

# 数字助听器中多通道响度补偿算法的研究

沐 城,张玲华

(南京邮电大学 通信与信息工程学院,江苏 南京 210003)

**摘要:**目前数字助听器中用的较多的响度补偿方法,通常是对语音的高频部分进行高增益,但由于听损患者在高频段的听觉范围比较小,直接对高频部分进行高增益容易超出听损患者的听觉范围。文中引进了宽动态压缩技术,首先将整个语音的动态范围按一定比例均匀压缩到患者的残余听力中,然后将频带非等宽划分,再在不同频段进行响度补偿和增益控制,最后对信号进行重构,并利用小波阈值去噪技术去除部分噪声。实验结果表明,该方法能够使补偿后语音的响度完全映射到听损患者的听觉范围内,从而有效提高患者的听力水平,特别是对高频部分语音的辨识,具有较好的效果。

**关键词:**响度补偿;宽动态压缩;正交镜像滤波器组;增益控制

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2016)06-0123-04

doi:10.3969/j.issn.1673-629X.2016.06.027

## Research on Multi-channel Loudness Compensation in Digital Hearing Aids

MU Cheng, ZHANG Ling-hua

(College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** Loudness compensation algorithm is a key technique in digital hearing aid. At present, digital hearing aid in multi-channel loudness compensation method, is usually high gain for the high frequency part of speech. However, because of the relatively small range of hearing loss in the high frequency range, the high gain of the high frequency part is easy to exceed the range of hearing loss. Wide dynamic compression technique is introduced in this paper. First the dynamic range of the speech according to a certain proportion is uniformly compressed to patients with residual hearing, then the band is partitioned with non equal width, and loudness compensation and gain control is carried out in different frequency band. Finally, the signal is reconstructed, removing partial noise by wavelet threshold denoising technology. Experiment shows that the method makes the loudness of speech after compensation completely mapped to range of hearing loss in patients, effectively improving the patient's hearing level, especially on speech identification of high frequency part with good effect.

**Key words:** loudness compensation; wide dynamic compression; quadrature mirror filter bank; gain control

### 0 引言

在听力损失的情况下,听阈普遍下移,从而造成听觉动态范围减小、分辨语言的能力下降等问题。响度补偿是数字助听器的关键算法之一,也是其核心任务之一。为了对声音进行压缩放大,将正常听力范围内的声音映射到听损患者的听力范围内<sup>[1]</sup>,就需要用到响度补偿。目前,大多数数字助听器都已实现多通道响度补偿功能<sup>[2]</sup>,但多数响度补偿方法主要用于处理等宽频率间隔。其实人耳对声音频率的感受与实际的

频率高低不是线性关系,而是接近对数关系,所以等宽频率间隔的补偿方法不能满足人耳的听觉特性。虽然不同患者的听力损失是不一样的,但基本上具有低频听力损失较小、高频听力损失较大的特点。有研究表明,大脑通过相对比率<sup>[3-4]</sup>来感受频率,而不是绝对频率。由于语素的频谱间存在一定位置关系<sup>[5]</sup>,通过按比例压缩可以保持这种关系。

考虑到上述情况,文中将宽动态压缩技术<sup>[6]</sup>应用于响度补偿。首先将整个语音信号的动态范围按一定

收稿日期:2015-08-26

修回日期:2015-11-27

网络出版时间:2016-05-05

基金项目:江苏省高校自然科学研究重大项目(13KJA510003);江苏高校优势学科建设工程资助项目(PAPD)

作者简介:沐 城(1990-),男,硕士研究生,研究方向为现代语音处理与通信技术;张玲华,教授,博士,博士生导师,研究方向为传感器和语音。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160505.0831.102.html>

比例均匀压缩到患者的残余听力中,再让信号顺次通过宽动态压缩模块、响度补偿模块以及增益控制模块,经过这一系列处理后,语音信号处于听损患者听阈范围内,再利用小波阈值去噪技术去除部分噪声。实验结果表明,该方法具有较好的效果。

