

基于Bark域的电子耳蜗频带划分分析和拟合研究

陈又圣,曹智亮,夏林中,张春晓,林炜臻
深圳信息职业技术学院,广东 深圳 518000

【摘要】目前电子耳蜗主流的言语处理策略是基于滤波器组的言语处理算法,该算法通过分频带进行信号处理并把参数传递到对应的电极上。电子耳蜗滤波器组的频带划分不是等分的,而是按一定规律进行并且符合人耳听觉特性的,其中,Bark域的频带划分是重要的参考。本研究基于Bark域的电子耳蜗频带划分方法,探讨Bark域频带划分的特性并结合目前的电子耳蜗滤波器组的频带划分进行分析,进而探讨频带划分中的曲线拟合方法,为电子耳蜗滤波器组中的频带划分提供重要的方法和参数。

【关键词】电子耳蜗;曲线拟合;频带划分;滤波器组

【中图分类号】R318;TH776

【文献标志码】A

【文章编号】1005-202X(2020)03-0350-05

Frequency band division analysis and fitting based on Bark domain for cochlear implant

CHEN Yousheng, CAO Zhiliang, XIA Linzhong, ZHANG Chunxiao, LIN Weizhen
Shenzhen Institute of Information Technology, Shenzhen 518000, China

Abstract: At present, the mainstream speech processing strategy for cochlear implants is based on the speech processing algorithm of filter banks. The signals processing by the algorithm is based on frequency band division and the parameters are transferred to the corresponding electrodes. The frequency band division of cochlear filter banks is not equal, but accords to certain rules and conforms to the human auditory characteristics. Among them, the frequency band division based on Bark domain is an important reference. Herein the method of frequency band division based on Bark domain for cochlear implants is studied. The characteristics of frequency band division based on Bark domain are discussed, and the current frequency band division is analyzed. Then the curve fitting method in frequency band division is further discussed, which provides important methods and parameters for frequency band division of cochlear filter banks.

Keywords: cochlear implant; curve fitting; frequency band division; filter bank

前言

Bark域是由Eberhard Zwicker提出的一种心理声学尺度和信号频带划分方法。在人耳中,存在24个频率位置会产生共振,因而可以划分出从1到24的共24个临界频带,即Bark域。Bark域对频率不均匀划分的思想,有助于对低频信号给予更高的分辨率,对高频信号给予相对较低的分辨率,在噪声分类和抑制^[1-3]、特征分析和特征提取^[4-7]、助听器^[8]、语音增强^[9-12]、语

音测试和评估^[13]、语音识别^[14]、语音端点检测^[15]等方面都有广泛应用。在电子耳蜗领域应用中,Bark域可用于信号处理和言语处理策略的设计^[16-17]。电子耳蜗包括体外器件和体内植入体两部分,体外器件的主要部分是言语处理器,该处理器进行频带划分、参数提取和信号调制等过程^[18-20]。目前电子耳蜗的频带划分也参考了Bark域的心理声学尺度,按照低频划分密、高频划分疏的方式来实现。但是,目前的频带划分是固定的,对于确定电极数,滤波器所划分的频带和子带中心频率是固定的,不利于针对不同采样率参数来设置不同的频带划分模式。另一方面,目前的电子耳蜗电极数不超过24个,但是,随着技术的进步和电极制作工艺的改善,未来植入电极数有望增加,所对应的体外言语处理器的滤波器组需要划分更多的频带。Bark域各个临界频带所对应的是离散的频率点,可对离散的频率点进行曲线拟合,然后通过插值方式来实现针对不同参数进行自适应的频带划分。Schroeder、

【收稿日期】2019-10-11

【基金项目】广东省自然科学基金(2018A030313146);2019年度广东省普通高校特色创新类项目(2019GKTSX094);广东大学生科技创新培育专项资金项目,攀登计划专项(pdjh2020b1197, pdjh2019b0866);深圳信息职业技术学院人工智能技术应用工程实验室项目(PT201701)

【作者简介】陈又圣,博士,讲师,主要研究方向:语音增强、麦克风阵列波束形成、电子耳蜗言语处理、医疗仪器设计,E-mail: chenyoushengtsinghua@aliyun.com

Zwicker、Sekey、Gersho等众多学者早期提出Bark域的拟合方法,但这些拟合公式不够精简,例如拟合公式中所含有的反三角函数不利于电子耳蜗在言语处理器中的计算实现。另外,目前的有关Bark域的拟合公式主要是从曲线本身进行分析,并没有根据电子耳蜗应用场景中的低频频段和低频频段不同特征来进行拟合。本文基于Bark域的特征研究电子耳蜗频带划分的方法,通过对Bark域进行拟合来形成简化的频带-频率分布曲线,并提出适合应用于电子耳蜗的频带划分方法。

