 吴玺宏, 迟惠生, 王楚. 基于听觉模型的语音信号听觉神经处理[J]. 生物物理学报, 1997, **13**(3): 427-434.
WU Xihong, CHI Huisheng, WANG Chu. An auditory model based neural processing of speech signal[J]. Acta Biophysica Sinica, **13**(3): 427-434.

[70] 迟惠生, 吴玺宏. 听觉计算模型在自动语音识别中的作用[J]. 自然科学进展, 2000, **10**(11): 969-976.
CHI Huisheng, WU Xihong. The auditory computing model in automatic speech recognition[J]. Progress in Natural Science, 2000, **10**(11): 969-976.

[71] 张焱, 张杰, 黄志同. 基于一种听觉模型的特征提取及语音识别[J]. 南京理工大学学报, 1998, **22**(2): 113-116.
ZHANG Yan, ZHANG Jie, HUANG Zhitong. Feature extraction and speech recognition based on an auditory model[J]. Journal of Nanjing University of Science and Technology, 1998, **22**(2):

113-116.

[72] 蔡悦斌, 张明之, 史习智, 等. 舰船噪声波形结构特征提取及分类研究[J]. 电子学报, 1999, **27**(6): 129-130.
CAI Yuebin, ZHANG Mingzhi, SHI Xizhi, et al. The feature extraction and classification of ocean acoustic signals based on wave structure[J]. Acta Electronica Sinica, 1999, **27**(6): 129-130.

[73] 陆振波, 章新华, 朱进. 基于 MFCC 的舰船噪声特征提取[J]. 舰船科学技术, 2004, **26**(2): 51-54.
LU Zhenbo, ZHANG Xinhua, ZHU Jin. Feature extraction of ship-radiated noise based on Mel frequency cepstrum coefficients[J]. Ship Science and Technology, 2004, **26**(2): 51-54.

[74] 陆振波, 章新华, 胡洪波. 水中目标辐射噪声的听觉特征提取[J]. 系统工程与电子技术, 2004, **26**(12): 1801-1803.
LU Zhenbo, ZHANG Xinhua, HU Hongbo. Auditory feature extraction of noise radiated from an underwater target[J]. Systems Engineering and Electronics, 2004, **26**(12): 1801-1803.

[75] 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, 19.
YANG Xiong, CHENG Yusheng. Extraction of ship-radiated noise characteristics based on analysis of short-term energy and subjective hearing[J]. Technical Acoustics, 2004, **23**(1): 11-13, 19.

[76] 彭圆, 王晟, 王科俊等. 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 Yuang, WANG Sheng, WANG Junke, et al. A study on underwater target classification applying perception linear prediction method[J]. Acta Acustica, 2006, **31**(2): 146-150.

[77] 杨宏晖, 彭圆, 曾向阳. 基于声信号人耳听觉谱特征和 SVME 的水下目标识别[C]// 水声综合技术学术会议, 2008.
YANG Honghui, PENG Yuan, ZENG Xiangyang. Underwater target recognition based on human auditory spectrum characteristics of acoustic signal and SVME[C]// Underwater Acoustic Technology Conference, 2008.

[78] 谢骏, 胡均川, 笪良龙等. 基于耳蜗模型的舰船噪声谱分析[J]. 传感器与微系统, 2009, **28**(5): 43-45.
XIE Jun, HU Junchuan, DA Lianglong, et al. Spectrum analysis for naval vessel radiation noise based on cochlea model[J]. Transducer and Microsystem Technologies, 2009, **28**(5): 43-45.

[79] 马元锋, 陈克安, 王娜等. 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 Yuanfeng, CHEN Kean, WANG Na, et al. Application of auditory spectrum—based features into acoustic target recognition[J]. Acta Acustica, 2009, **34**(2): 142-150.

[80] 马元锋, 陈克安, 王云山, 等. 自适应听觉感知时频分析模型[J]. 声学学报, 2010, **35**(4): 393-402.
MA Yuanfeng, CHEN Kean, WANG Yunshan, et al. An adaptive perception based time-frequency analysis model[J]. Acta Acustica, 2010, **35**(4): 393-402.

[81] Ma Y, Chen K. A time-frequency perceptual feature for classification of marine mammal sounds[C]// ICSP2008 Proceedings, 2008.

[82] 王娜, 陈克安. 听觉感知特征在目标识别中的应用[J]. 系统仿真学报, 2009, **21**(10): 3128-3132.
WANG Na, CHEN Kean. Applications of auditory perceptual features into targets recognition[J]. Journal of System Simulation, 2009, **21**(10): 3128-3132.

[83] 王娜, 陈克安. 水下噪声音色属性回归模型及其在目标识别中的应用[J]. 物理学报, 2010, **59**(4): 2874-2880.
WANG Na, CHEN Kean. Regression model of timbre attribute for underwater noise and its application to target recognition[J]. Acta Physica Sinica, 2010, **59**(4): 2874-2880.

[84] 王娜, 陈克安, 黄凰. 水下噪声听觉属性的主观评价与分析[J]. 物理学报, 2010, **58**(10): 7330-7338.
WANG Na, CHEN Kean, HUANG Huang. Subjective evaluation and analysis of auditory attributes for underwater noise[J]. Acta Physica Sinica, 2010, **58**(10): 7330-7338.

[85] 王娜, 陈克安. 水下噪声音色属性回归模型及其在目标识别中的应用[J]. 物理学报, 2010, **59**(4): 2874-2880.
WANG Na, CHEN Kean. Regression model of timbre attribute for underwater noise and its application to target recognition[J]. Acta Physica Sinica, 2010, **59**(4): 2874-2880.

[86] 王娜, 陈克安. 分段谱质心特征在水下目标识别中的应用[J]. 兵工学报, 2009, **30**(2): 144-149.
WANG Na, CHEN Kean. Application of sub-band spectral centroid features to recognizing underwater targets[J]. ACTA ARMAMENTARII, 2009, **30**(2): 144-149.

[87] 陈克安, 王娜, 王金昌. 人耳辨识非语言声目标能力的实验研究[J]. 物理学报, 2009, **58**(7): 5075-5082.
CHEN Kean, WANG Na, WANG Jinchang. Investigation on human ear's capability for identifying non-speech objects[J]. Acta Physica Sinica, 2009, **58**(7): 5075-5082.

[88] 陈克安, 王娜, 伍莹, 等. 基于音色属性辨识声目标的主观评价实验研究[J]. 科学通报, 2010, **55**(8): 651-659.
CHEN Kean, WANG Na, WU Ying, et al. Subjective evaluation experiments of timbre attribute based acoustic target identification[J]. Chinese Science Bulletin, 2010, **55**(8): 651-659.

[89] Duda R, Hart P, Stork D. Pattern Classification[M]. 2nd Edition. Beijing: China Machine Press, 2004.

[90] Bishop C. Pattern recognition and machine learning[M]. New York: Springer, 2007.