

可实现听觉场景匹配的智能数字助听器算法

魏 政¹, 尹雪飞¹, 陈克安²

(1. 西北工业大学电子信息学院, 西安 710129; 2. 西北工业大学航海学院, 西安 710072)

摘要: 听觉场景匹配是当前智能数字助听器研究的热点问题, 目前匹配精度大都较为理想, 然而过高的计算量还是影响了硬件处理实时性的实现。此外, 各类研究中所涉及的听觉场景类型和数量与人们实际接触较多的听觉场景也有差别。为此, 首先对日常听觉场景类型进行分析统计, 在此基础上提出了一种简单高效的智能数字助听器听觉场景匹配方法, 以临界带能量比例为特征, 以最小距离为分类算法, 计算量较小、精度高, 更适合数字助听器对计算量的苛刻要求。最后, 通过主观测听, 证明该方法与人耳的识别效果类似。

关键词: 智能数字助听器; 听觉场景匹配; 临界带能量比例; 主观测听

中图分类号: TN713

文献标识码: A

文章编号: 1000-3630(2012)-05-0511-06

DOI 编码: 10.3969/j.issn1000-3630.2012.05.013

A method of recognizing acoustic scenes for intelligent digital hearing aids

WEI Zheng¹, YIN Xue-fei¹, CHEN Ke-an²

(1. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China;

2. School of Marine Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Recently, the recognition of acoustic scenes has been a hot topic in the field of intelligent digital hearing aids. Though the recognition precision is high in most current studies, the huge computational load remains a big problem that influences the implementation of hardware real-time processing. In addition, there exist differences between researches and peoples' daily lives in classifying and quantifying acoustic scenes. In this study, the analysis and statistics to daily acoustic scenes are addressed firstly, and then an efficient method of recognizing acoustic scenes is presented, which uses the critical band ratio as features and the minimum distance as classifier. Compared with other methods, this method is more suitable for digital hearing aids with limited resource because of the lower load and the high precision. Finally, it is shown from subjective listening tests that the proposed method has similar effects compared with auditory perception.

Key words: intelligent digital hearing aids, recognition of acoustic scenes, critical band ratio, subjective listening tests

0 引 言

随着数字助听器技术的发展, 听觉场景匹配已成为智能数字助听器的核心功能, 其算法处于信号处理的前端, 能够识别出助听器使用者当前的听力环境, 自适应地调用相应的处理程序, 实现针对不同场景声信号的个性化处理。从本质上讲, 听觉场景匹配属于环境声识别问题, 主要包含两个方面: 特征提取和分类^[1]。特征提取是对声信号进行维数约减, 提取出能代表原始信号的数据; 而分类是指通过一定方法编码声特征, 并与模板数据库进行比

对以确定声信号所属类别^[2]。因此, 对声信号的特征提取及对分类算法的研究是智能数字助听器场景匹配的关键问题。

2005 年, Ravindran^[3]对听觉倒谱进行多尺度变换, 并以此为特征, 采用高斯混合模型(Gaussian Mixture Model, GMM)对 4 种 1s 长的环境声样本分类, 效果较好。2006 年, Ma^[4]以美尔倒谱系数(Mel-frequency Cepstral Coefficients, MFCC)为主要特征, 采用隐马尔科夫模型(HMM)对 12 种环境声进行分类, 精度可达 90%。2007 年, Alexandre^[5]使用基于遗传算法的约束搜索对助听器中的声音分类进行特征选择。2009 年, Chu^[6]采用匹配跟踪(Matching Pursuit)提取声信号时频特征, 与 MFCC 结合可提高识别精度。2010 年, Lamarche^[7]以声信号幅度调制深度和瞬时频率为特征, 提出了两种助听器自适应分类系统, 并可根据当前声环境的变化分裂或聚合类。总的来看, 国外在智能数字助听器

收稿日期: 2011-09-05; 修回日期: 2011-11-17

基金项目: 国家自然科学基金资助项目(11074202)、西北工业大学研究生创业种子基金资助项目(Z2011100)

