

语音识别在双相障碍中应用的研究进展

杨秋实 丰雷 王刚

100088 首都医科大学附属北京安定医院 国家精神疾病医学中心 国家精神心理疾病临床医学研究中心 精神疾病诊断与治疗北京市重点实验室(杨秋实、丰雷、王刚);

100069 首都医科大学人脑保护高精尖创新中心(丰雷、王刚)

通信作者: 王刚, Email: gangwangdoc@ccmu.edu.cn

DOI: 10.3969/j.issn.1009-6574.2023.10.007

【摘要】 双相障碍是一种常见精神疾病。目前双相障碍的诊断主要依靠临床经验,其早期识别率、诊断率较低。语言信号特征在双相障碍与健康人群中存在差异,且随着患者情绪和疾病状态变化而改变,具有成为双相障碍客观评估指标的潜力。现对语音识别技术的基本概念及方法,以及语音识别在双相障碍中的研究进展进行介绍,旨在从中发现反映双相障碍的语音特征类别,为建立双相障碍客观评估体系提供依据。

【关键词】 双相障碍; 语音识别; 诊断; 综述

基金项目: 国家自然科学基金(61602017);北京市属医院科研培育项目(PX2022074);北京市高层次公共卫生技术人才培养计划(学科骨干-01-12)

Research progress on speech recognition in bipolar disorder Yang Qiushi, Feng Lei, Wang Gang
Beijing Key Laboratory of Mental Disorders, National Clinical Research Center for Mental Disorders & National Center for Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing 100088, China (Yang QS, Feng L, Wang G); Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing 100069, China (Feng L, Wang G)
Corresponding author: Wang Gang, Email: gangwangdoc@ccmu.edu.cn

【Abstract】 Bipolar disorder is a common psychiatric disorder. At present, the diagnosis of bipolar disorder mainly relies on clinical experience, and its early recognition rate and diagnostic rate are relatively low. There are differences in language signal characteristics between individuals with bipolar disorder and healthy individuals, which change with patients' emotions and disease state, and have the potential to become objective evaluation indicators for bipolar disorder. This paper introduces the basic concepts and methods of speech recognition technology, as well as the research progress of speech recognition in bipolar disorder, with the aim of identifying speech feature categories that reflect bipolar disorder and providing a basis for establishing an objective evaluation system for bipolar disorder.

【Key words】 Bipolar disorder; Speech recognition; Diagnosis; Review

Fund programs: National Natural Science Foundation of China(61602017); Beijing Municipal Hospital Research and Cultivation Project(PX2022074); Beijing High Level Public Health Technology Talent Training Plan(xuekegugan-01-12)

双相障碍是一种既有躁狂或轻躁狂发作,又有抑郁发作的常见精神疾病^[1]。流行病学调查显示,我国双相障碍的年患病率为0.5%,终生患病率为0.6%^[2]。双相障碍具有发病年龄早、自杀率高等特点^[3-4],是一种严重的社会及公共卫生问题。双相障碍的诊断主要依赖医生的临床经验,了解患者病情的纵向变化。双相障碍早期特征不明显、临床表现复杂,导致该疾病的早期识别率、诊断率较低^[5],

因此寻求可靠的标志物作为双相障碍早期识别、诊断与鉴别诊断、评价预后与转归的工具尤为重要。

语言可被用来描述情感性疾病的特征,双相障碍患者的语言特点会随着情绪变化而改变^[5]。相比于常见的生物标志物,语音信号具有数量大、易获取、非侵入性等特点,语音信号特征可能对双相障碍的早期识别、鉴别诊断、严重程度评估、预后监测有应用价值,成为精神疾病临床评估的客观指标之一。

现介绍语音识别(speech recognition)技术的基本概念及方法、步骤,并对语音识别在双相障碍领域的研究进展进行综述。

一、语音识别技术的基本概念及方法

1. 语音识别的基本概念: 语音识别是让机器听懂人类语言的技术。其目的是开发出一种具备听觉能力的机器,能够收集人类语音,并通过算法和模型对语音进行识别、分析和解读,转化为文本或计算机指令^[6]。

