

一种视听融合的水下目标识别方法研究

马理想, 曾向阳

(西北工业大学航海学院, 陕西西安 710072)

摘要: 特征提取是水下目标识别研究中最关键的技术之一, 特征参数的优劣将直接决定分类识别系统的性能。将声信号的听觉与视觉感知特征结合, 应用于水下目标识别, 通过实验得出如下结论, 相比于单独应用听觉特征, 融合特征的平均识别率能提高4%~6%以上, 特别是将听觉特征与声谱图的Gabor小波变换特征、灰度-梯度共生特征进行融合后, 分类性能较好, 平均达到87%以上。

关键词: 水下目标识别; 听觉特征; 可视化; 图像特征; 特征选择

中图分类号: TB561

文献标识码: A

文章编号: 1000-3630(2015)-03-0209-05

DOI 编码: 10.3969/j.issn1000-3630.2015.03.004

Study of underwater targets recognition based on audiovisual feature integration

MA Li-xiang, ZENG Xiang-yang

(College of Marine Engineering, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China)

Abstract: Feature extraction is one of the most important techniques of underwater targets recognition. Feature parameters directly determine the performances of classification systems. In this paper, a joint method which combines the auditory and visualized feature extraction methods is proposed and applied to underwater target recognition after feature selection and fusion. Experimental results show that fusion features achieve a better performance than a single audio feature, and the enhancement of recognition rate is 4%~6%, and that the fusion of audio feature with gray gradient co-occurrence matrix and Gabor small wave exchange feature can obtain an even better performance and the recognition rate is over 87%.

Key words: underwater targets recognition; auditory feature; visualization; image feature; feature selection

0 引言

目标识别是现代声呐和水声对抗的重要组成部分, 而特征提取是目标识别的关键。经过多年的发展和积累, 目前已提出了不少特征提取方法。早期提出的有时域波形结构特征^[1], 主要参数有过零率、峰间幅值分布、波长差分布、波列面积等, 该方法实现过程较为容易, 但对一些复杂的水下噪声信号的分类识别显得无能为力。后来又发展了基于谱估计^[2,3]的特征提取方法, 包括经典谱估计、现代谱估计、倒谱分析、DEMON谱、LOFAR谱分析等。功率谱和相关函数是二阶统计特性, 在随机过程是正态分布时, 它们能完全代表过程的特性。但实际

的水声信号或噪声往往不是理想的高斯分布, 用二阶统计特性不能全面描述信号特性, 只有高阶统计特性(Higher-Order Statistics, HOS)才能更全面地反映非高斯信号的特性。

传统的傅里叶变换有时间积分作用, 平滑了非平稳随机信号中的时变信息, 因而其频谱只能代表信号中各频率分量的总强度。采用短时傅里叶变换(Short-Time Fourier Transform, STFT)对时变信号逐段进行分析, 虽具有时频局部化性质, 但其时间分辨率和频率分辨率是互相矛盾的, 不能兼顾。而小波变换通过对原小波的平移和伸缩, 能使基函数长度可变, 因而可获得不同的分辨率。基于其高分辨力的特点, 可以提取多种小波特征^[4]。

近年来, 随着人们对人耳听觉模型研究的深入及人工智能的迅速发展, 基于人耳听觉特性^[5]的特征提取方法成为了研究的热点, 由此得到的特征参数有: 梅尔频率倒谱系数(Mel Frequency Cepstrum Coefficient, MFCC)、感觉加权线性预测(Perceptual Linear Predictive, PLP)、响度、尖锐度、粗糙度以

收稿日期: 2014-02-11; 修回日期: 2014-03-20

基金项目: 国家自然科学基金(11374241)、陕西省自然科学基金(2012JM1010)资助项目。

作者简介: 马理想(1992—), 男, 安徽阜阳人, 硕士, 研究方向为声信号处理。

通讯作者: 马理想, E-mail: malixiang1234@163.com

及波动强度等。此外,在对人耳听觉模型进行深入研究后,例如 Gammatone 基底膜模型、Meddis 内毛细胞模型,可以得到更多相关的听觉特征参数。不过,由于人耳听觉机理还是一个尚未完全破解的难题,这方面的研究还有很大的空间。

