

电子耳蜗近场信号采集震荡及失配研究

陈又圣,王健,薛国伟,陈培培,张璞
深圳信息职业技术学院,广东 深圳 518000

【摘要】电子耳蜗是一种通过电极刺激耳聋患者的听觉神经来产生听觉信号和听觉感知能力的小型电子装置,包括体内植入电路和体外机两部分。体外机嵌入了微型麦克风用于采集声音信号,传统的电子耳蜗单麦克风及双麦克风信号采集往往假定信号源为远场信号,并以几何中平行波的方式来分析信号采集。实际电子耳蜗更常用的场景是面对面沟通及电话交流,在此应用场景下信号源距离较短。本文研究在电子耳蜗实际应用场景中信号采集的近场震荡特性及不同频率和距离的近场信号采集失配特征,为电子耳蜗处理算法的近场应用提供技术参数及理论分析方法。

【关键词】电子耳蜗;失配;信号采集;近场

【中图分类号】R318;TH776

【文献标志码】A

【文章编号】1005-202X(2017)07-0731-05

Vibration and mismatch of near-field signal acquisition for cochlear implant

CHEN Yousheng, WANG Jian, XUE Guowei, CHEN Peipei, ZHANG Pu
Shenzhen Institute of Information Technology, Shenzhen 518000, China

Abstract: Cochlear implant, including implanted circuit and external part, is a small electric device to stimulate the generation of auditory signal and voice sensation in the auditory nerve of deaf patients through electrode stimulation. A microphone is embedded in the external part of cochlear implant for signal acquisition. In traditional research of cochlear implant with one or two microphones for signal acquisition, the signal source is always assumed to be far-field signal, and the geometric method for parallel wave is applied for signal analysis. In actual common application, more demand of the use of cochlear implant is in the face-to-face communication and telephone communication, in which the distance of signal source is not very larger. The near-field vibration feature in actual applications and mismatch characteristic based on different frequencies and distances in signal acquisition were investigated in the paper, providing technical parameters and theoretical analysis method for the research of speech processing in cochlear implant.

Keywords: cochlear implant; mismatch; signal acquisition; near-field

前言

电子耳蜗最早是从国外的3大厂商(包括澳大利亚的COCHLEAR、美国的AB和奥地利的MED-EL公司)引入国内的,是一种通过植入体内并用电极阵列刺激耳蜗听神经来产生听觉感知的微型电子装置^[1-2]。目前电子耳蜗在安静环境下的言语识别率已经较高,但噪声环境下的识别率仍较低^[3-5],近年来双麦克

风或者多麦克风波束形成及语音增强技术逐渐应用在电子耳蜗前端信号采集上^[6-9]。麦克风阵列采集空间上方位存在差异信号,并通过特定波束形成方法来指向目标方位,在语音增强^[6,10-12]和声源定位^[13-14]方面具有较好的应用。但是,在电子耳蜗领域中所涉及的传统方法是把目标语音源假定为远场,通过远场的假设可以在信号采集和波束形成中把信号近似为平行波的特征,有助于信号分析和公式的推导。但是在电子耳蜗使用者最为常见的面对面和电话交流的应用场景中,实际声源的距离不会很大,此时从声源发出的信号不再是平行的特征,如果仍然按平行波的方法来进行推导及分析则会产生较大的误差。另一方面,语音信号也符合波的一般特征,会产生声波的叠加并产生幅度的震荡和失配的特征。本文通过研究近场圆片波源的信号采集强度特征和震荡失配特征,探索基于频率及距离变化的麦克风

【收稿日期】2017-04-05

【基金项目】深圳科技计划基础研究(JCYJ20160415114250896);广东省自然科学基金博士启动基金(2016A030310073);国家自然科学基金(81401539)

