

基于信号特征的电子耳蜗言语处理策略的频带选取

陈又圣, 杜军

深圳信息职业技术学院电子与通信学院, 广东 深圳 518000

【摘要】言语处理策略提取信号参数,并以特定的编码和刺激模式把参数传递到植入体内的电子阵列,是电子耳蜗各个环节中的关键部分。目前不同类型的电子耳蜗采用不同的言语处理策略,其中基于信号特征的言语处理策略是主流的方法。该模式进行参数提取和频带选择是基于所采集信号本身的特征,其刺激模式与连续交替采样刺激模式不同。本研究基于信号特征电子耳蜗言语处理策略的频带选取模式和特征,探讨频带选取对信号的影响,并总结基础的参数结果和信号变化结论,有助于电子耳蜗新算法的开发和改进。

【关键词】言语处理策略;频带选取;电子耳蜗;信号特征;频谱

【中图分类号】R318;TH776

【文献标志码】A

【文章编号】1005-202X(2019)12-1461-05

Frequency-band selection for signal feature-based speech processing strategy in cochlear implant

CHEN Yousheng, DU Jun

School of Electronics and Communication Technology, Shenzhen Institute of Information Technology, Shenzhen 518000, China

Abstract: Speech processing strategy, one of the key parts in cochlear implant, can extract signal parameters and transmit parameters to the inner electrode array through particular encodings and stimulating patterns. At present, different types of cochlear implants adopt different speech processing strategies, and the signal feature-based speech processing strategy is the most common method. Both parameter extraction and frequency-band selection of the signal feature-based speech processing strategy are based on the characteristics of the collected signals themselves, and the corresponding stimulating mode is different from continuous interleaved sampling stimulating mode. Based on the frequency-band selection modes and characteristics of signal feature-based speech processing strategy in cochlear implant, the effects of frequency-band selection on signals are discussed, and the analyses on basic parameters and signal changes are summarized. The research is helpful for developing and improving new algorithms for cochlear implant.

Keywords: speech processing strategy; frequency-band selection; cochlear implant; signal feature; frequency spectrum

前言

据世界卫生组织官网报告(数据更新日期为2018年3月15日),全世界患有残疾性听力损失的人数高达4.66亿,其中包含0.34亿儿童患者,并且听力损失的患者在迅速增长,到2050年,患有残疾性听力损失的人数将会超过9亿。2017年,第70届世界卫生大会通过一项有关预防耳聋以及听力损失的决议,协助世界卫生组织的成员制定相应的规划,并把耳部和听力初级保健纳入本国的卫生保健系统。

不同的听力损失情况可以采用不同的治疗方式,例如药物治疗、佩戴助听器、植入电子耳蜗等,其中,对于重度耳聋患者和极重度耳聋患者来说,植入电子耳蜗是目前最为有效的方式。电子耳蜗既可以单独使用,也可以和助听器进行结合并形成双耳的模式^[1-5]。电子耳蜗处于噪声环境以及应用场景中的各类失配都会导致言语识别率的大幅度降低,为了提高电子耳蜗的言语识别率,近年来学者们在电流导引^[6]、电子耳蜗的精细结构编码^[7]、虚拟电极^[8-9]、光学耳蜗^[10-11]、麦克风阵列^[12-13]等方面做了很多努力和研究工作。

电子耳蜗包括体外和体内两个模块,其中体外部分通过麦克风采集声音信号,传入言语处理器中,信号被特定的电子耳蜗言语处理策略处理,进行参数提取、信号调制、编码、压缩等系列过程,然后通过

【收稿日期】2019-07-08

【基金项目】广东省自然科学基金(2018A030313146)