1 Bark域的频带划分及拟合

Bark域所划分的24个临界频带及对应的中心频率如表1所示。

表1 Bark域的24个临界频带及对应的中心频率
Tab.1 Twenty-four critical frequency bands in Bark domain and the corresponding center frequencies

临界频带编号	所在频带的中心频率/Hz
1	60
2	150
3	250
4	350
5	450
6	570
7	700
8	840
9	1 000
10	1 170
11	1 370
12	1 600
13	1 850
14	2 150
15	2 500
16	2 900
17	3 400
18	4 000
19	4 800
20	5 800
21	7 000
22	8 500
23	10 500
24	13 500

目前的文献对Bark域有多种方式进行拟合,但表达式复杂,本文提出根据Bark域的低频频段和高频频段进行不同处理的方式来进行拟合。用临界频带编号为横坐标(用 n 表示),所在频带的中心频率为纵坐标(用 f_{cen} 表示)来探讨其函数关系,如图1所示。

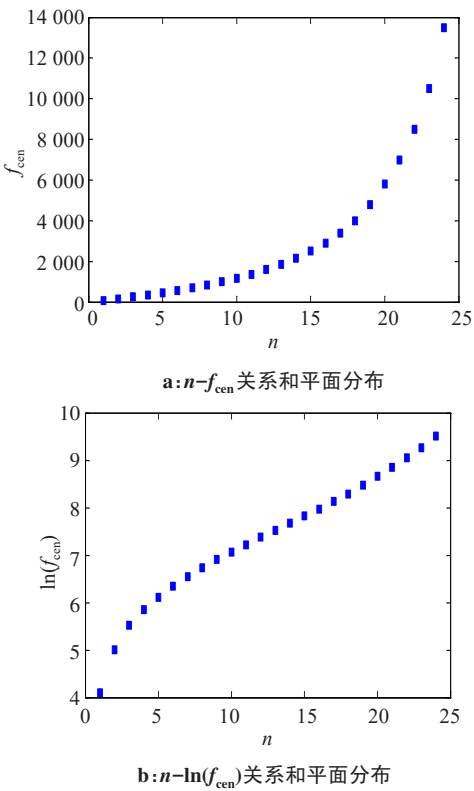


图1 临界频带编号(n)和频带的中心频率(f_{cen})的关系和平面分布
Fig.1 Relationship and planar distribution of critical frequency band number (n) and the center frequency (f_{cen})

图1a表示临界频带编号(n)和频带的中心频率(f_{cen})的函数关系, n 越大, f_{cen} 也越大。从图中进一步观察可以看到,低频频段的 n 和 f_{cen} 具有较好的线性关系,高频频段的 n 和 f_{cen} 则呈现指数关系。如果对 f_{cen} 取对数,则得到图1b中的 n 和 $\ln(f_{cen})$ 关系,从图中可以看到, f_{cen} 取自然对数后的高频频段具有良好的线性关系,而低频频段则不具备线性关系。

从图1可知,低频频段和高频频段的 n 和 f_{cen} 呈现不同的函数关系。通过比较和计算图1a和图1b中点的分布,可以设置一个频率分界点来划分出高低频段的具体位置,让图1a中的低频各点和图1b中的高频各点都近似线性分布,进而可以拟合曲线的简化。通过计算和比较,第9频带的频率值(数值为1 000 Hz)可作为分界点,通过该频率点划分后的各频率点的分布呈良好的线性分布,如图2所示。从图2a可以看出, n 和 f_{cen} 在1~9频带具有非常好的线性关系,从图2b可以看出, n 和 $\ln(f_{cen})$ 在9~24频带具有非常好的线性关系。因此,可以