本文有别于传统的声学特征提取方法,尝试先将获得的声信号转化为图像,再提取图像特征^[6],然后将提取出的图像特征与传统的听觉特征进行融合,最后将融合后的特征用于水下目标识别。由于图像属于二维信号,相比于一维时间序列可能包含更多的特征信息,且图像特征提取与识别已有大量成功的应用,这种视听特征融合的方法有望获得更好的识别性能。

本文在对三类水下目标(A:舰船;B:潜艇;C:UUV)听觉特征和可视化特征进行识别正确率分析的基础上,最终选择了识别率较高的三种听觉特征,即 16 维响度特征、9 维 PLP 特征和 11 维 MFCC 特征,分别与信号声谱图的 5 维 Gabor 小波特征以及 15 维的灰度梯度共生矩阵特征融合,比较各融合特征的识别正确率以及鲁棒性。

1 听觉特征提取

1.1 响度特征

响度描述的是声音的响亮程度,表示人耳对声音的主观感受。Zwicker 计算方法^[7]是经典的响度计算方法。在计算响度时要考虑掩蔽效应,相关的一个重要物理量是临界频带,单位: Bark。通常把 20 Hz~16 kHz 的频率范围划分为 24 个临界频带。频率小于 500 Hz 时,临界频带 1 Bark 带宽约等于 100 Hz;频率大于 500 Hz 时,1 Bark 带宽为该临界频带中心频率的 20%。转换公式如下:

$$z=13\arctan(0.00076f)+3.5\arctan(f/7500)^2 \quad (1)$$

式中: f 为临界带宽的中心频率; z 为临界频带的带宽。

1.2 PLP 特征

PLP 特征参数是一种基于听觉模型的特征参数。该特征参数是全极点模型预测多项式的一组系数,等效于一种线性预测系数(Linear Prediction Coefficient, LPC)特征。不同之处是 PLP 技术将人耳听觉实验获得的一些结论,通过近似计算的方法进行了工程化处理,将输入的语音信号经听觉模型处理后所得的信号替代传统的 LPC 分析所用的时域信号。这样处理后的语音频谱考虑了人耳的听觉特

点,因而有利于抗噪特征提取。

1.3 MFCC 特征参数

与普通倒谱分析不同的是, MFCC 分析与人耳的听觉特征紧密联系,这是因为人耳所听到声音的高低与声音频率并不呈线性正比关系,而用 Mel 频率尺度则更符合人耳的听觉特征。Mel 频率尺度的值大体上对应于实际频率的对数分布关系,两者的具体关系可用式(2)表示:

$$\text{Mel}(f)=2595\lg(1+f/700) \quad (2)$$

式中: f 为频率;临界带宽随频率的变化而变化,并与 Mel 频率的增长一致,在 1000 Hz 以下大致呈线性分布,带宽为 100 Hz 左右;在 1000 Hz 以上呈对数增长。图 1 显示了 Mel 频率与线性频率的关系。

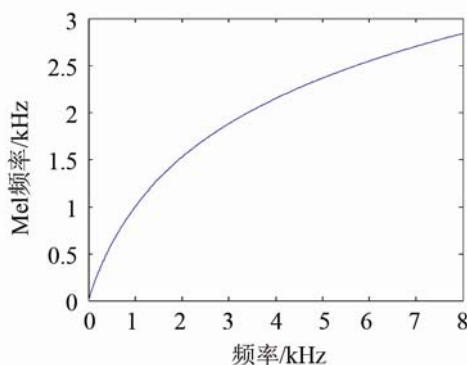


图 1 Mel 频率与线性频率的关系
Fig.1 Relationship between Mel and linear frequencies

2 信号的可视化及可视化特征提取

2.1 信号声谱图

声谱图是一种重要的时频图,它反映了声信号的动态频谱特征,在语音信号中被视为可视语言。声谱图一般分为宽带声谱图和窄带声谱图,是一张有颜色或灰度变化的二维图。图 2 为某类水下目标信号的声谱图。图中横坐标表示时间,纵坐标表示