【通信作者】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

无线方式把有效信号和能量传输到体内植入体中,体内植入体的线圈接收信号,并通过解码等过程把目标参数提取并传输到电极阵列上,电极阵列以特定的刺激参数和模式刺激对应位置的听觉神经并让听力患者产生听觉感知,恢复部分的听力。目前国外电子耳蜗的生产商主要有3家,即澳大利亚的COCHLEAR公司、奥地利的MED-EL公司和美国的AB公司,而国内则只有诺尔康公司推出电子耳蜗临床产品,其中,COCHLEAR公司是全球最大的电子耳蜗生产商,占了全球市场份额的60%~70%。由于电子耳蜗的复杂性,其价格一直很高,限制了其应用的广度和深度。据报导,目前电子耳蜗的市场渗透率在发达国家也只有20%,而在发展中国家则不到1%。近几年这种情况有所改观,有学者研究和统计指出^[14],近年来3家国外公司(COCHLEAR、MED-EL和AB)和一家国内公司(诺尔康)在中国政府投标价格中有了大幅度的下降,从2011年的每个25 000美元到2016年的6 030美元再到2017年的5 490美元。价格的大幅度降低有助于促进电子耳蜗在中国大陆的进一步使用,但是汉语本身是音调语言,音调对汉字的识别比英语重要得多,因此,直接把国外电子耳蜗中的言语处理策略移植并应用在汉语人群中是不适合的。本研究选取目前广泛使用的基于信号特征

1 基于信号特征

的电子耳蜗言语处理策略以及信号采集实验的设计

从大类来说,电子耳蜗的言语处理策略可以划分为基于滤波器组的算法和不基于滤波器组的算法,其中,基于滤波器组的算法是目前的主流方法。而从刺激模式来看,基于滤波器组电子耳蜗的言语处理策略可以进一步划分为非信号特征的模式(即规律性循环刺激模式)和基于信号特征

刺激模式。规律性循环刺激模式与信号本身的特性无关,该言语处理策略是按预先设计好的刺激模式进行,例如连续交替采样模式^[15-17]、连续交替等间隔刺激模式^[18]。而基于信号特征

的刺激模式是基于信号本身的实时变化特征的,例如可以选取频带中若干个最大的频带用于电极刺激,常见模式包括SPEAK策略^[19]、ACE策略^[20]、n of m策略、FFC策略、信号精细结构策略等。目前研究领域和应用领域应用最广泛的是基于信号特征的言语处理策略,而该类策略的主要特点是基于信号特征的频带选取模式,本文的实验测试和参数分析是基于该类言语处理策略进行研究的。

2 频谱特性及参数变化

用所搭建的硬件系统采集目标信号,保存在计算机,然后分别用Matlab所编写的电子耳蜗言语处理进行信号处理,生成并合成用于表征原始信号编码和传入电极阵列的新信号,进而用于算法比较。本研究选用目前应用最为广泛的两种基于信号特征

的电子战耳蜗言语处理策略,即ACE策略和SPEAK策略,来进行算法研究和参数比较。其中,ACE策略选用常见的22通道滤波器组并选取8个频带用于电极刺激,而SPEAK选用常见的20通道滤波器组并选取6个频带用于电极刺激。

原始采集的目标信号、经过ACE策略和SPEAK策略处理后并合成的新信号,它的信号频谱图的对比如图1所示。

图1表示信号经过ACE策略和SPEAK策略处理后的信号和原始信号的对比。从图1a可以看到,原始信号在每个时间帧均包含不同频率成分的信号,其中红色是高能量的部分,颜色越深对应的能量越大。而对比图1b和图1c可以看到,ACE策略和SPEAK策略处理后的信号在每个时间帧开始集中在特定的几个频率上。在图1b和图1c中标记了对应滤波器组每个频带的中心频率,通过比较可以看到,经过算法处理后的信号能量集中在对应的若干频率位置上,该频率位置对应所在滤波器组每个通道的中心频率。而进一步比较图1b和图1c中每个时间帧所选取的频带来看,算法选取的是能量最大的若干个频带,频谱中显示的是红色和深红色区域,其中,