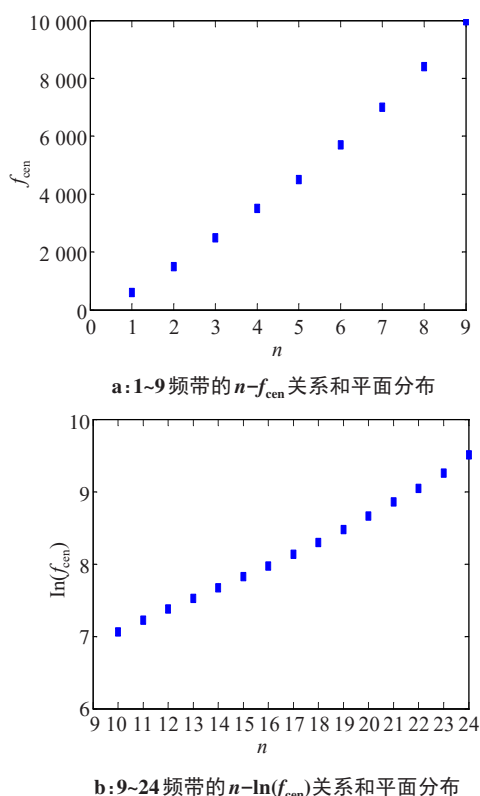


图2 选取第9频带划分出高低频带后的临界频带编号(n)和中心频率(f_{cen})的关系和平面分布

Fig.2 Relationship and planar distribution of critical frequency band number (n) and the center frequency (f_{cen}) after dividing the high and low bands from the 9th frequency band

把频带分成低频段和高频段分别进行近似简化,并拟合出Bark域的 $n \sim f_{cen}$ 函数。

在图2a中,通过两个端点(1, 60)和(9, 1 000)计算出直线的函数解析式。用 n 表示自变量, $f_{cen}(n)$ 表示函数(因变量),则可以得到通过端点(1, 60)和(9, 1 000)的直线解析式,如式(1)所示:

$$f_{cen}(n) = 117.5n - 57.5 \quad (1)$$

在图2b中,以 n 为自变量, $\ln(f_{cen}(n))$ 为因变量的直线经过两个端点(9, $\ln 1000$)和(24, $\ln 13500$), n 为和 f_{cen} 满足以下的关系:

$$\frac{\ln(f_{cen}(n)) - \ln 1000}{n - 9} = \frac{\ln 13500 - \ln 1000}{24 - 9} \quad (2)$$

对式(2)经过简化并求解 $f_{cen}(n)$ 的表达式,可以得到式(3):

$$f_{cen}(n) = 1000 \times 13.5^{\frac{n-9}{15}} \quad (3)$$

综合式(1)和式(2)可以得到经过简化和拟合后的Bark域临界频带编号(n)和频带中心频率(f_{cen})的关系表达式,如式(4)所示:

$$f_{cen}(n) = \begin{cases} 117.5n - 57.5 & (n \leq 9) \\ 1000 \times 13.5^{\frac{n-9}{15}} & (n > 9) \end{cases} \quad (4)$$

2 拟合误差及分析

本文所提出的对Bark域的拟合表达式具有简洁的优点,并且从公式的推导和分析中对低频和高频有明确不同的处理模式,探讨了心理声学尺度和信号频带划分,拟合后的频率分布对比如图3所示。从图3可以看出,本文提出对Bark域的拟合方法所得到的各个频带的频率值与Bark域本身的值非常接近。该方法的拟合公式简单,能保持足够的准确度,并且对低频频段的信号和高频频段的信号有不同的分析模式,可应用于电子耳蜗的频带划分。

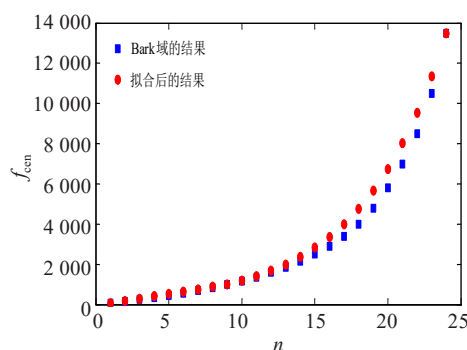


图3 曲线拟合后的中心频率(f_{cen})与Bark域的对比

Fig.3 Comparison of the center frequency (f_{cen}) of the frequency band after curve fitting with that in Bark domain

3 电子耳蜗的频带划分及分析

Bark域是离散的点,通过拟合的曲线可用于自适应地确定所划分的频带。目前电子耳蜗所划分的频带基本是固定的,例如,诺尔康的18通道电子耳蜗的中心频率分别是:213.0、335.0、456.5、578.0、700.0、822.0、944.0、1 066.0、1 249.0、1 523.0、1 858.0、2 254.0、2 711.0、3 259.5、3 960.5、4 814.0、5 849.5和7 098.5 Hz;而22通道电子耳蜗的中心频率分别是:213.0、335.0、456.5、578.0、700.0、822.0、944.0、1 066.0、1 218.5、1 401.0、1 583.5、1 797.0、2 071.5、2 406.5、2 772.0、3 168.5、3 625.5、4 173.5、4 813.5、5 514.5、6 337.0和7 281.5 Hz。可以看到,诺尔康的18通道电子耳蜗和22通道电子耳蜗的滤波器组所对应的各子带的中心频率是固定的,而且前8个通道的频率值是相同的。