【作者简介】陈又圣,男,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

【通信作者】王健,女,博士,副教授,主要研究方向:听觉音高感知、语音信号处理, E-mail: wangj01@szit.edu.cn

采集特性,为电子耳蜗实际应用场景的语音处理算法提供实际参数分析方法。

1 电子耳蜗近场信号采集

1.1 麦克风远场和近场的信号采集

实际电子耳蜗应用在近场环境中,即麦克风与目标声源的距离较小。在近场环境中,声源波不再假定为平行特征,如图1所示。

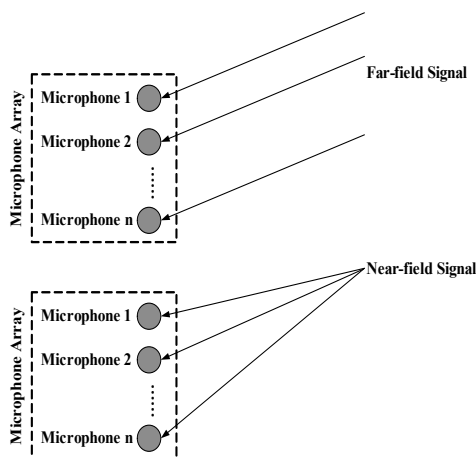


图1 麦克风远场和近场信号采集的比较

Fig.1 Comparison of signal acquisition in far-field and near-field

对于远场声源来说,信号距离较大,如果麦克风间距相对于声源距离来说相对极小的话,则可近似认为平行声波,远场情形的平行波分析有助于简化理论分析^[4-5, 12]。而对于图1中的近场情形,信号可认为是点声源发出,麦克风阵列采集的各声波不具有平行特征。

1.2 近场圆片声源相对强度的计算

对于单个麦克风的信号采集来说,由于声源的距离小,声源本身不是点信号源,麦克风采集了信号源各个位置的信号并产生了叠加^[15]。在电子耳蜗使用者近距离的面对面的交谈中,说话人的发声器官不再是点信号源,把声源近似简化为圆片声源,单个麦克风采集信号如图2所示。

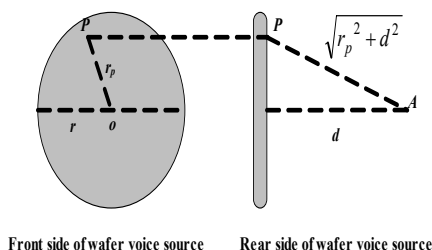


图2 近场圆片声源示意图

Fig.2 Sketch of near-field wafer voice source

图2表示近场圆片声源信号采集的示意图,圆片声源的半径为 r 。圆片声源可以看成是由无数个点声源组成,声源内部或者边沿任意一个点声源 P 距中心点 o 的距离为 r_p ($r_p \leq r$),被放置在 A 点的麦克风采集到(麦克风距离圆片声源中心 o 点的距离为 d)。按照几何关系可以计算出 PA 的距离(如图1所示),距离越大则 A 点采集到的信号的衰减更大。圆片声源上无数个点声源在 A 点叠加形成实际采集的信号,为了排除信号源本身强度的影响,通过惠更斯原理积分计算 A 点的信号强度与圆片信号源本身强度的比值 G ,该比值表示麦克风处于不同距离的信号采集相对强度,其表达式如式(1)所示^[15]:

$$G = \sin^2 \left(\frac{\pi}{\lambda} \left(\sqrt{r^2 + d^2} - d \right) \right) \quad (1)$$

式中, λ 为声音波长,代入 $\lambda = \frac{c}{f}$ 可以得到相对强度与声音频率 f 和距离 d 之间的函数关系,如式(2)所示:

$$G = \sin^2 \left(\frac{\pi f}{c} \left(\sqrt{r^2 + d^2} - d \right) \right) \quad (2)$$

从式(2)表达式可以看到,声源信号相对强度是有关信号频率和的函数。不同的频率相对函数不同,本文分析电子耳蜗实际参数条件的信号强度特征,以常见的8、16、22通道和24通道电子耳蜗为例探讨近场声源强度,其滤波器组中心频率如表1所示。