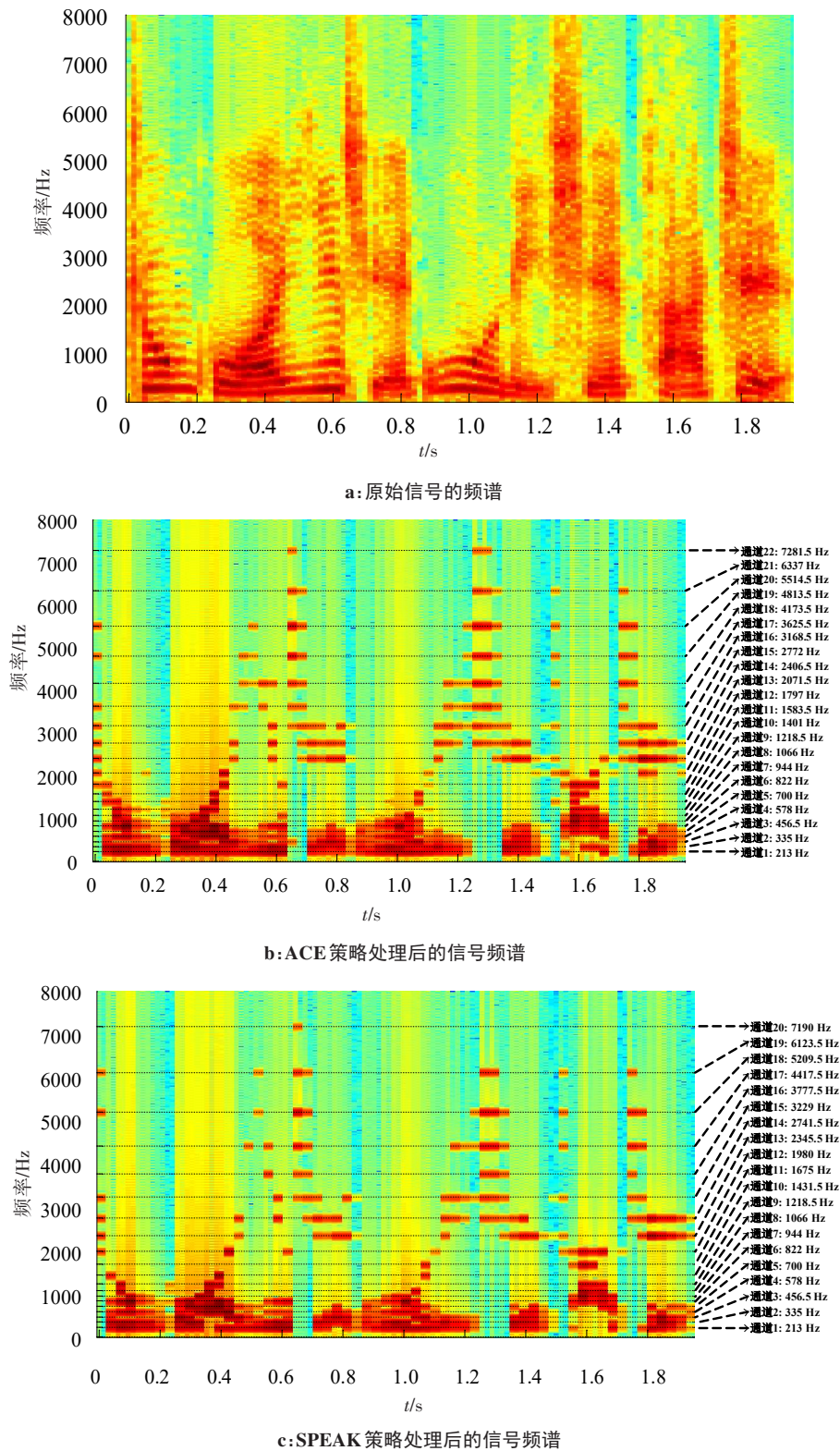


图1 信号的频谱比较

Fig.1 Comparison of signal frequency spectra

ACE策略选取的是能量最大的8个通道的信号,而SPEAK策略选取的是能量最大的6个通道的信号。对于不同的时间帧来说,由于原始信号的能量分布随着时间的变化而不同,因此,算法所提取的频带也不同,频带的选取与原始信号的能量强度分布有关。

3 频带选取特性及分析

不同时间帧所选取的信号的频带不同,可以对所选取的各个频带进行统计。在所采集的语料为“中国医学物理学杂志”的语音信号里,该语音信号的长度为1.95 s,算法中的每一帧的长度为512点,整