2. 语音识别技术方法: 语音识别大致可以分为3个步骤。第1步是语音信号的数字化及预处理。语音信号经过采样、量化、编码过程转化为数字信号后进行预处理,将自然语音转化为可分析的数字信号^[7]。其中如何高效地采集语音信号对后续研究至关重要。目前,应用于精神疾病语音识别的语音采集方法可分为两类:一类是结构性语音信息。如阅读中性文字、看图说话^[8],或是半定式访谈,患者回答问题或量表评估时录制患者的反馈^[9],此类方法操作简单、内容清晰,便于医院场景使用,但该方法形式单一、内容固定,可能无法反映患者的真实情绪变化。另一类是自然状态下的语音采集,包括手机后台录音和主动自我监控^[10-13]。相关研究表明,自然语音数据提取的特征对躁狂及混合发作的分类更加准确,可以更好地反映患者的情绪变化^[14],且自然语音具有易于采集、非侵入性的特点^[15],适用于长时、动态监测。但采集自然语音存在录音设备不同、环境噪音不可控、无效信息量大等问题。

第2步是语音特征参数的提取。语音特征可分为韵律特征、声门波特征、音质特征和频谱特征。常用的韵律特征包括基频(pitch)、能量(energy)、时长(duration)等。韵律特征被认为是最主要的用于描述情感的语音特征^[16],其多用于描述情感维度的差别,且与情感的效价有关。如能量在效价高的情绪中偏高,在效价低的情绪中偏低^[17]。声门波特征主要包括共振峰和元音空间,共振峰(formant)代表了声道变化的各种信息,常被用于认知负荷分类和抑郁的识别与评估^[18],第一、第二、第三共振峰(F1、F2、F3)的峰值随情绪波动的变化较为明显^[19]。音质特征以呼吸音、喉化音、明亮度等表现形式存在^[18],常见的参数是频率微扰(jitter)和振幅微扰(shimmer),其中频率微扰可能反映了影响肌肉张力和发音控制的自主神经系统失调,可用于表征情绪状态^[20]。频谱特征体现了声道形状变化与发声运动间的相关性^[21],包括Mel频率倒谱系数(Mel

frequency cepstrum coefficient, MFCC)、线性预测倒谱系数(linear prediction cepstrum coefficient, LPCC)、线性预测系数(linear prediction coefficient, LPC)等。其中LPC、LPCC是以模拟发声的声管模型为基础提取的参数,反映了声道响应的特性,MFCC是利用人耳听觉感知特性和语音产生机制得到的参数,通过对不同频段声波的敏感度差异,证明人耳对语音信号的感知与语音频率变化的关系^[18]。

第3步是语音识别模型的建立与训练。根据不同的需求,选择合适的统计方法进行建模。目前常用的语音识别建模算法主要分为两类:一类是模式匹配方法,常见动态时间规整(dynamic time warping)、隐马尔科夫模型(hidden Markov models);另一类为基于深度学习的模型,如循环神经网络(recurrent neural networks)、卷积神经网络(convolutional neural networks)。

二、双相障碍的语音识别研究现状

目前基于语音信号特征的研究主要集中在抑郁症、精神分裂症,相比于前两者,双相障碍语音特征的研究起步较晚、数量有限。目前语音识别技术在双相障碍中的研究,主要集中在疾病状态识别、辅助疾病诊断、严重程度评估及语音识别模型的构建。

1. 疾病状态识别: Vanello等^[8]记录6例双相障碍患者(3例抑郁发作、3例轻躁狂发作)处于发病期(抑郁或轻躁狂)及缓解期中的音频信号,发现基频均值、标准差在发病期与缓解期中的差异有统计学意义,提示基频特征的变化可以区分不同的疾病状态。Karam等^[12]对6例双相I型障碍患者进行了6个月至1年的追踪研究,通过手机软件收集221 h不同情绪状态下的自然语音数据,共提取51个维度的特征进行统计分析,结果显示处于轻躁狂或抑郁状态的患者与缓解状态患者,在基频标准差、过零率的段均值等基频及时长特征中的差异有统计学意义。该研究提示通过手机软件采集的自然语音数据对情绪状态的识别有一定意义。