## 1 多通道响度补偿原理及系统设计

### 1.1 多频段语音信号响度补偿原理

一般来说,人耳能感觉到频率在 16 ~ 20 000 Hz 间的声音。在这个范围内,可引起人的听觉的最小声音强度被称为听阈。如果振动强度继续增加,当达到一定限度时,将会引起鼓膜疼痛,这个限度称为痛阈。对于不同频率,人耳的感觉是不同的,对正常听力者和听损患者进行声强测试,可以分别得到正常人和听损患者的听力检查曲线图。听损患者的听力测试曲线与正常人相比是不同的,他们对声强的敏感度更低,所以他们的听阈要比一般人高,而痛阈却比一般人低,因此听力范围比普通小。一般通过多次测试特殊频点的阈值,求各频点阈值的平均值作为听损程度,连接各阈值所得曲线图作为听损患者的听力检查图。图 1 是一个听损患者的听力检查曲线图。

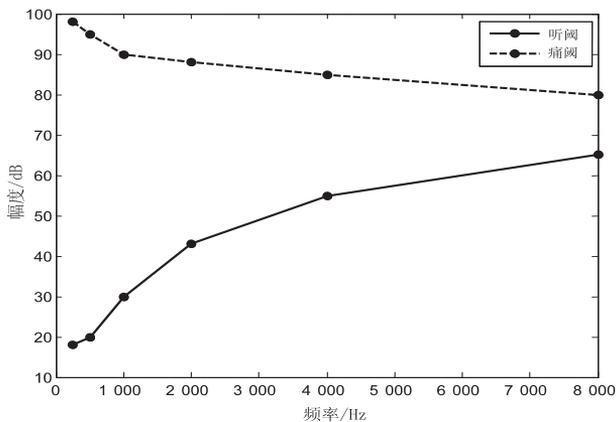


图 1 听损患者的听力检查曲线

从图 1 中可以看到,听损患者在不同频率损伤情况是不一样的,一般来说,在低频区域损伤较小,听阈也较小,在高频区域的听力损失更为严重,使得在高频区域阈值较大,导致高频区域听力范围较小。就算是同一患者,在不同频段上也会有不同程度的听力损伤,因此使用单通道响度补偿方式是不合适的。

对听损患者进行响度补偿,不仅要考虑不同频率信号的听力敏感,还要考虑患者可以承受的信号强度。因此,在对声音进行响度补偿的同时,还要根据声音的强度进行控制,以确保小信号能被病人听清楚,同时保护病人的耳朵在强信号下不受伤害<sup>[7]</sup>。

图 2 是增益控制曲线。其中,  $UCL_n$  和  $UCL_u$  分别为正常人和患者的痛阈值,  $THR_n$  和  $THR_u$  分别为正常

人与患者的听阈值,  $MCL_n$  和  $MCL_u$  分别为正常人和患者的最适阈值。这些参数的值并不是固定的,不同患者在不同频率上的值都是不同的,因此需要根据各患者的具体情况来测定听力测试曲线(见图 1)。根据图 2 的增益控制曲线对响度补偿后的语音进行增益调节,正常听力范围内的声音被映射到听障者的听力范围上。

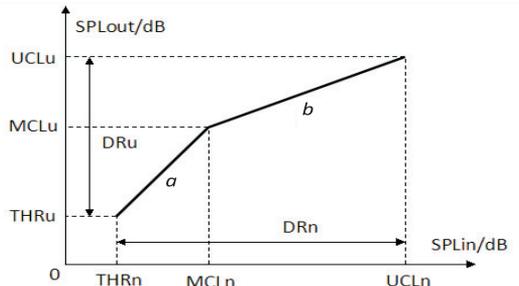


图 2 增益控制曲线

### 1.2 宽动态压缩方法

所谓宽动态压缩,意思是将整个声音按照一定的比例进行均匀压缩,最终压缩到患者的残余听力中。由于声音具有特殊性质,语素之间有一定关系,语素之间的联系可以通过按比例压缩来保持。将语音按比例压缩到患者较好的听力范围内,可以使患者更好地听到声音的高频部分,有助于改善患者的语音识别率<sup>[8]</sup>。