个信号一共划分为61帧。然后对每一帧所选取的频带编号进行统计和分析,可以发现以下的规律:ACE策略中每一个时间帧选取8个频带,其中低频的频带经常被选取,同样,SPEAK策略中每一个时间帧选取6个频带,也是低频的频带经常被选取。进一步统计每个频带被选取的次数并由大到小排列,则ACE策略中被选取次数最多的前5个频带依次是:通道2、通道1、通道4、通道5、通道3,对应被选取的次数分别是:51、50、45、44、43。而SPEAK策略中被选取次数最多的前5个频带依次是:通道1、通道2、通道3、通道4、通道5,对应被选取的次数分别是:48、45、36、35、30。由于实验中的语音信号共划分为61帧,因此,计算得到ACE策略的通道1、通道2、通道3、通道4、通道5的被选取的概率分别是:81.97%、83.61%、70.49%、73.77%和72.13%,SPEAK策略的通道1、通道2、通道3、通道4、通道5的被选取的概率分别是:78.69%、73.77%、59.02%、57.38%和49.18%。因此,从分析和统计上看,基于信号特征电子耳蜗言语处理策略在频带选取上倾向于选取低频的频带,这种频带选取的模式一方面可以保留语音中低频的较高的能量信息,另一方面容易丢失中高频的细致信息。

从前面的分析看到,低频的频带容易被选择,而中频和高频的频带选取情况也有所差异,进一步详细统计各个频带的选取概率,其频带选取概率曲线如图2所示。

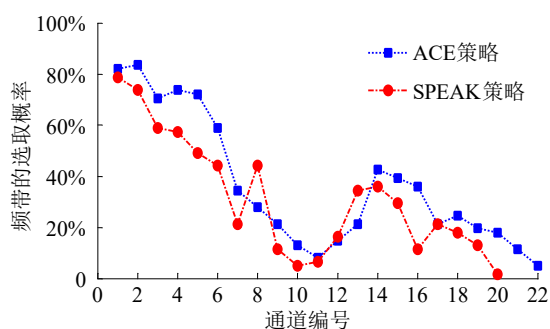


图2 ACE策略和SPEAK策略的频带选取概率曲线

Fig.2 Probability curves of frequency-band selection for ACE strategy and SPEAK strategy

图2为ACE策略和SPEAK策略的频带选取概率曲线,从曲线上可以看到,低频通道所对应的频带选取概率较高,中频和高频的选取概率较低。另一方面,两组曲线均在中间位置有一个下凹,表示ACE策略和SPEAK策略对中频频段的选取概率是最低的。从图1a中原始信号的频谱中可以看到,原始信号的能量分布总体上是沿着从低频到高频有下降的趋

势,中频频段的能量并没有明显比高频频段的能量少。但是,由于ACE策略和SPEAK策略在频带划分上都是高频的频带宽,低频的频带窄。因此,高频频段所包括的频带内的信号成分更多,也有助于提高所在通道被选取的概率,而中频频段既没有低频通道的能量集中特性,也没有高频通道的频带宽特性,反而被选取概率最低的。

4 结论

本研究基于信号特征电子耳蜗言语处理策略的频带选取特性,总结了ACE策略和SPEAK策略等算法在频带选取上倾向于低频频带的特征。通过谱分析和电子耳蜗通道选取的统计,对比高低频频带的电子耳蜗不同时间帧的通道选取和参数。基于信号特征电子耳蜗言语处理策略在频带选取上更多选取能量集中的低频信号,部分选取高频信号,较少选择中频信号。该选择模式有助于传递目标语音信号的主要信息,但中高频通道被选取的概率偏低,导致中高频信息丢失,后续的算法研究和改进需要结合该特性研究降低中高频信号失配的补偿模式,尤其是中频信号的提升方法。