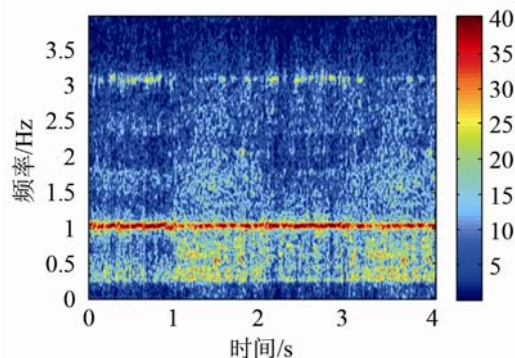


图 2 某类水下目标信号的声谱图
Fig.2 Spectrogram of certain types of underwater targets

频率, 图像中的灰度值代表某时刻对应该频率处的能量, 此处为短时傅里叶变换幅度的平方。

这种时频分布图经过灰度处理后, 可以利用图像特征提取方法提取特征参数。

2.2 信号的可视化特征

2.2.1 灰度-梯度共生矩阵特征

基于灰度共生矩阵可以得到灰度-梯度共生矩阵^[8], 从中提取的特征参数对于边界不清晰的图像具有较好的分类效果^[9]。灰度-梯度共生矩阵集中反映了图像中像素的灰度和梯度的相互关系, 各像素的灰度是构成一幅图像的基础, 而梯度则是构成图像边缘轮廓的要素。灰度-梯度空间可以清晰地表现图像内像素灰度与梯度的分布规律, 同时也体现了各像素与其领域像素的空间关系, 对图像的纹理能很好地描绘。灰度-梯度共生矩阵纹理分析方法是利用图像的灰度和梯度的综合信息提取纹理特征。表1所示为几种灰度-梯度特征的计算公式。

表1 灰度-梯度共生矩阵的中部分参数的数字特征

Table 1 Expressions of some parameters in the grayscale-gradient co-occurrence matrix

小梯度优势	大梯度优势
$T_1 = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y) / (y+1)^2}{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y)}$	$T_1 = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y) y^2}{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y)}$
灰度分布的不均匀性	梯度分布的不均匀性
$T_3 = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} [\hat{H}(x, y)]^2}{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y)}$	$T_3 = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} [\hat{H}(x, y)]^2}{\sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y)}$
能量	灰度平均
$T_5 = \sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} \hat{H}(x, y)^2$	$T_6 = \sum_{x=0}^{L-1} \sum_{y=0}^{L_g-1} x \hat{H}(x, y)$

2.2.2 Gabor 小波特征

Gabor 小波特征是根据模拟人类视觉系统而产生的。通过模拟人类视觉系统, 可以将视网膜成像分解成一组滤波图像, 每个分解的图像能够反映频率和方向在局部范围内的强度变化。通过一组多通道 Gabor 滤波器, 可以获得纹理特征。

由于小波变换在数字图像中局部区域的频率和方向信息提取方面有优异的性能, 因此, 一些学者们将其用于图像的纹理特征提取中, 并且取得了非常满意的效果。最为常用的做法就是对图像进行小波变换后, 从小波子带中提取平均值以及方差作为纹理特征。虽然小波分析取得了不错的效果, 然而与 Gabor 滤波器组相比而言, 还存在不足。Gabor

变换已被证明是在二维测不准情况下, 对信号空间和频率域的最有力描述。这些滤波器可以当作方向和尺度都可变化的边缘和直线的检测, 并且对一个给定区域中的这些微观特征的统计, 经常可以用来表示纹理信息的特征。其优点在于, 能够很好地模拟哺乳动物视皮层简单细胞的感受域, 符合视觉生理特点, 能获得频域和空域的最佳联合分辨率。

3 实验研究

3.1 融合特征的实验对比

由所选取的听觉和可视化特征自由组合可以得到六种融合特征, 分别是响度(16 维)-灰度梯度共生矩阵特征(15 维)、响度(16 维)-Gabor 小波变换特征(5 维)、PLP(9 维)-灰度梯度共生矩阵特征(15 维)、PLP(9 维)-Gabor 小波变换特征(5 维)、MFCC(11 维)-灰度梯度共生矩阵特征(15 维)、MFCC(11 维)-Gabor 小波变换特征(5 维)。

针对 A、B、C 三类水下目标辐射噪声信号, 分别选取 100 个训练样本和 100 个测试样本, 即训练样本数与测试样本数之比为 1:1。分类器为 SVM 算法。