首先根据患者的听力情况,选择要进行压缩的频段和目标频段,然后计算出高频部分的压缩比<sup>[9]</sup>。设压缩比为  $\gamma$ , 则:

$$\gamma = \frac{f_{\text{omax}} - f_{\text{omin}}}{f_{\text{tmax}} - f_{\text{tmin}}} \quad (1)$$

其中:  $f_{\text{tmax}}$  和  $f_{\text{tmin}}$  分别为目标频段的最高频率和最低频率;  $f_{\text{omax}}$  和  $f_{\text{omin}}$  分别为原频段的最高频率和最低频率。

压缩后信号频谱表达式为:

$$\begin{cases} X(f) = 0, f < f_{\text{tmin}} \\ X(f) = X(\gamma \times (f - f_{\text{tmin}}) + f_{\text{omin}}), f_{\text{tmin}} < f < f_{\text{tmax}} \\ X(f) = 0, f > f_{\text{tmax}} \end{cases} \quad (2)$$

图 3 为从实验语音中截取的一帧语音,采用了宽动态压缩方法,选择压缩比  $\gamma = 2$ , 获得压缩前后的频谱比较图。其中,压缩前原始语音的频段为 0 ~ 8 000 Hz,压缩后的目标频段取 0 ~ 4 000 Hz。实线表示原始语音的频谱,虚线表示压缩后语音的频谱。

### 1.3 基于 QMFB 的多通道子带滤波器组设计

从图 1 可以看出,不同频段上的听损程度是不同的,因此需要将语音划分为多个频段,不同频段使用不同的补偿因子。还可以看出,听损患者在 4 000 ~

8 000 Hz 频段上的听力范围较小,而在 0 ~ 4 000 Hz 频段上的听力范围较大。因此先将信号整体压缩到 0 ~ 4 000 Hz 频段上,这样使得 4 000 ~ 8 000 Hz 频段上的信号对应到 2 000 ~ 4 000 Hz 频段上,这时的信号处于患者的残余听力中,最后再使用正交镜像滤波器组(QMFB)对信号进行非等带宽划分<sup>[10-12]</sup>。

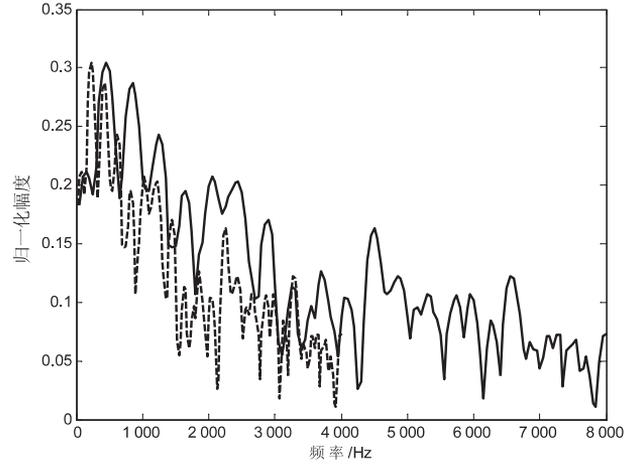


图3 原频谱及压缩后频谱

信号经过 QMFB 的分析,滤波器组被分成若干子带信号,各子带信号经过处理后,再经过合成滤波器组输出合成信号。假设分析滤波器组的低通滤波器为  $H_0(\omega)$ ,高通滤波器为  $H_1(\omega)$ ;合成滤波器组的低通滤波器为  $G_0(\omega)$ ,高通滤波器为  $G_1(\omega)$ <sup>[13]</sup>。若  $H_0(\omega) = H(\omega)$ ,由于  $H_0(\omega)$  和  $H_1(\omega)$  关于  $\omega = \pi/2$  对称,两者频响关系满足:

$$H_1(\omega) = H(\omega - \pi) \quad (3)$$

$$|H_0(\omega)|^2 + |H_1(\omega)|^2 = 1 \quad (4)$$

合成滤波器组跟分析滤波器组的差别只是因子不同,两者结构基本相同。因此合成滤波器组的低通滤波器  $G_0(\omega) = 2H(\omega)$ ,高通滤波器  $G_1(\omega) = -2H(\omega - \pi)$ 。