2 近场圆片声源信号采集特征分析和参数提取

2.1 相对强度函数曲线和震荡区的划分

不同的声源尺寸条件下声源相对强度也不同,以 $r=8$ cm为例说明上述电子耳蜗4种频率参数下的声源相对强度函数曲线,如图3所示。

图3表示所选的电子耳蜗8、16、22和24通道参数条件下的声源相对幅度曲线。在图3a的8通道电子耳蜗中,通道1~5(红色曲线)的声源相对幅度曲线为单调下降,而通道6~8(蓝色曲线)则出现高低变化,即发声幅度的震荡,而且震荡趋于分布在距离接近0的位置,该情形表示震荡发生在靠近声源位置。对比图3各子图,图3b、c和d也有类似的结果。16通道电子耳蜗中的通道11~16(蓝色曲线)发生震荡,22通道电子耳蜗中的通道14~22(蓝色曲线)发生震荡,24通道电子耳蜗中的通道15~24(蓝色曲线)发生震荡。发声源幅度的高低波动的位置称为震荡区,在震荡区中,声源距离的微小变化会导致信号幅度大幅变化,因此信号采集区域应避免在震荡区中。对比图3中不同频率的声源相对幅度曲线可以看到,

表1 8、16、22和24通道电子耳蜗的滤波器组中心频率参数
Tab.1 Center frequency parameters of 8, 16, 22 and 24 channel filter bank in cochlear implant

Channel number	Center frequency parameter in filter bank
8	Channel 1: 274.0 Hz; Channel 2: 517.5 Hz; Channel 3: 761.0 Hz; Channel 4: 1 005.0 Hz; Channel 5: 1 462.0 Hz; Channel 6: 2 376.0 Hz; Channel 7: 3 869.0 Hz; Channel 8: 6 276.0 Hz
16	Channel 1: 216.0 Hz; Channel 2: 343.0 Hz; Channel 3: 485.0 Hz; Channel 4: 645.0 Hz; Channel 5: 826.0 Hz; Channel 6: 1 030.0 Hz; Channel 7: 1 259.0 Hz; Channel 8: 1 516.5 Hz; Channel 9: 1 806.5 Hz; Channel 10: 2 132.5 Hz; Channel 11: 2 499.5 Hz; Channel 12: 2 912.5 Hz; Channel 13: 3 377.0 Hz; Channel 14: 3 899.5 Hz; Channel 15: 4 487.0 Hz; Channel 16: 5 148.0 Hz
22	Channel 1: 213.0 Hz; Channel 2: 335.0 Hz; Channel 3: 456.5 Hz; Channel 4: 578.0 Hz; Channel 5: 700.0 Hz; Channel 6: 822.0 Hz; Channel 7: 944.0 Hz; Channel 8: 1 066.0 Hz; Channel 9: 1 218.5 Hz; Channel 10: 1 401.0 Hz; Channel 11: 1 583.5 Hz; Channel 12: 1 797.0 Hz; Channel 13: 2 071.5 Hz; Channel 14: 2 406.5 Hz; Channel 15: 2 772.0 Hz; Channel 16: 3 168.5 Hz; Channel 17: 3 625.5 Hz; Channel 18: 4 173.5 Hz; Channel 19: 4 813.5 Hz; Channel 20: 5 514.5 Hz; Channel 21: 6 337.0 Hz; Channel 22: 7 281.0 Hz
24	Channel 1: 213.0 Hz; Channel 2: 335.0 Hz; Channel 3: 456.5 Hz; Channel 4: 578.0 Hz; Channel 5: 700.0 Hz; Channel 6: 822.0 Hz; Channel 7: 944.0 Hz; Channel 8: 1 066.0 Hz; Channel 9: 1 188.0 Hz; Channel 10: 1 340.5 Hz; Channel 11: 1 523.0 Hz; Channel 12: 1 705.5 Hz; Channel 13: 1 919.0 Hz; Channel 14: 2 193.5 Hz; Channel 15: 2 498.0 Hz; Channel 16: 2 802.5 Hz; Channel 17: 3 138.0 Hz; Channel 18: 3 534.0 Hz; Channel 19: 3 991.0 Hz; Channel 20: 4 509.0 Hz; Channel 21: 5 087.5 Hz; Channel 22: 5 758.0 Hz; Channel 23: 6 520.0 Hz; Channel 24: 7 342.5 Hz