【参考文献】

- [1] EHLERS E, GOUELL M J, ZHENG Y, et al. Binaural sensitivity in children who use bilateral cochlear implants[J]. J Acoust Soc Amer, 2017, 141(6): 4264-4277.
- [2] LOPEZ-POVEDA E A, EUSTAQUIO-MARTIN A, STOHL J S, et al. Intelligibility in speech maskers with a binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex[J]. Hear Res, 2017, 348: 134-137.
- [3] SHEFFIELD B M, SCHUCHMAN G, JGW B. Pre- and postoperative binaural unmasking for bimodal cochlear implant listeners[J]. Ear Hear, 2017, 38(5): 554-567.
- [4] GOUELL M J, STAKHOVSKAYA O A, JGW B. Contralateral interference caused by binaurally presented competing speech in adult bilateral cochlear-implant users [J]. Ear Hear, 2018, 39(1): 110-123.
- [5] 罗鑫,傅前杰,王仁华.联合使用助听器 and 增强电子耳蜗的使用者的中文语音识别[J].北京生物医学工程,2005,24(4): 250-253. LUO X, FU Q J, WANG R H. Enhancing Chinese speech recognition for cochlear implant users by using hearing aid in the contralateral ear [J]. Beijing Biomedical Engineering, 2005, 24(4): 250-253.
- [6] BUECHNER A, BRENDEN M, KRUEGER B, et al. Current steering and results from novel speech coding strategies[J]. Otol Neurotol, 2008, 29(2): 203-207.
- [7] MAGNUSSON L. Comparison of the fine structure processing (FSP) strategy and the CIS strategy used in the MED-EL cochlear implant system: speech intelligibility and music sound quality[J]. Int J Audiol, 2011, 50(4): 279-287.
- [8] DUCHATEAU N, MANGADO N, CERESA M, et al. Virtual cochlear electrode insertion via parallel transport frame[C]. New York, 2015: 1398-1401.
- [9] MA X J, SUDANTHI W, ZHOU Y, et al. Simulation for training cochlear implant electrode insertion [C]//IEET International

- Symposium on Computer-based Medical System. IEEE, 2017: 1-6.
- [10] JIANG B, XIA N, WANG X, et al. Auditory responses to short-wavelength infrared neural stimulation of the rat cochlear nucleus[C]. Seogwipo: EMBC, 2017: 1942-1945.
- [11] TEUDT I U, MAIER H, RICHTER C P, et al. Acoustic events and "Optophonic" cochlear responses induced by pulsed near-infrared Laser[J]. IEEE Trans Biomed Eng, 2011, 58(6): 1648-1655.
- [12] SANGHAMITRA V A, REKHA V. Comparison of speech intelligibility parameter in cochlear implants by spatial filtering and coherence function methods[C]. Ghaziabad: ICMETE, 2016: 573-577.
- [13] KALLEL F, GHORBEL M, FRIKHA M. A noise cross PSD estimator based on improved minimum statistics method for two-microphone speech enhancement dedicated to a bilateral cochlear implant[J]. Appl Acoust, 2012, 73(3): 256-264.
- [14] ZENG F G. Challenges in improving cochlear implant performance and accessibility[J]. IEEE Trans Biomed Eng, 2017, 64(8): 1662-1664.
- [15] BOEX C, PELIZZONE M, MONTANDON P. Speech recognition with a CIS strategy for the Ineraid multichannel cochlear implant[J]. Amer J Otol, 1996, 17(1): 61-68.
- [16] CHEN Y S, GONG Q. A normalized beamforming algorithm for broadband speech using a continuous interleaved sampling strategy [J]. IEEE Trans Audio Speech Lan P, 2012, 20(3): 868-874.
- [17] 陈又圣, 王健, 薛国伟, 等. 电子耳蜗 CIS 言语处理策略参数特征研究[J]. 深圳信息职业技术学院学报, 2017, 15(3): 12-18.
- CHEN Y S, WANG J, XUE G W, et al. Research on parameter feature of continuous interleaved sampling strategy for cochlear implant[J]. Journal of Shenzhen Institute of information technology, 2017, 15(3): 12-18.
- [18] 陈又圣, 王健, 陈艳, 等. 基于连续交替采样的电子耳蜗等间隔多刺激算法研究[J]. 中国医学物理学杂志, 2018, 35(5): 385-591.
- CHEN Y S, WANG J, CHEN Y, et al. Continuous interleaved sampling-based multiple equal-interval stimulation algorithm for cochlear implant[J]. Chinese Journal of Medical Physics, 2018, 35(5): 385-591.
- [19] SKINNER M W, HOLDER L K, WHITFORD L A. Speech recognition with the nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults[J]. Ear Hear, 2002, 23(3): 207-223.
- [20] PSARROS C E, PLANT K L, LEE K. Conversion from the SPEAK to the ACE strategy in children using the nucleus 24 cochlear implant system: speech perception and speech production outcomes[J]. Ear Hear, 2002, 23(18): 18-27.

(编辑:陈丽霞)