作者简介: 魏政(1986—), 男, 陕西西安人, 硕士研究生, 研究方向为现代数字信号处理。

通讯作者: 魏政, E-mail: maomao2136@126.com

听觉场景匹配这方面研究已做了大量工作,而国内研究较少。尽管这些特征提取和分类算法的识别精度较高,但计算复杂度也相对较高,对于处理能力有限的数字助听器来讲,在满足精度的前提下,算法计算量越小越好^[7]。此外,由于目前国内外还没有专门针对日常听觉场景类型的统计研究,因而大多数研究对日常听觉场景类型的假定都不尽相同,这有可能导致研究类别与实际助听器使用者日常所处场景存在较大差异。

针对上述问题,本文首先对人们日常所处听觉场景类型进行分析统计,归类出五类日常听觉场景,在此基础上,通过对这五类场景的典型环境声进行频谱分析,提出一种简单高效的听觉场景匹配方法。该方法以临界带能量比例为特征,以最小距离为分类算法,计算量相对较小。最后,通过主观测听,证明该方法与人耳的识别效果类似,整体识别精度较高。

1 日常听觉场景类型的统计

通过与西门子助听器陕西服务中心合作,以发放问卷和网络投票等方式进行调查统计,调查内容包括预先假定的几类日常听觉场景作为参考,同时被调查者也可根据自身特点补充场景并做选择。被调查者是助听器使用者和听力正常的各类人群。累计回收有效问卷 168 份,其中网络投票 50 份^[8],被调查者中男女比例约 1.2:1,平均 47.2 岁,助听器使用者占 20.8%。通过对问卷进行统计,有 20 种听觉场景是人们日常接触的,如图 1 所示。

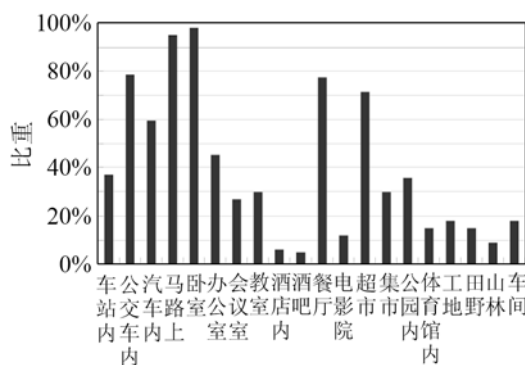


图 1 20 种场景在所有被调查者日常所处的听觉场景中的比重
Fig.1 Percentages of 20 kinds of acoustic scenes selected by investigators

经过对比高于 30% 的听觉场景进行归类,有五类听觉场景是人们日常接触最多的,分别是:(1) 车辆内,如公交车内和小汽车内,简称类 1;(2) 嘈杂的公共场所内,如餐厅、超市等,简称类 2;(3) 安

静的场所,如卧室、办公室等,简称类 3;(4) 户外交通,如公路上、车站等,简称类 4;(5) 安静的语音环境,如会议室和教室等,简称类 5。为了明确这五类听觉场景的具体含义,本文对每类场景根据自身特点进行限定。从声学特性来看,类 1 主要是指已启动的机动车内部环境,其声学特点是单一且持续性的低频引擎声;类 2 主要是指人流量大的场所内部环境,其特点是大量语音和乐音的混响;类 3 主要是指可听声源较少环境,其声学特点是声信号声压级较低、能量较小;类 4 主要是指有车辆行驶的道路和交通枢纽等,其特点是同时或先后出现多个移动的机械声源,并可能伴有高频鸣笛声;类 5 主要是指具备单纯语音声源的环境,其特点是声信号包络起降清晰,有声段和无声段间隔清晰,频谱能量较多集中在中低频。

2 环境声特征提取和分类

通过对典型环境声信号频谱特性进行分析,确定出临界带特征矢量(Critical Band Energy, CBE)可以作为特征。随后根据 CBE 在应用中的不足,本文对其进行归一化改进为临界带能量比例(Critical Band Ratio, CBR),并与最小距离分类算法配合实现听觉场景分类。

2.1 频谱统计分析

研究表明,人耳对音色的感知与声信号时域和频域特征密切相关^[9]。其中频域特征包括谱质心、谱通量、谱下降值、谱不规则性、CBE 等;时域特征包括时域质心、零交点率、短时能量、短时平均幅度、振幅包络和线性预测倒谱系数(Linear Predictive Cepstrum Coefficients, LPCC)等。