对于滤波器而言,一般来说,阶数提高了,可以使过渡带变窄,阻带波动变小,当然滤波效果更接近理想情况,缺点是成本较高。经过综合考虑,文中选取阶数  $N = 99$ ,归一化截止频率  $f_p = 0.45\pi$  的滤波器。由于 QMFB 是线性相位的,可以消除混叠失真,能够完美重构信号<sup>[14]</sup>。

### 1.4 系统结构框图

首先对原始语音进行响度衰减,模拟患者听到的语音,然后对该语音进行宽动态压缩,将语音按一定比例均匀地压缩到目标频段,再使用分析滤波器组对语音进行非等带宽划分,并对各子带信号进行响度补偿和增益调节,最后使用合成滤波器组将子带信号合成为最终补偿后的语音。

系统框图如图 4 所示。

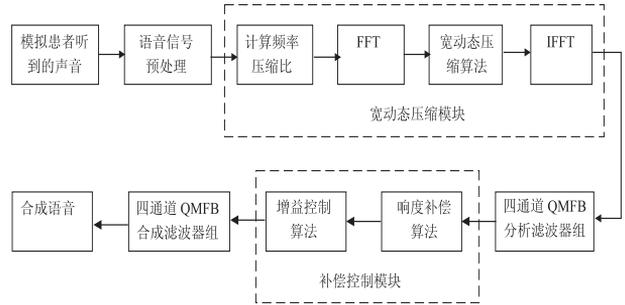


图4 系统框图

具体步骤如下:

第 1 步:根据图 1 听力检查曲线图模拟患者听到的声音。听力曲线图是分段曲线,先计算各分段点的声压级,相邻分段频率点之间是线性关系,可以计算各频段声压级的衰减曲线,然后在各频段内依据得到的衰减曲线进行幅度衰减,从而得到患者听到的声音。声压级计算公式为:

$$SBL = 20 \times \lg \left( \frac{1}{N} \sqrt{\sum_{n=1}^N |x_i(n)|^2} \right) \quad (5)$$

其中:  $x$  是采样信号;  $N$  是信号长度;  $P_0$  是基准声压级,一般取  $20 \mu\text{Pa}$ 。

第 2 步:对语音进行分帧。帧长取 20 ms,帧移为 10 ms,加窗为汉明窗。

第 3 步:对分帧后的语音进行压缩。从图 1 听力检查图可以看到,患者在 4 000 ~ 8 000 Hz 上的听觉范围较小,而在 0 ~ 4 000 Hz 上的听觉范围较大,可以取压缩比  $\gamma = 2$ 。根据式(1)将语音进行压缩,压缩到目标频段 0 ~ 4 000 Hz,这样原始语音在 4 000 ~ 8 000 Hz 的频段就映射到了 2 000 ~ 4 000 Hz 频段上。要根据不同患者的具体情况来选取原频段与目标频段。

第 4 步:对语音进行分频处理。使用四通道 QMFB 分析滤波器组将压缩后的语音帧分频,语音被分成 0 ~ 500 Hz, 500 ~ 1 000 Hz, 1 000 ~ 2 000 Hz 以及 2 000 ~ 4 000 Hz 这四个频段。