对于采样率可变的情形来说,采用固定的频率就会不合适,而本文所提出的Bark域拟合曲线可适用于采样率变化的情形。对于采样率 f_s ,按照采样定理,信号的最高频率不能超过 $0.5f_s$,只要在 $0 \sim 0.5f_s$ 范围内选取一个最低频率 f_L 和最高频率 f_H ,可以在此范围内通过拟合曲线确定中间的其他通道的频率值。由于本文通过离散点拟合出曲线,因此式(4)在具体

应用时 n 可以取非整数值。本文提出的对电子耳蜗进行频带划分的方式可以依据拟合的Bark曲线进行,具体的算法流程通过以下几步实现:

第一步: $0 \sim 0.5 f_s$ 范围内选取 f_L 和 f_H ,具体选取时要求 $0 < f_L < 1\ 000, 1\ 000 < f_H < 0.5 f_s$ 。

第二步: 根据拟合曲线及式(4)计算得到 f_L 和 f_H 对应的 n_L 和 n_H 值,具体计算公式如式(5)和式(6)所示:

$$n_L = \frac{f_L + 57.5}{117.5} \quad (5)$$

$$n_H = 15 \log_{13.5} \left(\frac{f_H}{1\ 000} \right) + 9 \quad (6)$$

第三步: 根据通道数 i 确定中间的 $i-1$ 个 n 的值,计算公式如式(7)所示:

$$n_j = n_L + \frac{n_H - n_L}{i - 1} \times (j - 1), \quad j = 1 \sim i \quad (7)$$

第四步: 根据各个 n_j 的值代入式(4)计算出对应的频率值 f_j ,进而可以得到全部的 i 通道电子耳蜗滤波器组的子带中心频率值。

从上述频带划分的步骤和具体的实现方式来看,本文提出的频带划分和曲线拟合方法是高效的。目前已有的Bark域的拟合方法,一方面,计算量更大,另一方面,文献检索并未查询到有学术论文阐述有关电子耳蜗频带划分的具体实现方式。因此,上述的方法可为自适应频带划分提供思路,也为未来更加微型化电极和更多电极条件下匹配的滤波器组的频带划分提供实施方案。

4 结 论

本文研究Bark域的心理声学尺度特征和信号频带划分模式,并提出基于Bark域的曲线拟合算法。本文所提出的算法具有简洁、准确度高、易于算法实现等优点。Bark域的拟合曲线对低频频段采取线性化的处理方式,对高频频段采取指数化的处理方式,并将拟合曲线用于电子耳蜗滤波器组的频带划分。本文所提出的基于Bark域的拟合曲线的电子耳蜗频带划分算法具有自适应的特征,并结合心理声学尺度特征,可用于多模式复杂场景的频带划分,相对于固定频率的划分模式具有更强的适应性。

【参考文献】

[1] 袁文浩,林家骏,陈宁,等. 一种基于Bark域能量分布的噪声分类方法[J]. 华东理工大学学报(自然科学版), 2013, 39(4): 472-476.
YUAN W H, LIN J J, CHEN N, et al. A noise classification method based on noise energy distribution in Bark domain[J]. Journal of East China University of Science and Technology (Natural Science Edition), 2013, 39(4): 472-476.

[2] 田玉静,左红伟,董玉民,等. Bark子带小波包自适应阈值语音去噪方法[J]. 计算机应用, 2010, 30(11): 3111-3114.
TIAN Y J, ZUO H W, DONG Y M, et al. Adaptive threshold speech denoising based on Bark scale wavelet package [J]. Journal of Computer Applications, 2010, 30(11): 3111-3114.

[3] 田玉静,董玉民,左红伟. 时频结合自适应阈值小波包消噪算法[J]. 应用声学, 2010, 29(4): 256-262.
TIAN Y J, DONG Y M, ZUO H W. A time-frequency combination adaptive threshold algorithm of wavelet package de-noising [J]. Applied Acoustics, 2010, 29(4): 256-262.

[4] 李响,李国正,石俊刚,等. 基于语音心理声学分析的驾驶疲劳检测[J]. 仪器仪表学报, 2018, 39(10): 167-175.
LI X, LI G Z, SHI J G, et al. Fatigue driving detection based on speech psychoacoustic analysis [J]. Chinese Journal of Scientific Instrument, 2018, 39(10): 167-175.

[5] 王晓华,屈雷,张超,等. 基于Fisher比的Bark小波包变换的语音特征提取算法[J]. 西安工程大学学报, 2016, 30(4): 452-457.
WANG X H, QU L, ZHANG C, et al. Speech feature extraction algorithm based on the Bark wavelet packet transform with Fisher [J]. Journal of Xi'an Polytechnic University, 2016, 30(4): 452-457.