一些研究者认为自然语音数据更能反映患者的情绪变化,如Farrús等^[22]对手机录音进一步研究,发现韵律特征(如基频特征)对躁狂状态的检测更加准确,而抑郁状态的检测更依赖于非韵律特征。但语音特征与患者当前的疾病状态并无对应。一项结构性访谈研究表明,第四共振峰(F4)及线性预测系数在躁狂与缓解期的差异有统计学意义,基频特征的差异无统计学意义^[23]。这种结果的差异可能与两项研究采用的语音采集方法不同有关,但在双相

抑郁中是否也存在差异,有待进一步研究验证。

2. 辅助疾病诊断: Guidi 等^[24]采集 11 例双相障碍患者及 18 名健康对照者阅读中性文本的语音信息进行比较,结果显示轻躁狂组的长停顿时长和全停顿时长的中位数较健康组降低,而以浊音/清音比组合的浊音清音片段时长提高;抑郁组与健康人群相比全停顿时长中位数及两个连续语段时间间距中位数升高,而浊音片段率降低,提示时长特征可区分双相轻躁狂患者与健康人群。Zhang 等^[23]对 30 例双相躁狂患者及 30 名健康对照的语音信息进行对比,发现 F1、F2 和线性预测系数是区分双相躁狂患者和健康人群的关键因素。

Yamamoto 等^[25]采集了 97 例抑郁症、68 例双相抑郁和 76 名健康对照的半定式访谈内容,结果发现与健康个体相比,抑郁患者的语速较慢,停顿时间及反应时间较长;而双相抑郁患者与健康个体间没有明显的特征差异。而 2021 年 Faurholt-Jepsen 等^[26]对 121 例双相障碍患者和 38 名健康对照进行研究,发现双相障碍患者与健康人群的语音特征的差异有统计学意义,但该研究在比较过程中未对双相障碍的不同状态进行分类,因此对双相抑郁患者与健康人群的识别意义较局限。

Yamamoto 等^[25]在上述研究中对抑郁症与双相抑郁患者的语音特征也进行了分析,结果发现,抑郁症患者较双相抑郁患者反应速度慢、语速和停顿时长未见差异,该研究认为利用语音特征区分抑郁症与双相抑郁的能力相对有限。此后一项基于自然手机通话的研究认为抑郁症与双相障碍抑郁发作患者的语音特征没有明显差异^[27]。而 Higuchi 等^[28]的研究对 8 例双相抑郁、14 例抑郁症患者和 32 名健康对照者朗读固定文字时的声学特征进行分析,选择了韵律和频谱两个维度的语音特征构建分类模型,其整体分类准确率为 90.79%。该结果说明语音特征可用于区分抑郁症与双相抑郁患者,然而研究的样本量较小,且未提供详细的临床评估方法,降低了其研究结果的临床参考价值。

3. 疾病严重程度评估: Zhang 等^[23]在研究中发现线性预测系数与患者的 Bech-Rafaelsdn 躁狂评分量表(Bech-Rafaelsen Mania Rating Scale, BRMS)得分之间呈正相关,提示线性预测系数的增加可能对双相躁狂严重程度的评估有重要作用。又有研究指出,抑郁与双相障碍抑郁发作患者的语音速率、停顿时间和反应时间与个体 HAMD-17 评分之间呈偏相关,其中语音速率的变化与 HAMD-17 的评分变化

呈偏相关,由此认为时间相关特征可以作为监测抑郁严重程度的标志^[25]。

4. 语音识别模型的构建: Karam 等^[12]在对手机自然语音识别的研究中采用支持向量机(support vector machines, SVM)建模,其平均识别率在躁狂发作患者中为 61%,抑郁发作患者中只有 59%。Weiner 等^[29]将机器学习应用于语言流畅性任务,该任务区分抑郁和混合抑郁的准确率为 84%,区分轻躁狂和混合轻躁狂的准确率为 86%。Weintraub 等^[30]也利用机器学习实现了一种识别率至少为 75.2% 的算法,用于检测情绪的表达。