一般而言,环境声是由多种声源组成的非结构化数据,并不像单一的乐声或语音,它不能假定信号中的重复部分和谐音结构^[6]。而本文将日常听觉场景分为五类,既有通常意义上的环境声又有单一的语音,因此对这五类听觉场景声特征的提取不能简单依赖于描述结构化语音的特征,如 MFCC 或 LPCC 等^[6]。研究表明^[9],对于水下环境声,频谱结构最能反映声信号的声学特性,而对日常环境声的描述较多的也是基于改进的频域特征,但计算量往往较大。本文对定义的五类听觉场景环境声进行了频谱分析,发现各个类型内环境声信号频谱结构类似,而类间差别较大。图 2 给出了这五类典型的环境声时域和频域图,其中第一列为时域图,图 2(a1)声信号取自公交车,2(a2)取自大型超市,2(a3)取自

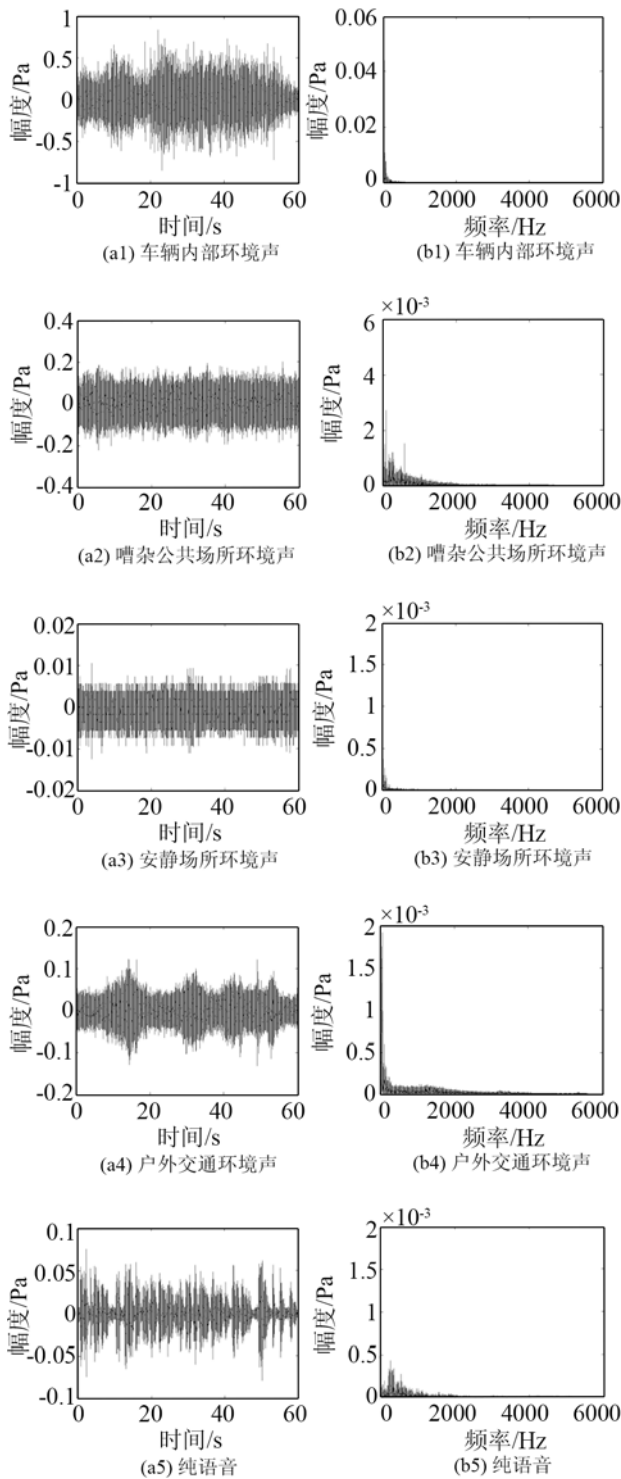


图 2 五类典型听觉场景环境声时域和频域图

Fig.2 Time domain graphics and spectrograms of 5 kinds of typical environment sounds