实验结果见表 2 和图 3, 表 3 则是在单听觉特征情况下系统的识别正确率。

图 3 中系列 1、2、3、4、5、6 分别代表:

表2 听觉特征与声谱图可视化特征的融合特征的识别率(%)

Table 2 The recognition rates of the fusion of audio feature with spectrogram visualized feature(%)

融合特征	A	B	C	总识别率
MFCC-Gabor 小波变换	79.0	86.0	93.0	86.0
MFCC-灰度梯度共生矩阵	78.0	87.0	91.0	85.3
PLP-Gabor 小波变换	80.0	89.0	98.0	89.0
PLP-灰度梯度共生矩阵	80.0	90.0	98.0	89.0
响度-Gabor 小波变换	80.0	87.0	93.0	86.7
响度-灰度梯度共生矩阵	78.0	86.0	100.0	88.0

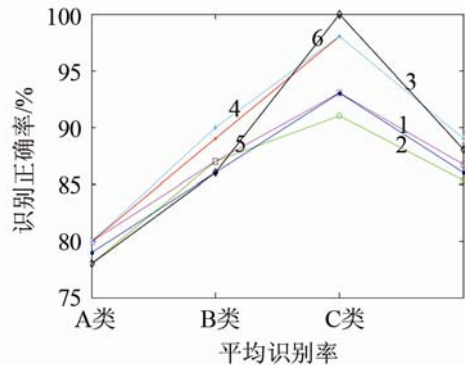


图3 视听融合特征的正确识别率

Fig.3 Recognition rates of the audio-visual fusion features

表 3 单听觉特征的识别率(%)

Table 3 The recognition rates of a single auditory feature(%)

听觉特征	A	B	C	总识别率
响度特征	79.0	83.0	87.0	83.0
MFCC 特征	68.7	77.9	83.9	76.8
PLP 特征	65.8	81.7	84.3	77.3

MFCC-Gabor 小波变换特征、MFCC-灰度梯度共生矩阵特征、PLP-Gabor 小波变换特征、PLP-灰度梯度共生矩阵特征、响度-Gabor 小波变换特征、响度-灰度梯度共生矩阵特征。

由以上结果可以看出,融合特征的识别率基本都在 85%以上,其识别率都较单听觉特征时的识别率高,最高达到 89%,相比于单听觉特征的最高识别率 83%,提高了 6%。

3.2 视听融合特征的识别性能实验研究

由 3.1 节可知,对于融合特征,听觉特征与可视化特征进行融合所得的识别结果相对更好,其中表现最好的融合特征是 PLP-Gabor 小波变换特征、PLP-灰度梯度共生矩阵特征和响度-灰度梯度共生矩阵特征。为进一步检验这三组融合特征的稳健性和抗噪性,以下进一步开展实验研究。

(1) 训练-测试样本比对识别结果的影响

3.1 节中介绍的融合特征测试时,使用的训练和测试样本比为 1:1,这里再分别对测试样本和训练样本为 1:2 和 2:1 时进行测试。实验结果见表 4。

由表 4 可知,PLP-Gabor 小波变换特征和响度-灰度梯度共生矩阵特征在训练和测试样本变化的情况下,识别率变化相对较小,说明这两类特征的鲁棒性较好。

表 4 不同训练测试比下融合特征的的识别正确率(%)

Table 4 Recognition rates of the fusion features under different training-test ratio(%)

融合特征	训练样本与测试样本的比例		
	1:2	1:1	2:1
PLP-Gabor 小波变换特征	73.0	89.0	83.0
PLP-灰度梯度共生矩阵特征	61.0	89.0	75.8
响度-灰度梯度共生矩阵特征	76.0	88.0	82.0

(2) 融合特征的抗噪性能实验

将测试信号分别加上高斯白噪声,得到信噪比分别为-10、-5、0、5、10 dB 的信号,利用测试实验得出 PLP-Gabor 小波变换特征、PLP-灰度梯度共生矩阵特征和响度-灰度梯度共生矩阵特征的分类正确率,结果见表 5,对应的曲线如图 4 所示。图 4 中,1、2、3 分别代表 PLP-Gabor 小波变换特征、PLP-灰度梯度共生矩阵特征、响度-灰度梯度共生矩阵特征。