目前研究者可选的语音识别算法众多,但缺乏统一标准,因此,有研究者对目前常见的语音识别算法在双相障碍状态识别方面的优劣进行比较。Pan 等^[31]选取 21 例双相障碍住院患者,提取了基频、共振峰、MFCC、LPCC、GFCC 等参数,比较 SVM 和高斯混合模型(Gaussian mixture models, GMM)在单例和多例双相障碍躁狂状态检测中的准确性,结果显示单一特征参数中 LPCC 的识别效率最高,在 SVM 及 GMM 中分别达到 87.66% 和 80.70%, SVM 对小样本特定人群的检测效果较好,而在非特定人群检测中 GMM 准确性更高。需要注意的,性别可能影响双相障碍的识别率,在一项基于声音和表情特征的研究中,语音特征(如能量和频谱谐波)为女性双相障碍患者的识别提供了更高的准确率^[32]。

三、总结和展望

目前,对双相障碍的语音识别的研究已取得一定的进展,如语音特征在双相障碍的不同状态及与健康人群之间的差别;将常用量表与语音特征结合,评估疾病的严重程度;利用语音识别模型对疾病状态识别的准确率也有一定提升。但仍存在一定的问题:(1)样本量较少,多数研究样本量局限于几人到几十人的区间,研究结果缺乏代表性,无法在人群中进行推广;(2)自然语音信息量大,采集设备及环境噪声对预处理及特征提取的影响较大;结构性语音信息模式较单一,可能无法真实地反映患者的情绪变化;(3)语音识别模型各有优劣,缺乏统一标准,此前有研究者将不同建模算法结合建立混合识别模型,将不同种类算法的优势进行综合,但建模过程复杂、训练运算量大,很难获得全局优化的模型^[33-34];(4)对汉语语境下双相障碍患者语音识别的研究不足。

今后可以关注以下研究方向:(1)扩充样本量,提高研究结果在人群中的普遍性;(2)改进语音特征提取方法及建模方法,提高对不同来源语音的适应

性和识别准确率; (3) 加强对汉语语音的研究, 与英语等非声调语言不同, 汉语是一种声调语言, 不同声调所蕴含的语音特征对汉语语音识别的意义重大, 因此对汉语双相障碍人群及不同语言之间的对比研究对提高语音识别模型的识别率有一定价值。