[6] 潘欣裕,赵鹤鸣. 中文耳语音的声调特征研究[J]. 信号处理, 2011, 27(10): 1525-1530.
PAN X Y, ZHAO H M. Research on tone feature of Chinese vowel whispers [J]. Signal Processing, 2011, 27(10): 1525-1530.

[7] 尹许梅,何逸森. 基于Bark子波变换的MFCC特征提取[J]. 计算机工程, 2011, 37(11): 192-194.
YIN X M, HE X S. MFCC feature extraction based on Bark wavelet transform [J]. Computer Engineering, 2011, 37(11): 192-194.

[8] 李如玮,曹长涛,鲍长春,等. 基于Gammatone滤波器组分解的数字助听器频响补偿算法[J]. 北京生物医学工程, 2016, 35(2): 143-150.
LI R W, CAO L T, BAO C C, et al. Loudness compensation algorithm based on Gammatone filter bank decomposition for digital hearing aid [J]. Beijing Biomedical Engineering, 2016, 35(2): 143-150.

[9] 任永梅,张雪英,贾海蓉. 一种新阈值函数的小波包语音增强算法[J]. 计算机应用研究, 2013, 30(1): 114-116.
REN Y M, ZHANG X Y, JIA H R. New threshold function of wavelet packet speech enhancement algorithm [J]. Application Research of Computers, 2013, 30(1): 114-116.

[10] 李野,吴亚锋,刘雪飞. 基于BARK子波变换的语音增强方法研究[J]. 计算机工程与应用, 2009, 45(33): 128-131.
LI Y, WU Y F, LIU X F. Research on speech enhancement based on BARK wavelet transform [J]. Computer Engineering and Applications, 2009, 45(33): 128-131.

[11] 赵欢,熊敏,侯卫国. 基于Bark域噪声估计及掩蔽效应的语音增强[J]. 计算机工程, 2009, 35(12): 261-263.
ZHAO H, XIONG M, HOU W G. Speech enhancement based on noise estimation and masking property in Bark domain [J]. Computer Engineering, 2009, 35(12): 261-263.

[12] NIKOLAOS D, MIKE B. Speech enhancement using Kalman filtering in the logarithmic Bark power spectral domain [C]//2018 26th European Signal Processing Conference (EUSIPCO). Rome, 2018: 1642-1646.

[13] 汤霖,彭土有,尹俊勋. 普通话水平客观测试中的韵母测试研究[J]. 湘潭大学自然科学学报, 2012, 34(1): 95-100.
TANG L, PENG T Y, YIN J X. Study of objective evaluation of final in putonghua proficiency testing [J]. Natural Science Journal of Xiangtan University, 2012, 34(1): 95-100.

[14] 刘菁华. 基于Bark子波变换和RAS-MFCC特征的语音识别系统研究[J]. 语音技术, 2011, 35(9): 47-54.
LIU J H. Speech recognition system based on the Bark wavelet and

- RAS-MFCC features[J]. Voice Technology, 2011, 35(9): 47-54.
- [15] 尹晨晓, 郭英, 张碧锋, 等. 基于Bark小波的语音端点检测算法[J]. 计算机工程, 2011, 37(12): 276-278.
- YIN C X, GUO Y, ZHANG B F, et al. Voice activity detection algorithm based on Bark wavelet[J]. Computer Engineering, 2011, 37(12): 276-278.
- [16] SHEN Y, TAO Z, GU J H, et al. A novel signal-processing strategy based on Bark wavelet transform for cochlear implant [C]//2011 First International Workshop on Complexity and Data Mining. Nanjing, 2011: 10-13.
- [17] WEI M F, WANG H Q, FAN Q Y, et al. Improvement of cochlear implant coding strategy based on Chinese speech boundary Information [C]//2018 International Conference on Information Systems and Computer Aided Education (ICISCAE). Changchun, 2018: 402-405.
- [18] CHEN Y S, CHEN Y, XUE G W. ACE speech processing strategy and feature analysis in CI device [C]//2018 IEEE 3rd International Conference on Integrated Circuits and Microsystems (ICICM). Shanghai, 2018: 318-321.
- [19] SUNGMIN L, SARA A, SATNAM S, et al. A speech processing strategy based on sinusoidal speech model for cochlear implant users [C]//2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). Honolulu, 2018: 393-397.
- [20] ZENG F G. Challenges in improving cochlear implant performance and accessibility[J]. IEEE Trans Biomed Eng, 2017, 64(8): 1662-1664.
- (编辑:陈丽霞)