办公室, 2(a4)取自城市街道, 2(a5)取自讲课的教室。第二列图 2(b1)~2(b5)为相应的频域图。

由图 2 可以看出, 每类听觉场景环境声的频域能量存在较大差异。图 2(b1)表明类 1 环境声的频谱能量较多集中在 1kHz 以下, 其中 200Hz 以下的低频能量较为突出; 2(b2)表明类 2 环境声频带较

宽, 频谱能量较多集中在 3kHz 以下; 2(b3)表明类 3 环境声能量较小, 主要分布在 100Hz 左右的低频; 2(b4)表明类 4 环境声频带分布较广, 其中 2kHz 以下占主要分量, 3kHz 和 5kHz 附近频谱略微突出, 其主要原因是道路上机动车运行的机械声响、轮胎摩擦声和鸣笛声等。2(b5)则表明类 5 频谱能量较多的集中在 1.5 kHz 的中低频段且分布较均匀。综上所述, 五类典型听觉场景环境声频谱分布各有不同, 频谱结构能够反映环境声信号声学特性的差异。综合考虑频谱特性和特征提取的复杂度, 音色特征中代表频谱分布的 CBE 可以作为环境声的识别特征。

2.2 CBE 及其改进

一个纯音可以被以它为中心频率、并且具有一定频带宽度的窄带噪声所掩蔽, 当调节窄带噪声带宽使纯音处于刚能被听到的临界状态时, 即称这一带宽为临界带宽(Critical Band, CB)^[10]。Zwicker 等^[11]在早期研究中给出了计算临界带宽的经验公式:

$$CB_c = 25 + 75 \times [1 + 1.4f^2]^{0.69} \quad (1)$$

以及临界带与信号频率的转化公式:

$$Z_c = 13 \times \arctan(0.76f) + 3.5 \times \arctan\left(\frac{f}{7.5}\right)^2 \quad (2)$$

其中: CB_c 为临界频带宽, 单位 Hz; f 为临界带中心频率, 单位 kHz; Z_c 为临界带率, 单位 Bark。

CBE 是指将一帧声信号功率谱按频率高低分成若干临界带, 对每个临界带中功率谱求和, 即得 CBE^[12]。下面给出将线性频谱转化为 CBE 的步骤:

第一步, 对声信号 $x(n)$ 进行预加重和分帧加窗, 得到各帧信号 $x_n(m)$, $m=0 \sim (N-1)$, N 是 hamming 窗窗长;

第二步, 对每帧信号进行 DFT 得到 $X_n(k)$, 并计算该帧的功率谱 $|X_n(k)|^2$ 。其中 $|X_n(k)|^2$ 与原始加窗模拟信号频谱模平方 $|X_n(\exp(j\omega_k))|^2$ 具有下列关系:

$$|X_n(k)|^2 = |X_n(\exp(j\omega_k))|^2, \quad k=0 \sim N-1 \quad (3)$$

式中, $\omega_k = 2\pi f_k$, $f_k = \frac{f_s}{N}k$, 这里 f_s 为采样率;

第三步: 在 $f=0 \sim f_s/2$ 中根据式(1)和(2)确定 L 个临界带。当 $f_s=24\text{kHz}$ 时, 临界带为 23 个;

第四步: 将每个临界带内 $|X_n(k)|^2$ 取和, 如式(4)所示, 得到 $CBE=[g_1, g_2, \dots, g_L]$ 。

$$g_l = \sum_{f_k < f_{k+1}} |X_n(k)|^2, \quad l=1 \sim L \quad (4)$$

至此, 一帧信号的 CBE 便可求出。由于本文将日常听觉场景进行了归类, 每类听觉场景包含很多类似的环境, 而这些类似的环境包含的环境声各段

CBE 大小并非非常相近,例如属于类 1 的公交车内和小汽车内声信号各段 CBE 大小就有所差异。如果单纯考虑临界带内能量的大小,则很有可能导致类内差异较大。经过统计分析可知,同类型环境声的能量分布比例更为接近,而不同类的能量分布差异较大,因此若将每类环境声各段 CBE 与 CBE 总和的归一化比值作为特征,即 CBR,如式(5)所示:

$$G_{rate}=[g_1/g_{all}, \dots, g_l/g_{all}, \dots, g_L/g_{all}] \quad (5)$$

其中, $g_{all}=\sum_{l=1}^L g_l$, 那么,相同场景环境声特征向量间差异就会减小,类间差异就会增大,更易分类。

2.3 CBR 计算复杂度比较

本文对比了 CBR 与传统 MFCC 及 LPCC 的计算量。以一幅信号为例,提取 CBR 先需一次加窗运算,之后是 N 点 FFT 变换、功率谱计算和比例运算。其中加窗运算需要 N 次实数乘法, N 点 FFT 需要约 $2M\log_2 N$ 次实数乘法和 $3M\log_2 N$ 次实数加法^[13],功率谱计算需要 N 次实数乘法,比例运算需 L 次实数乘法,当 N 较大时 CBR 需要约 $2M\log_2 N+2N+L \approx 2M\log_2 2N$ 次实数乘法和 $M\log_2 2N^3$ 次实数加法。LPCC 算法是由 LPC(线性预测系数)算法推导出的,当取 p 阶 LPC 时,LPC 算法需要约 $(p+N)N+3p^2/2$ 次实数乘法, $p+N+p^2/2$ 次实数加法^[14]。由于计算机中乘法运算是通过加法运算完成的,因此算法中乘法运算开销远大于加法^[14],故本文只对比乘法运算量。当 $N=1024$ 、 $p=12$ 时,CBR 的乘法次数约 22528,远小于 LPCC 的约 1060864 次。与求 CBR 前两步相同,求一幅信号的 p 阶 MFCC 同样需要先进行加窗和 N 点 FFT 运算,但之后还需相继进行多通道滤波、对数运算和 DCT 等复杂运算,所以 MFCC 的计算量也远大于 CBR,即求一幅信号的 p 阶 MFCC 的乘法运算次数远大于 22528。因而,CBR 的计算量比 LPCC 和 MFCC 都要小。

2.4 听觉场景环境声分类算法

CBE 是从人耳对频率高低的非线性心理感觉角度来反映声音短时幅度谱分布特征,其改进特征 CBR 继承了这一特性,且所包含的信息量纲一致,它的畸变仍可用基于距离的方法来度量^[12]。考虑到计算量,本文以最小距离为分类算法作为分类器。最小距离 (Minimum Distance, MD) 分类器的基本思想是根据训练集按照统计平均生成一个代表该类的聚类中心,计算测试样本与训练样本聚类中心的欧式距离,最后判定测试样本属于与之较近距离的类^[15]。实际应用中,训练样本聚类中心已被事先存储在数字助听器中。

3 实验仿真及结果分析

通过仿真对本文方法的性能进行验证。主要包括三部分,首先简要介绍样本数据库的制作,其次重点阐述本文算法和其它常用算法效果间的对比,最后详细介绍了主观测听实验及其效果。

3.1 环境声样本数据库

为了获取更丰富的环境声样本,本文实际采集了上述五类听觉场景多种环境声实例。这些实例是在西安市内和周边录制的,采用 B&K 2850 声级计,采样率为 48kHz,16 位量化。之后将这些声信号降采样到 24 kHz,分成时长为 3 s 的声样本,提取其中部分制作样本数据库,库中样本 1/3 用于训练,2/3 用于测试,如表 1。

表 1 五类听觉场景环境声样本数据个数
Table 1 Number of samples for 5 kinds of environment sounds

样本类别	训练样本/个	测试样本/个	总样本/个
类 1	300	600	900
类 2	600	1200	1800
类 3	100	200	300
类 4	300	600	900
类 5	100	200	300
样本和	1400	2800	4200

3.2 实验结果比对及分析

为了充分考虑声信号的短时平稳性,特征提取帧长为 $N=1024$,帧移 $\Delta N=512$,由于 $f_s=24\text{kHz}$,所以实际帧长为 42.7 ms。本文 MD 分类器存储和计算的均是样本特征的聚类中心。为了比较 CBR 与原始 CBE 的分类效果,本文首先以 CBE 为特征进行分类。表 2 给出了分类结果的混淆矩阵(Confusion Matrix),这里定义整体识别率为

$$\text{整体识别率} = \frac{\text{分类正确的测试样本数}}{\text{测试样本的总数}} \times 100\% \quad (6)$$