利益冲突 文章所有作者共同认可文章无相关利益冲突

作者贡献声明 文献收集与整理、论文撰写为杨秋实, 文章构思为王刚, 论文修订为丰雷、王刚

参 考 文 献

- [1] 美国精神医学学会, 张道龙. 精神障碍诊断与统计手册[M]. 北京: 北京大学出版社, 2016.
- [2] Huang Y, Wang Y, Wang H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study[J]. *Lancet Psychiatry*, 2019, 6(3): 211-224. DOI: 10.1016/S2215-0366(18)30511-X.
- [3] Nowrouzi B, McIntyre RS, MacQueen G, et al. Admixture analysis of age at onset in first episode bipolar disorder[J]. *J Affect Disord*, 2016, 201: 88-94. DOI: 10.1016/j.jad.2016.04.006.
- [4] Dong M, Lu L, Zhang L, et al. Prevalence of suicide attempts in bipolar disorder: a systematic review and meta-analysis of observational studies[J]. *Epidemiol Psychiatr Sci*, 2020, 29: e63. DOI: 10.1017/S2045796019000593.
- [5] 陆林. 沈渔邨精神病学[M]. 6 版. 北京: 人民卫生出版社, 2018.
- [6] 韩纪庆, 张磊, 郑铁然. 语音信号处理[M]. 3 版. 北京: 清华大学出版社, 2019.
- [7] 胡航. 现代语音信号处理[M]. 北京: 电子工业出版社, 2014.
- [8] Vanello N, Guidi A, Gentili C, et al. Speech analysis for mood state characterization in bipolar patients[J]. *Annu Int Conf IEEE Eng Med Biol Soc*, 2012, 2012: 2104-2107. DOI: 10.1109/embc.2012.6346375.
- [9] Morrison D, Wang R, De Silva LC. Ensemble methods for spoken emotion recognition in call-centres[J]. *Speech Communication*, 2007, 49(2): 98-112. DOI: 10.1016/j.specom.2006.11.004.
- [10] Grünerbl A, Muaremi A, Osmani V, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients[J]. *IEEE J Biomed Health Inform*, 2015, 19(1): 140-148. DOI: 10.1109/jbhi.2014.2343154.
- [11] Guidi A, Salvi S, Ottaviano M, et al. Smartphone application for the analysis of prosodic features in running speech with a focus on bipolar disorders: system performance evaluation and case study[J]. *Sensors (Basel)*, 2015, 15(11): 28070-28087. DOI: 10.3390/s151128070.
- [12] Karam ZN, Provost EM, Singh S, et al. Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech[J]. *Proc IEEE Int Conf Acoust Speech Signal Process*, 2014, 2014: 4858-4862. DOI: 10.1109/icassp.2014.6854525.
- [13] Majid S, Reeves S, Figueredo G, et al. The extent of user involvement in the design of self-tracking technology for bipolar disorder: literature review[J]. *JMIR Ment Health*, 2021, 8(12): e27991. DOI: 10.2196/27991.
- [14] Faurholt-Jepsen M, Busk J, Frost M, et al. Voice analysis as an objective state marker in bipolar disorder[J]. *Transl Psychiatry*, 2016, 6(7): e856. DOI: 10.1038/tp.2016.123.
- [15] Hecker P, Steckhan N, Eyben F, et al. Voice analysis for neurological disorder recognition—a systematic review and perspective on emerging trends[J]. *Front Digit Health*, 2022, 4: 842301. DOI: 10.3389/fgdh.2022.842301.
- [16] Bhatti MW, Wang Y, Guan L. A neural network approach for human emotion recognition in speech[C]// *Proceedings of the 2004 IEEE International Symposium on Circuits and Systems (ISCAS)*. New York: IEEE, 2004. DOI: 10.1109/ISCAS.2004.1329238.
- [17] 王薇, 杨丽萍, 魏丽, 等. 语音情感特征的提取与分析[J]. *实验室研究与探索*, 2013, 32(7): 91-94. DOI: 10.3969/j.issn.1006-7167.2013.07.023.
- [18] Wang W, Yang LP, Wei L, et al. Extraction and analysis of speech emotion characteristics[J]. *Research and Exploration in Laboratory*, 2013, 32(7): 91-94.
- [18] 郑纯军, 王春立, 贾宁. 语音任务下声学特征提取综述[J]. *计算机科学*, 2020, 47(5): 110-119. DOI: 10.11896/j.sjcx.190400122.
- [19] Zheng C, Wang C, Jia N. Survey of acoustic feature extraction in speech tasks[J]. *Computer Science*, 2020, 47(5): 110-119.
- [19] Titze IR, Baken RJ, Bozeman KW, et al. Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization[J]. *J Acoust Soc Am*, 2015, 137(5): 3005-3007. DOI: 10.1121/1.4919349.
- [20] Ozdas A, Shiavi RG, Silverman SE, et al. Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk[J]. *IEEE Trans Biomed Eng*, 2004, 51(9): 1530-1540. DOI: 10.1109/tbme.2004.827544.
- [21] Benesty J, Sondhi MM, Huang Y. *Springer handbook of speech processing*[M]. Berlin: Springer, 2008.
- [22] Farrús M, Codina-Filbà J, Escudero J. Acoustic and prosodic information for home monitoring of bipolar disorder[J]. *Health Informatics J*, 2021, 27(1): 1460458220972755. DOI: 10.1177/1460458220972755.
- [23] Zhang J, Pan Z, Gui C, et al. Analysis on speech signal features of manic patients[J]. *J Psychiatr Res*, 2018, 98: 59-63. DOI: 10.1016/j.jpsychires.2017.12.012.
- [24] Guidi A, Schoentgen J, Bertschy G, et al. Features of vocal frequency contour and speech rhythm in bipolar disorder[J]. *Biomedical Signal Processing and Control*, 2017, 37: 23-31. DOI: 10.1016/j.bspc.2017.01.017.
- [25] Yamamoto M, Takamiya A, Sawada K, et al. Using speech recognition technology to investigate the association between timing-related speech features and depression severity[J]. *PLoS One*, 2020, 15(9): e0238726. DOI: 10.1371/journal.pone.0238726.
- [26] Faurholt-Jepsen M, Rohani DA, Busk J, et al. Voice analyses using smartphone-based data in patients with bipolar disorder, unaffected relatives and healthy control individuals, and during different affective states[J]. *Int J Bipolar Disord*, 2021, 9(1): 38. DOI: 10.1186/s40345-021-00243-3.
- [27] Faurholt-Jepsen M, Rohani DA, Busk J, et al. Discriminating between patients with unipolar disorder, bipolar disorder, and healthy control individuals based on voice features collected from naturalistic smartphone calls[J]. *Acta Psychiatr Scand*, 2022, 145(3): 255-267. DOI: 10.1111/acps.13391.