频率越大,震荡越多(即出现更多的峰和谷),以人耳能听到的声音频率上限 20 kHz 来分析高频的信号幅度震荡情形,如图4所示。

从图4的声源相对幅度在人耳听音频率上限频率 20 kHz 的曲线变化可以看到,在 0.5 m 内产生了5个震荡峰,尤其是在靠近声源的 0~0.11 m 范围内包含了4个震荡峰。在震荡区中采集的信号幅度具有不确定性,信号幅度变化过快,因此电子耳蜗的麦克风应避免放置在震荡区中进行信号采集。

2.2 临界距离的计算和分析

从前面的图3和图4可以看到,高频信号的震荡区域更宽,采集信号应该在非震荡区中。本文所定义的非震荡区为相对幅度函数 G 最后一个峰值所对应的距离 d_0 的值,在该位置之后的函数曲线不再震荡,本文称震荡区和非震荡区的分界线所对应的 d_0 值为震荡区临界距离。

声带是人发声的主要器官及声的信号源,位于喉腔之间。人喉直径 2~3 cm,因此以声带作为圆片声源的 r 为 1.0~1.5 cm。对于人耳听音范围 20 Hz~20 kHz 频率范围来说,该尺寸条件下的圆片声源的震荡区临界距离如图5所示。

对于 1 cm 半径的圆片声源来说(蓝线),当频率小于 17.02 kHz 时震荡区临界距离为 0,也就是 $r=1$ cm 半径的圆片声源当频率小于 17.02 kHz 时不存在震荡区。而当频率大于 17.02 kHz 时,临界距离逐渐增大,在 20 kHz 的人耳听觉上限频率下临界距离最大为 0.001 62 m。同样,对于 1.5 cm 半径的圆片声源来说(蓝线),在频率 0~11.34 kHz 范围内不存在震荡区,频率在 11.34~20.00 kHz 范围内最大临界距离为 0.008 98 m。表1中各通道电子耳蜗最高频率为 7 342.5 Hz,因此在电子耳蜗滤波器组频带范围 0~7 342.5 Hz 内不存在震荡区。

考虑到声带发出的声音还需要经过复杂机制并在喉腔、咽部、口腔、唇部和鼻腔等部位综合形成复杂声音,因此由声带源出来的信号源扩大,可以把这些发音器官综合起来形成一个大的圆片声源来进行分析,其宽度为 5~8 cm,即使半径 r 为 2.5~4.0 cm,该尺寸条件下的圆片声源的震荡区临界距离如图6所示。

图6的以发音器官为圆片声源($r=2.5\sim4.0$ cm)的震荡区临界距离曲线特征与图5的声带圆片声源类似,分析并计算图中的参数,结果如表2所示。

从表2可以看到,对于半径为 2.5~4.0 cm 的圆片声源来说,综合各个参数,当频率大于 4 251 Hz 时就会出现震荡区,最大临界距离(对应 20 kHz 的人耳听觉上限频率)约为 9 cm(0.089 86 m),而对于电子耳蜗通道最大频率 7 342.5 Hz 来说,最大临界距离约 2.3 cm(0.022 96 m)。对于电子耳蜗使用者日常使用来说,面对面交流的距离一般在 50 cm 左右,而手机打电话时手机贴近耳朵距离较小,一般距离小于 9 cm,而距离是否小于 2.3 cm 则要看使用时手机发音源是否直接贴近电子耳蜗的麦克风。由于在声源震荡区进行信号采集会导致信号幅度突变和确定性,因此实际使用中应放置非震荡区外。对于电子耳