表 2 基于 CBE 特征的 MD 分类混淆矩阵
Table 2 The confusion matrix of MD classifier with CBE

	类 1	类 2	类 3	类 4	类 5
类 1	63.1%			36.9%	
类 2		73.6%	1.4%	25.0%	
类 3			100%		
类 4		28.3%		66.7%	5%
类 5		7.5%	10%		82.5%

注:类 1: 车辆内;类 2: 嘈杂的公共场所内;类 3: 安静环境;类 4: 户外交通;类 5: 安静的语音环境。

由表 2 可以看出,以 CBE 为特征,MD 分类器整体分类精度为 72.4%。其中,类 2、类 3 和类 5 的分类效果较为理想,均达到 73% 以上。而类 1 和

类 4 的分类效果较差, 其中类 1 有 36.9% 的声样本被分类为类 4, 分析原因主要是因为类 4 和类 1 样本声特性的相似性, 而类 1 声样本 CBE 自身较分散, 导致很多样本靠类 4 聚类中心较近。同时, 可能由于类 4 声样本 CBE 自身聚合性较好, 因而没有样本被分类成类 1。此外, 类 4 声样本有 28.3% 被分类为类 2, 而类 2 中有 25% 被分类为类 4, 这主要是因为两类特征间数量级较小且接近, 使用欧式距离分类时, 公式中的求和运算容易平滑掉 CBE 中变量间的差异。总的来看, 采用 CBE 为特征容易导致特征自身聚合性不强, 特征彼此间分散性较小。表 3 给出了以 CBR 为特征的混淆矩阵, 训练集和测试集同上。

表 3 基于 CBR 特征的 MD 分类混淆矩阵
Table 3 The confusion matrix of MD classifier with CBR

	类 1	类 2	类 3	类 4	类 5
类 1	99.3%			0.7%	
类 2		84.4%	8.5%	6.4%	0.7%
类 3			100%		
类 4		33.3%	5%	60%	1.7%
类 5		7.5%			92.5%

由表 3 可以看出, 采用 CBR 作为特征向量, MD 分类器整体分类精度为 84.1%, 比采用 CBE 时提高约 12%。比较可知, 类 1 的分类精度提高明显, 达 99.3%。此外, 类 2 和类 5 各自的识别率也均提高约 10%, 但类 4 的识别率下降了 6.7%, 分析原因也是由于分类器中距离公式的求和平滑掉 CBR 中变量间的差异, 这是采用 MD 分类器自身的不足。由于 CBR 作为 CBE 的归一化改进, 两者间计算量接近, 在使用 MD 分类时, 以 CBR 为特征的整体精度却远高于 CBE。图 3 给出了这两种特征分类情况的对比。

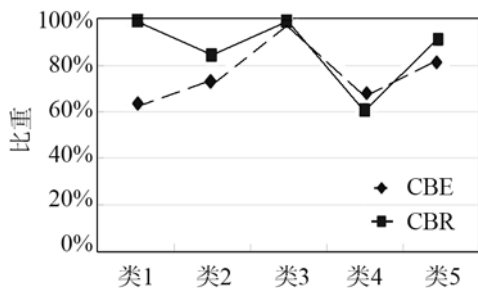


图 3 以 CBE 和 CBR 为特征的认识率对比

Fig.3 Comparison of classification accuracy between CBE and CBR

为了进一步说明 CBR 作为五类环境声特征的有效性, 本文选用两种常用的分类算法进行验证, 分别是隐马尔科夫模型(Hidden Markov Models, HMM)和人工神经网络(Artificial Neural Networks,

ANN)。其中, HMM 选择运算量较小的离散隐马尔科夫模型 (Discrete Hidden Markov Models, DHMM), 其状态数为 8, 矢量量化时码本长度为 16, 最大迭代 5000 次。ANN 选择 BP 神经网络^[16], 其输入层神经元个数与特征向量维数相同, 隐层神经元个数为 20, 采用 S 型激励函数, 输出神经元个数等于分类数 5。训练和测试样本同上。表 4 和表 5 分别给出了以 CBR 为特征, 以 DHMM 和 BP 神经网络为分类器的混淆矩阵。