表 5 不同信噪比情况下融合特征的识别正确率(%)

Table 5 Recognition rates of the fusion feature under different SNRs(%)

融合特征	信噪比/dB				
	-10	-5	0	5	10
PLP-Gabor 小波变换	66.7	65.0	67.7	72.7	85.0
PLP-灰度梯度共生矩阵	56.3	68.0	70.3	74.0	76.0
响度-灰度梯度共生矩阵	51.7	63.0	68.7	70.0	72.3

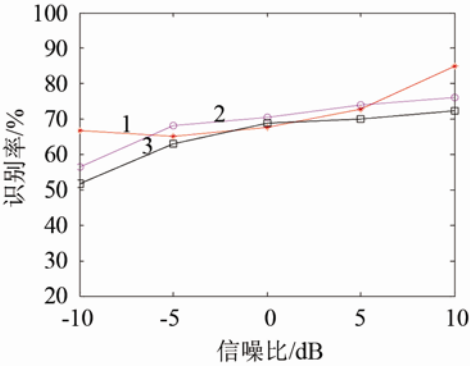


图 4 不同信噪比情况下融合特征的正确识别率

Fig.4 Recognition rates of the fusion features under different SNRs

从图 4 可以看出,对各种特征而言,总体呈现出随信噪比的增加识别率也逐渐增加的趋势。其中,PLP-Gabor 小波变换特征抗噪能力相对更优。综合以上两方面可以认为,PLP-Gabor 小波变换特征在视听融合特征中具有相对更好的识别性能。

4 结 语

通过本文的实验结果可以看出,水下目标辐射噪声信号听觉特征和可视化特征的融合能够有效地提高目标信号的正确识别率。下一步还将深入探讨声学特征与可视化特征之间的内在联系,这有助于寻找更加有效的视听特征。

参 考 文 献

[1] 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): 129-131.
CAI Yuebing, ZHANG Mingzhi, SHI Xizhi, et al. The feature extraction and classification of ocean acoustin signals baned on wave stucture[J]. Acta Electronica sinica, 1999, 27(6): 129-131.
[2] 余秋星. 水下目标识别的相关技术研究[D]. 西安: 西北工业大学博士学位论文, 2004.
YU Qiuxing. Research on underwater target recognition[D]. Xi'an: Northwestern Polytechnical University dissertation, 2004.
[3] 余秋星, 李志舜. 舰船辐射噪声的特征提取方法[J]. 鱼雷技术, 2002, 10(6): 23-25.
YU Qiuxing, LI Zhishun. Characteristics extraction method of ship radiative noise[J]. Torpedo technology, 2002, 10(6): 23-25.
[4] 章新华. 基于小波变换的舰船辐射噪声信号特征提取[J]. 声学学报, 1997, 22(2): 139-144.
ZHANG Xinhua. Feature extraction of ship radiated noises based on wavelet tansform[J]. Acta Acustica, 1997, 22(2): 139-144.
[5] 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]. 西安: 西北业大学硕士学位论文, 2009.

- WANG Na. A feature extraction approach based on auditory properties and psychoacoustic model and its application in targets recognition[D]. Xi'an: Northwestern Polytechnical University dissertation, 2009.
- [6] 何佳若, 曾向阳. 声信号的可视化特征提取方法[J]. 电声技术, 2011, 35(7): 61-64.
- HE Jiaruo, ZENG Xiangyang. Method of extracting visualized features for sound signals[J]. Audio Engineering, 2011, 35(7): 61-64.
- [7] Zwicker E, Fastl H. Psychoacoustics facts and models[M]. Springer Verlag Berlin Heidelberg, 1999.
- [8] 洪继光. 灰度-梯度共生矩阵纹理分析方法[J]. 自动化学报, 1984, 10(1): 15-18.
- HONG Jiguang. Gray level-gradient cooccurrence matrix texture analysis method[J]. Acta automatica sinica, 1984, 10(1): 15-18.
- [9] ZENG Xiangyang, HE Jiaruo, MA Lixiang. Image representation of acoustic features for the automatic recognition of underwater noise targets[C]// GCIS, 2012.