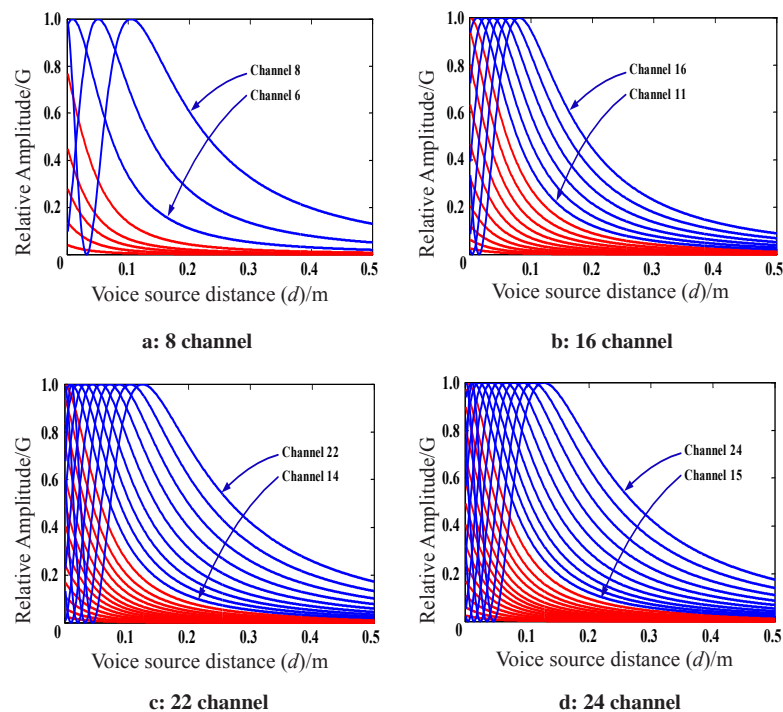
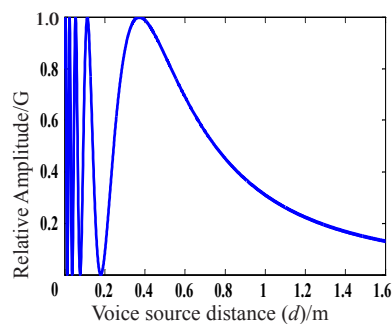
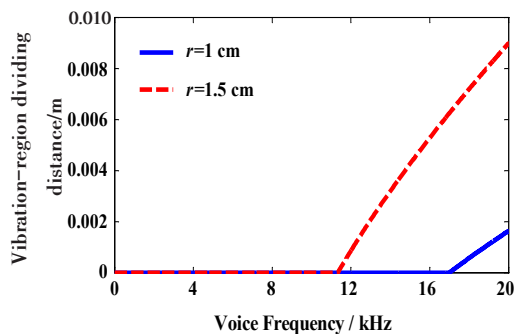
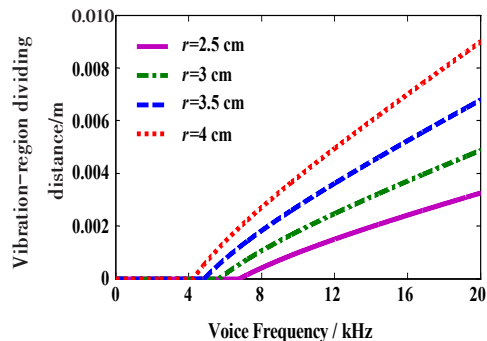
图3 在 $r=8$ cm 圆片声源时不同距离的相对信号幅度比较Fig.3 Comparison of relative amplitude based on wafer voice source with $r=8$ cm

图4 人耳听音频率上限频率 20 kHz 的相对幅度曲线

Fig.4 Relative amplitude curve at the human auditory upper limiting frequency of 20 kHz

图5 以声带为圆片声源($r=1.0\sim1.5$ cm)的震荡区临界距离曲线Fig.5 Vibration-region dividing distance curve of vocal cords based on wafer voice source ($r=1.0\sim1.5$ cm)

蜗使用者面对面交流情形来说,信号采集不受影响,而对于电话交流情形来说,部分高频信号的信号采集

图6 以发音器官为圆片声源($r=2.5\sim4.0$ cm)的震荡区临界距离曲线Fig.6 Vibration-region dividing distance curve of vocal organs based on wafer voice source ($r=2.5\sim4.0$ cm)

会受到影响。

3 结论

多麦克风信号采集和语音增强技术逐渐应用在电子耳蜗前端信号采集中^[2-5,9,12]。本文研究了麦克风近场信号采集的震荡及失配特征,通过理论上分析和工程上分析电子耳蜗麦克风在近场圆片声源信号采集信号强度变化特征,并进一步分析了震荡区和非震荡区的临界距离计算。最后通过结合电子耳蜗实际滤波器频率参数分析电子耳蜗使用者在不同情形的信号采集影响情形,具有较大的应用价值。