表 4 以 CBR 为特征的 DHMM 分类混淆矩阵
Table 4 The confusion matrix of DHMM classifier with CBR

	类 1	类 2	类 3	类 4	类 5
类 1	90.8%	1.6%	2.3%	5.3%	
类 2		89.3%	1.4%	5%	4.3%
类 3	2.5%	10%	47.5%	40%	
类 4		23.3%	3.3%	73.4%	
类 5		7.5%			92.5%

表 5 以 CBR 为特征的 BP 神经网络分类混淆矩阵
Table 5 The confusion matrix of BP classifier with CBR

	类 1	类 2	类 3	类 4	类 5
类 1	100%				
类 2		86.5%	2.1%	11.4%	
类 3			100%		
类 4		1.7%		85%	13.3%
类 5		7.5%	27.5%		65%

由表 4 和表 5 可以看出, DHMM 和 BP 神经网络的整体分类精度较高, 分别是 83.5% 和 88.5%。其中, DHMM 在对类 3 声样本的分类上精度较差, 而 BP 神经网络对类 5 的分类也不是很理想, 均低于表 3 相应类别的分类精度。但不可否认 BP 神经网络整体的分类精度优于 DHMM 和 MD 分类器, 而在语音识别中应用较广的 DHMM 只给出了与 MD 相当的分类精度。

基于欧氏距离测量的 MD 分类算法较常规基于概率统计的分类算法计算复杂度低, 这主要是因为概率统计的分类思想是基于对大量样本进行统计分析, 并抽象出样本模型, 继而采用距离测量等准则进行判别; 而基于最小距离的 MD 算法不需进行概率统计抽象, 直接对样本使用距离测量准则, 因而其算法复杂度明显较小。表 6 给出了以 CBR 为特征(3s 长声信号), 在 MATLAB 下 MD、DHMM 和 BP 分类算法计算时间上的对比。

总的来看, 以 CBR 作为环境声样本的分类特

表 6 三类分类算法执行时间对比
Table 6 The operation times of MD, BP and DHMM

分类算法	MD	BP	DHMM
时间/ms	0.611	1.008	60.606

征具备一定的鲁棒性,且计算量远小于其它类型的常用特征。由于 MD 分类器相比于 HMM 和 BP 神经网络等分类算法更为简便,且分类精度差别不大,所以 CBR 与 MD 分类器配合更适合数字助听器对计算量的要求。

3.3 主观测听实验

主观测听的目的在于了解人类听觉系统识别环境声的性能,可作为评价本文算法优劣的参考。测听时需要被试听声样本并判断其类别,最后计算出所有被试的平均识别率,并与算法识别情况进行对比。

实验中采用数据库中部分样本进行测试,五类样本每类 18 个,共 90 个。被试共 9 名(7 男 2 女),平均 24.6 岁,听力正常。实验前先给被试分别播放 5 种场景的典型声音各一次,每次 10s,以便被试对各种环境声建立一定的主观感受。实验时每次播放的样本随机且不重复,由被试佩戴耳机测听。图 4 给出了 9 位被试的平均识别率与上述以 CBR 为特征的 3 类分类算法识别率的对比,实验中被试整体识别率为 90.38%,高于上述 3 类分类算法。其中,被试对类 1、类 3 和类 5 的识别率较高,而对类 2 和类 4 的识别率相对偏低,这与本文以 CBR 为特征,采用 MD 分类器进行分类的效果类似,从而进一步说明了本文方法的有效性。