青少年非自杀性自伤行为与反刍思维相关性的研究进展

金茜 尹斐 朱婉 李梦甜 葛菲 曹建琴

150081 哈尔滨医科大学护理学院

通信作者: 曹建琴, Email: cj338@163.com

DOI: 10.3969/j.issn.1009-6574.2023.10.008

【摘要】 非自杀性自伤(NSSI)是自杀意念和自杀行为的重要预测因素,其在全球青少年中的患病率呈逐渐上升趋势,成为青少年常见的公共卫生问题之一。反刍思维作为一种认知过程和情绪调节策略,与NSSI密切相关,高水平的反刍思维易诱发NSSI。本文从青少年NSSI和反刍思维的相关理论、反刍思维对青少年产生的影响及相关干预措施进行综述,旨在为早期发现青少年NSSI并进行针对性的干预提供方向。

【关键词】 青少年; 非自杀性自伤; 反刍思维; 综述

基金项目: 国家自然科学基金(72204065); 黑龙江省自然科学基金(LH2020H030); 黑龙江省哲学社会科学规划项目(19EDB091)

Research progress in rumination in adolescents with non-suicidal self-injury behaviors Jin Xi, Yin Fei, Zhu Wan, Li Mengtian, Ge Fei, Cao Jianqin
School of Nursing, Harbin Medical University, Harbin 150081, China
Corresponding author: Cao Jianqin, Email: cj338@163.com

【Abstract】 Non-suicidal self-injury (NSSI) is an important predictor of suicidal thoughts and behaviors. The prevalence of NSSI in the global youth population is gradually increasing, and it is one of the common public health problems among adolescents. As a cognitive process and emotional regulation strategy, rumination is closely related to NSSI. High-level rumination can easily induce NSSI. This paper summarizes the related theories of adolescent NSSI and rumination, the influence of rumination on adolescents and intervention

- [28] Higuchi M, Tokuno S, Nakamura M, et al. Classification of bipolar disorder, major depressive disorder, and healthy state using voice[J]. Asian Journal of Pharmaceutical and Clinical Research, 2018, 11(3): 89-93. DOI: 10.22159/ajpcr.2018.v11i3.30042.
- [29] Weiner L, Guidi A, Doignon-Camus N, et al. Vocal features obtained through automated methods in verbal fluency tasks can aid the identification of mixed episodes in bipolar disorder[J]. Transl Psychiatry, 2021, 11(1): 415. DOI: 10.1038/s41398-021-01535-z.
- [30] Weintraub MJ, Posta F, Arevian AC, et al. Using machine learning analyses of speech to classify levels of expressed emotion in parents of youth with mood disorders[J]. J Psychiatr Res, 2021, 136: 39-46. DOI: 10.1016/j.jpsychires.2021.01.019.
- [31] Pan Z, Gui C, Zhang J, et al. Detecting manic state of bipolar disorder based on support vector machine and gaussian mixture model using spontaneous speech[J]. Psychiatry Investig, 2018, 15(7): 695-700. DOI: 10.30773/pi.2017.12.15.
- [32] Birnbaum ML, Abrami A, Heisig S, et al. Acoustic and facial features from clinical interviews for machine learning-based psychiatric diagnosis: algorithm development[J]. JMIR Ment Health, 2022, 9(1): e24699. DOI: 10.2196/24699.
- [33] Mohammadi HS, Saeidi R. Speaker identification performance enhancement using Gaussian mixture model with GMM classification post-processor[C]// Proceedings of the 2007 IEEE International Conference on Signal Processing and Communications. New York: IEEE, 2007. DOI: 10.1109/ICSPC.2007.4728366.
- [34] Fu L, Mao X, Chen L. Speaker independent emotion recognition based on SVM/HMMs fusion system[C]// Proceedings of the 2008 International Conference on Audio, Language and Image Processing. New York: IEEE, 2008. DOI: 10.1109/ICALIP.2008.4590144.

(收稿日期: 2023-03-27)

(本文编辑: 赵金鑫)