表2 以发音器官为圆片声源($r=2.5\sim4.0$ cm)的震荡区范围和临界距离参数Tab.2 Vibration region range and dividing distance for vocal organs based wafer voice source ($r=2.5\sim4.0$ cm)

Wafer source radius/cm	Vibration region range/kHz	Maximal dividing distance/m	Dividing distance for 7 342.5 Hz/m
2.5	0-6.800	0.032 50	0.001 91
3.0	0-5.670	0.048 68	0.007 85
3.5	0-8.859	0.067 80	0.014 87
4.0	0-4.251	0.089 86	0.022 96

【参考文献】

- [1] 陈又圣, 王健, 陈伟芳. 电子耳蜗n of m算法的特性[J]. 中国医学物理学杂志, 2017, 34(3): 266-271.
CHEN Y S, WANG J, CHEN W F. Feature of n of m algorithm for cochlear implant[J]. Chinese Journal of Medical Physics, 2017, 34(3): 266-271.
- [2] 陈又圣. 基于Matlab的电子耳蜗信号采集研究[J]. 深圳信息职业技术学院学报, 2016, 14(3): 6-10.
CHEN Y S. Research on Matlab based signal acquisition for cochlear implant[J]. Journal of Shenzhen Institute of Information Technology, 2016, 14(3): 6-10.
- [3] CHUNG K, ZENG F G. Using hearing aid adaptive directional microphones to enhance cochlear implant performance[J]. Hear Res, 2009, 250(2): 27-37.
- [4] CHEN Y S, GONG Q. Broadband beamforming compensation algorithm in CI front-end acquisition[J]. Biomed Eng Online, 2013, 12(1): 18.
- [5] CHEN Y S, GONG Q. Real-time spectrum estimation-based dual-channel speech-enhancement algorithm for cochlear implant[J]. Biomed Eng Online, 2012, 11(1): 74.
- [6] MIRZAHASANLOO T, KEHTARNAVAZ N. Real-time dual-microphone noise classification for environment-adaptive pipelines of cochlear implants[C]//2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2013: 5287-5290.
- [7] NIMA Y, PHILIPOS C L. A dual-microphone speech enhancement algorithm based on the coherence function[J]. IEEE Trans Audio Speech Lang Processing, 2012, 20(2): 599-609.
- [8] HER A A, GRAYDEN D B, FALLON J B, et al. Perceptual effect of reverberation on multi-microphone noise reduction for cochlear implant[C]//IEEE International Conference on Acoustics. IEEE, 2015: 5853-5857.
- [9] 陈又圣, 宫琴. 基于双TP麦克风的电子耳蜗前端指向性语音增强系统的研制[J]. 仪器仪表学报, 2010, 31(9): 1952-1958.
CHEN Y S, GONG Q. Study on dual TP microphone-based directional speech enhancement system for cochlear implant[J]. Chinese Journal of Scientific Instrument, 2010, 31(9): 1952-1958.
- [10] ELKO G W, PONG A T. A simple adaptive first-order differential microphone [C]//IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. IEEE, 1995: 169-172.
- [11] FISCHER S, SIMMER K U. Beamforming microphone arrays for speech acquisition in noisy environments[J]. Speech Commun, 1996, 20(3): 215-227.
- [12] CHEN Y S, GONG Q. A normalized beamforming algorithm for broadband speech using a continuous interleaved sampling strategy [J]. IEEE Trans Audio Speech Lang Processing, 2012, 20(3): 868-874.
- [13] IVAN J T, LONG L, VANI G, et al. Cost function for sound source localization with arbitrary microphone arrays[C]//2017 Hands-free Speech Communications and Microphone Arrays. IEEE, 2017: 76-80.
- [14] MADMONI L, BARFUSS H, RAFAELY B, et al. A unified framework for multiple arrays on a robot and application to sound localization [C]//2017 Hands-free Speech Communications and Microphone Arrays. IEEE, 2017: 66-70.
- [15] 高杰. 数字助听器中语音增强算法的研究[D]. 北京: 清华大学, 2005.
GAO J. Research on speech enhancement algorithm for digital hearing aids [D]. Beijing: Tsinghua University, 2005.

(编辑: 薛泽玲)