• 简 讯 •

声学领域两个重要学术交流会议将在 2015 年下半年度先后召开

1. 第八届海峡两岸声学学术交流研讨会将在 2015 年 8 月 26~29 日在我国首都北京市内召开。海峡两岸声学学术交流会议自 2008 年在香港成功召开首届会议以来, 根据两岸声学学会商定每年一届轮值筹备召开的原则, 已先后召开共七届会议。其中一届、五届在香港召开, 二届、六届在台湾召开, 三届、四届和七届先后在上海、杭州及云南普洱市召开。据去年七届普洱会议期间商定, 第八届、第九届及第十届的三次学术交流会议将先后安排在北京、澳门和台湾三地举行。现经商议已决定第八届海峡两岸声学学术交流会议已确定在 2015 年 8 月 26 日(周三为报到日)至 8 月 29 日在北京召开。会议由中国声学学会环境声学分会、上海市声学学会和上海市章奎生声学工程顾问有限公司三家联合主办; 并由中国环境保护产业协会噪声与振动控制委员会及北京市绿创环保集团有限公司共同协办, 并组成了会议筹委会及秘书组。四月下旬已发出第一轮通知即征文通知。要求在五月底前报送论文题目及三百字以内的论文摘要, 六月底前报送全文论文, 限中文 A4 纸每篇不超过四页。若需在论文集封页印制彩色企业宣传广告或黑白插页企业介绍版页, 请提前与会议秘书处联系预约, 并应在 7 月 15 日前将宣传版面电子版稿样电邮发至秘书处邮箱。

本届会议主题为“重视声学科科技, 促进创新发展”, 会议注册费为: 1200 元/人, 主要用于会议厅租用及布置、会议用餐、论文集编辑及印刷出版、会议租车及集体照相相等。

会议秘书处主要成员及联系方式:

北京 许欣(中国声学学会环境噪声分会)

电话: 010-62554489 手机: 13641386876

电邮: xuxin@mail.ioa.ac.cn

张玉敏(中国环保产业协会噪声与振动控制委员会)

电话: 010-83523485-814 手机: 13811558991

电邮: zhangyumin002@sina.com

上海 宋拥民(上海章奎生声学工程顾问有限公司)

电话: 021-65891758 手机: 13052233060

电邮: asong1102@sina.com

章奎生(上海章奎生声学工程顾问有限公司)

电话: 021-65891289 手机: 13801987251

电邮: zhangkuisheng268@163.com

参加会议应付的注册费 1200 元/人, 论文版面费 300 元/篇(限四页内)及论文集的企业宣传版面费等请按规定汇付至上海声学学会办公室, 汇款账号为:

户 名: 上海市声学学会

开户行: 招商银行四平支行

账 户: 214180080610001

汇款后请将汇款底联、发票抬头、邮寄地址及收款人名电邮发至秘书处邮箱。

2. 第十四届全国噪声与振动控制工程学术会议已定于 2015 年 11 月中旬在江苏省会南京市召开。本届会议仍由中国环保产业协会噪声与振动控制专业委员会牵头组织, 联合中国声学学会环境声学分会、中国建筑学会建筑物理分会建声专业委员会、中国环境科学学会环境工程分会、中国环境科学学会环境物理专业委员会、中国职业安全健康协会噪声与振动控制专业委员会、中国振动工程学会振动与噪声控制专业委员会、中国机械工程学会环境保护分会、中国劳动保护科学技术学会噪声与振动控制委员会等国内与噪声振动相关的各大学会/协会, 以及台湾地区声学学会、台湾地区振动噪音工程学会, 以及香港地区声学学会等共计 12 个学会组织联合组织召开。自 1982 年在安徽黄山市成功召开首届会议至今已历时 23 年, 先后共召开了 13 届学术会议。声学界的一代宗师马大猷教授曾先后出席了六届会议, 每次都作了大会报告, 对会议极为支持。近年来会议已改为每两年召开一届。规模已达到每届有 200 人左右参会, 交流论文也达 150 篇左右, 会期一般为三天左右。本届会议预计在六、七月间发正式邀请函(即征文通知), 请注意会议信息及时报名参会。

上海章奎生声学工程顾问有限公司 章奎生、宋拥民