第 5 步:对子带信号进行响度补偿和增益控制。子带补偿的方法跟第一步的过程相反。语音经过补偿后,再对其强度进行增益调节。先根据式(5)计算输入信号的声压级,再确定各子带的增益调节因子,也就是  $a$  与  $b$  两段直线的斜率,如图 2 所示。假设  $a$ ,  $b$  两段直线的斜率分别为  $CR_a =$  和  $CR_b$ ,则:

$$CR_a = \frac{MCL_n - THR_n}{MCL_u - THR_u} \quad (6)$$

$$CR_b = \frac{UCL_n - MCL_n}{UCL_u - MCL_u} \quad (7)$$

由图 2 可知,当  $THR_n \leq SPL_{in} < MCL_n$  时,有:

$$SPL_{out} = THR_u + \frac{SPL_{in} - THR_n}{CR_a} \quad (8)$$

当  $MCL_n \leq SPL_{in} < UCL_n$  时,有:

$$SPL_{out} = MCL_u + \frac{SPL_{in} - MCL_n}{CR_b} \quad (9)$$

第 6 步:将子带信号进行合成。先用 QMFB 的合成滤波器组将子带信号合成为一帧完整的信号,再将各帧信号叠加得到响度补偿后的语音,并利用小波阈值去噪技术去除部分噪声。

### 2 实验与仿真

录制一段语音作为实验使用语音。先对语音进行分帧,帧长取 20 ms,帧移取 10 ms。再根据图 1 的听力曲线对语音进行衰减,获得患者听到的声音。然后根据图 4 的系统框图,将衰减语音依次通过宽动态压缩模块、QMFB 分析滤波器组、补偿控制模块以及 QMFB 合成滤波器组,最终获得符合听损患者听觉特性的语音,并利用小波阈值去噪方法去除部分噪声。

图 5 为其中一帧语音经过衰减以及响度补偿后的波形图。

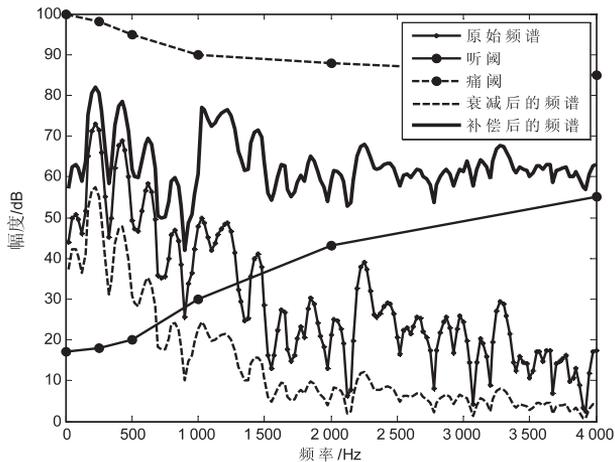


图 5 响度补偿及增益控制图

从图 5 可以看出,衰减后的语音在大多数频段内的响度值都远低于听损患者的听阈值。信号处于低频段时衰减相对较小,处于中高频段时的衰减较大。经过响度补偿后,原先衰减的语音大部分都处于患者的听觉范围,虽然有少部分处于听阈以下,但是经过增益控制后,所有频段的语音都被调节到患者的听觉范围内,有效实现响度补偿。

如图 6 所示,这是原始语音、经过衰减后语音、经过增益控制后语音的时域波形图。

从图 6 可以看出,衰减后的语音幅度非常小,并且波形发生变化,甚至无法分辨出原始语音。而由增益控制后语音波形图可以看出,经过响度补偿和增益控制后的语音与原始语音的波形基本相同,不仅如此,语音的幅度也有了明显的增强,从而实现了语音进行响度补偿的目的。