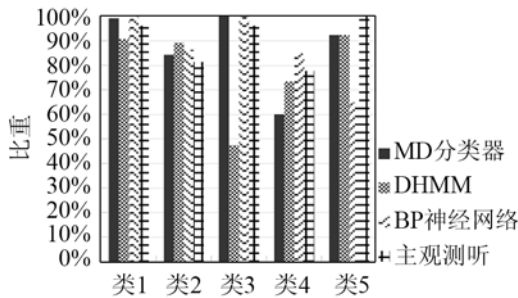


图 4 主观测听实验结果对比

Fig.4 Comparison between subjective listening and experimental results

4 结 论

本文根据当前智能数字助听器听觉场景匹配研究中存在的研究类别不统一以及计算量大的不足,首先对人们日常所处的听觉场景进行了分析统计,在此基础上给出了一种简单高效的助听器听觉场景匹配方法,该方法以临界带能量比例为特征,以最小距离为分类算法,计算量小、精度高,更适合应用在功率有限的智能数字助听器中。通过主观测听实验,证明了该方法与人耳在识别环境声方面

具有类似的效果。

参 考 文 献

- [1] Cowling M, Sitte R. Comparison of techniques for environmental sound recognition[J]. Pattern Recognition Letters, 2003, 24(1): 2895-2907.
- [2] Nordqvist P, Leijon A. An efficient robust sound classification algorithm for hearing aids[J]. Acoust. Soc. Am., 2004, 115(6): 3033-3041.
- [3] Ravindran S, Anderson D V. Audio classification and scene recognition and for hearing aids[C]// IEEE International Symposium on Circuits and Systems. Kobe, 2005, 2: 860-863.
- [4] Ma L, Milner B, Smith D. Acoustic environment classification [J], ACM Transactions on Speech and Language Processing, 2006, 3(2): 1-22.
- [5] Alexandre E, Cuadra L, Rosa M, et al. Feature selection for sound classification in hearing aids through restricted search driven by genetic algorithms [J], IEEE Trans. on Audio, Speech, and Language Processing, 2007, 15(8): 2249-2256.
- [6] Chu S, Narayanan S, C.-C, et al. Environmental sound recognition with time-frequency audio features[J]. IEEE Trans. Audio, Speech, and Language Processing, 2009, 17(6): 1142-1158.
- [7] Lamarche L, Giguère C, Gueaieb W, et al. Adaptive environment classification system for hearing aids[J]. Acoust. Soc. Am., 2010, 127(5): 3124-3135.
- [8] <http://www.openlab.net.cn/forums/threads/2516009.aspx>[EB/O L]. 2010.
- [9] 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 Ke'an. Application of sub-band spectral centroid features to recognizing underwater targets[J]. Acta Armamentarii, 2009, 30(2):144-149.
- [10] 张雪英. 语音信号处理及 MATLAB 仿真[M]. 北京: 电子工业出版社, 2010: 14-16.
ZHANG Xueying. Speech signal processing and MATLAB simulation[M]. Beijing: Electronics Industry Press, 2010: 14-16.
- [11] Zwicker E, Terhardt E. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency[J]. Acoust. Soc. Am., 1980, 68(5): 1523-1525.
- [12] 程佩青. 数字信号处理[M]. 第 2 版. 北京: 清华大学出版社, 2001: 144-150.
CHENG Peiqing. Digital signal processing[M]. Version 2. Beijing: Tsinghua University Press, 2001: 144-150.
- [13] 赵毅, 尹雪飞, 陈克安. 一种新的基于倒谱的共振峰频率检测算法 [J]. 应用声学, 2010, 29(6): 416-424.
ZHAO Yi, YIN Xuefei, CHEN Kean. A multi-channel loudness compensation algorithm based on formant extraction[J]. Applied Acoustics, 2010, 29(6): 416-424.
- [14] 赵力. 语音信号处理[M]. 第 2 版. 北京: 机械工业出版社, 2010: 44-47.
ZHAO Li. Speech signal processing[M]. Version 2. Beijing: Machinery Industry Press, 2010: 44-47.
- [15] 齐敏, 李大健, 郝重阳. 模式识别导论[M]. 北京: 清华大学出版社, 2009: 20-23.
QI Min, LI Dajian, HAO Chongyang. Introduction to pattern recognition[M]. Beijing: Tsinghua University Press, 2009: 20-23.
- [16] MATLAB 中文论坛. MATLAB 神经网络 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2010: 1-10.
MATLAB CHINESE FORUM. MATLAB Neural Network with 30 case studies[M]. Beijing: Beijing Aerospace University Press, 2010: 1-10.