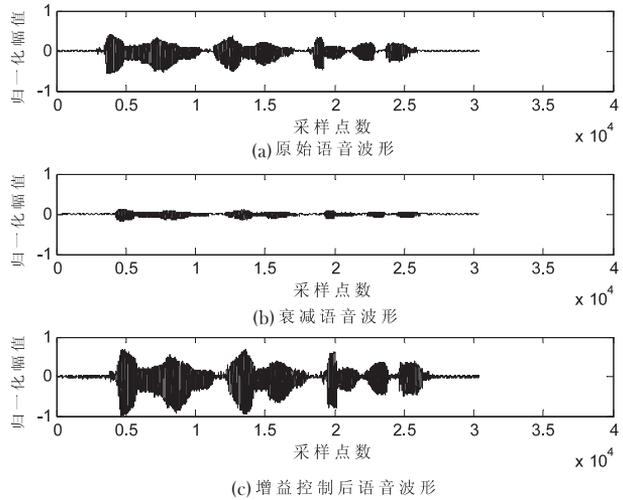


图 6 原语音、衰减语音、增益后语音时域波形图

### 3 结束语

上述实验表明,使用宽动态压缩,可将患者听觉范围较小的高频区域信号移动到患者听觉范围较大的低频区域,同时对该低频区域的语音进行响度补偿,这样可使患者获得较好的听觉效果。补偿后,语音信号基本上与原来的信号相同,而且声音的强度明显增强。再经过增益调整,将处于患者听阈以下的信号调整到患者听阈以上,同时将幅度过大的信号进行适当削弱,这样既可以使声音映射到听损患者的听力范围内,又能防止声音过大对患者造成二次伤害。实验结果表明,该方法能有效实现响度补偿,改善听觉效果,有效提高患者的听力水平。

#### 参考文献:

- [1] 赵力,张昕然,梁瑞宇,等. 数字助听器若干关键算法研究现状综述[J]. 数据采集与处理,2015,30(2):252-265.
- [2] Chong K S, Gwee B H, Chang J S. A 16-channel low-power nonuniform spaced filter bank core for digital hearing aids [J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2006, 53(9):853-857.
- [3] Miller-Hansen D R, Nelson P B, Widén J E, et al. Evaluating the benefit of speech recoding hearing aids in children [J]. American Journal of Audiology, 2003, 12(2):106-113.
- [4] Stelmachowicz P G, Pittman A L, Hoover B M, et al. The importance of high-frequency audibility in the speech and language development of children with hearing loss [J]. Arch Otolaryngol Head Neck Surg, 2004, 130:556-562.
- [5] Goldbaum S M, Halpin C. Exploring the damaged ear: the NIDCD national temporal bone registry [J]. ASHA, 1999, 41(1):29-33.
- [6] 陈广飞,应俊. 数字助听器宽动态压缩算法研究[J]. 北京生物医学工程,2006,25(5):454-456.

求取特征向量,在分类阶段通过交叉验证寻求最佳分类器参数。实验结果表明,基于小波包基与能量熵的 MEG 自动分类方法有较高的分类准确度,分类准确率可达 97.586 8%。

#### 参考文献:

- [1] 李 坤,褚蕾蕾,朱世东,等. 基于 mu 节律能量的运动意识分类研究[J]. 计算机技术与发展,2006,16(8):157-159.
- [2] 罗志增,李亚飞,孟 明,等. 脑电信号的混沌分析和小波包变换特征提取算法[J]. 仪器仪表学报,2011,32(1):33-39.
- [3] 文 念,黄丽亚,于 涵,等. 基于 ICA 和聚类的 EEG 脑源定位研究[J]. 计算机技术与发展,2015,25(5):228-232.
- [4] Zhukov L, Weinstein D, Johnson C. Independent component analysis for EEG source localization[J]. IEEE Engineering in Medicine and Biology Magazine,2000,19(3):87-96.
- [5] 侯凤贞,戴加飞,刘新峰,等. 基于网络连接度指标的脑梗死患者脑电信号相同步分析[J]. 物理学报,2014,63(4):59-64.
- [6] Huberle E, Lutzenberger W. Temporal properties of shape processing by event-related MEG adaptation[J]. NeuroImage, 2013,67:119-126.
- [7] Cravo A M, Rohenkohl G, Wyart V, et al. Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex[J]. The Journal of Neuroscience,2013,33(9):4002-4010.
- [8] Mirzaei A, Ayatollahi A, Gifani P, et al. EEG analysis based on wavelet-spectral entropy for epileptic seizures detection[C]//Proc of 3rd international conference on biomedical engineering and informatics. [s. l.]:IEEE,2010:878-882.
- [9] 王 登,苗夺谦,王睿智. 一种新的基于小波包分解的 EEG 特征抽取与识别方法研究[J]. 电子学报,2013,41(1):193-198.
- [10] 王 攀,沈继忠,施锦河. 基于小波变换和时域能量熵的 P300 特征提取算法[J]. 仪器仪表学报,2011,32(6):1284-1289.
- [11] Kremer J, Pedersen K S, Igel C. Active learning with support vector machines[J]. Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery,2014,4(4):313-326.
- [12] Hsu W Y, Lin C H, Hsu H J, et al. Wavelet-based envelope features with automatic EOG artifact removal: application to single-trial EEG data[J]. Expert Systems with Applications, 2012,39(3):2743-2749.
- [13] Coifman R R, Wickerhauser M V. Entropy-based algorithms for best basis selection[J]. IEEE Transactions on Information Theory,1992,38(2):713-718.
- [14] 张德丰. MATLAB 小波分析[M]. 北京:机械工业出版社,2009:158-163.
- [15] Clochon P, Fontbonne J M, Lebrun N, et al. A new method for quantifying EEG event-related desynchronization: amplitude envelope analysis[J]. Electroencephalography and Clinical Neurophysiology,1996,98(2):126-129.
- [16] 胡剑锋,穆振东,肖 丹. 基于能量熵的运动想象脑电信号分类[J]. 计算机工程与应用,2008,44(33):235-238.
- [17] 陈 悦,张少白. LM 算法在神经网络脑电信号分类中的研究[J]. 计算机技术与发展,2013,23(2):119-122.
- [18] Cortes C, Vapnik V. Support-vector networks[J]. Machine Learning,1995,20(3):273-297.
- [19] Zhang X J, Wang Y T, Wang Y J, et al. Ultra-slow frequency bands reflecting potential coherence between neocortical brain regions[J]. Neuroscience,2015,289:71-84.
- [20] Lin C T, Wang Y K, Chen S A. An EEG-based brain-computer interface for dual task driving detection[C]//Proc of neural information processing. Berlin:Springer,2011:701-708.
- [21] Nakayama K, Kaneda Y, Hirano A. A brain computer interface based on FFT and multilayer neural network-feature extraction and generalization[C]//Proc of ISPACS. Xiamen:IEEE, 2008:826-829.
- (上接第 126 页)
- [7] 雍雅琴. 数字助听器中主要语音信号处理方法研究[D]. 北京:北京协和医学院,2013.
- [8] 王青云,赵 力,赵立业,等. 一种数字助听器多通道响度补偿方法[J]. 电子与信息学报,2009,31(4):832-835.
- [9] 张宝琳,张玲华. 数字助听器中多通道响度补偿方法的研究[J]. 信号处理,2013,29(5):656-661.
- [10] Rothweiler J H. Polyphase quadrature filters—a new subband coding technique[C]//Proc of IEEE ICASSP. [s. l.]:IEEE, 1983.
- [11] 张 晨,李双田. 设计 QMF 组的一种新算法及基于 GA 的优化[J]. 信号处理,2005,21(2):120-125.
- [12] 孟 君. 基于 DSP 的数字助听器多通道响度补偿方案[J]. 南京信息工程大学学报:自然科学版,2010,2(5):420-425.
- [13] 李 丽. 非均匀滤波器组的研究[D]. 西安:西安电子科技大学,2009.
- [14] Huang J, Gu G X. A direct approach to the design of QMF banks via frequency domain optimization[J]. IEEE Transaction on Signal Processing,1998,46(8):2131-2138.