

耳语音情感数据库的设计与建立

金 贇^{1,2}, 赵 艳¹, 黄程韦¹, 赵 力¹

(1. 东南大学信息科学与工程学院, 南京 210096; 2. 徐州师范大学物电学院, 江苏徐州 221116)

摘要: 为了研究耳语音情感识别, 制定了耳语音情感数据库的制作流程和制作规范, 设计并建立了一个包含 4000 个单词、5000 条语句和 600 个段落的耳语音情感数据库。利用模糊数学中的隶属度函数和层次分析法对语句的有效性进行了检验, 得到一个满足研究要求的耳语音情感数据库。利用库中的语料进行初步的耳语音情感识别, 得知正常语音情感识别所用的音长、能量等参数, 在耳语音中仍然可以使用; 而正常音的其它涉及基音的一些参数, 不适用于耳语音, 需要进行改进或者提取新的特征参数。

关键词: 耳语音; 情感识别; 数据库; 隶属度函数; 层次分析法

中图分类号: TN912

文献标识码: A

文章编号: 1000-3630(2010)-01-0063-06

DOI 编码: 10.3969/j.issn1000-3630.2010.01.015

The design and establishment of a Chinese whispered speech emotion database

JIN Yun^{1,2}, ZHAO Yan¹, HUANG Cheng-wei¹, ZHAO Li¹

(1. School of Information Science and Engineering, Southeast University, Nanjing 210096, China;

2. School of Physics and Electronics Engineering, Xuzhou Normal University, Xuzhou 221116, Jiangsu, China)

Abstract: In order to study the new issue of emotion recognition of whispered speech, the construction flow and standards are developed, and the whispered speech emotion database, which includes 4000 words, 5000 sentences and 600 paragraphs, is designed and established. With the concepts of Membership Function and Analytic Hierarchy Process in fuzzy math, the validity of each sentence has been checked out to obtain a qualified database with which a preliminary experiment of whispered speech emotion recognition is carried out. It is concluded that the parameters such as sound length and sound power, which are suitable for normal speech emotion recognition, are still useful in whispered speech, while the other parameters related to pitch are not suitable for whispered speech emotion recognition. So the old parameters should be improved or new parameters would be extracted.

Key words: whispered speech; emotion recognition; database; membership function; AHP

1 引言

耳语音是人们常见的语言交流方式之一, 在会场、音乐厅、图书馆等禁止大声喧哗的场所被广泛应用。早期由于技术的限制, 关于耳语音的研究主要停留在基础语音学研究^[1-3]和医学工作领域^[4]。随着科学技术的发展, 近年来对耳语音的研究逐渐走向多领域和实际应用, 例如, 耳语音的声学性质^[5-9]、耳语音转换为正常音^[10-12]、耳语音的声调语调感知^[13,14]、耳语音的语音增强^[15,16]、耳语音的语音识别^[17]等等。

与正常语音一样, 人们在用耳语音交流时, 也可从耳语音中感受到说话人所带有的情感信息。正常语音的情感计算早在上世纪 90 年代就已经展开, 并且随着情感研究的深入, 正常语音的情感特征分析与识别得到了很大的发展^[18], 而且为语音识别和语音合成^[19]等研究提供了新的思路和方法。

耳语音作为一种常见而特殊的语音信号, 对其进行情感信息处理技术的研究是非常有意义的。首先, 耳语音情感信息的分析处理与利用, 可以提高耳语交流处理系统的处理能力, 辅助和促进耳语识别系统、耳语合成系统与耳语通信系统的鲁棒性和其它性能^[20]。其次, 研究如何让计算机具有从耳语音信号中自动识别话者情感状态的能力, 是语音情感识别的一个重要组成部分。它可以进一步提高人机交互能力, 结合语义来判断操作者的内心活动, 让计算机真正理解操作者的想法、情绪和态度, 实现真正意义上的人工智能。分析和处理耳语音

收稿日期: 2009-01-12; 修回日期: 2009-03-23

基金项目: 国家自然科学基金项目(60872073, 60472058)、江苏省自然科学基金项目(BK2008291)

作者简介: 金贇(1979-), 男, 浙江乐清人, 博士研究生, 研究方向为语音信号处理。

通讯作者: 金贇, E-mail: jinyun9999@yahoo.com.cn

信号中的情感特征,判断和模拟耳语话者的喜怒哀乐等方面的研究具有理论和应用两方面的重要意义。然而,目前国际国内还没有专门用于耳语音情感识别的数据库。为了进行耳语音情感识别的研究,本文建立了一个耳语音情感数据库和其对应的正常音情感数据库。

2 正常音与耳语音情感特征的对比研究

研究表明,语音的情感信息包含在多种声学参数变化中。由于耳语音独特的发音机理和声学特性,使得耳语音的情感声学参数变化不同于正常语音。因此,有许多正常语音已经研究清楚的问题,在耳语中却不一定适用。比如:

(1) 语音情感识别的声学特征从时间的角度可分为基于语音帧的时序特征和基于语句的全局统计特征。前者广泛运用于语音的内容识别领域,而后者是前者在语句时长内的统计量。由于统计计算可以部分去除由语音内容差异造成对情感表达的影响,因此正常语音的全局统计特征对情感区分度更大。然而对于耳语语音来讲,全局统计特征主要反映韵律特征以及在激励维度上的情感区分能力,其激励源是噪声,声带不振动,韵母部分和浊声母部分基频特性不明显,所以在激励维度上的情感区分能力较正常语音弱。因此全局统计特征对情感区分度是否比正常语音要弱?

(2) 在以语音帧和语句为时长的语音情感特征的选取上,如果特征的时长过小可能会导致识别系统过于关注样本的语义信息,而时长过大会丢失掉重要的韵律学信息,因此有必要确定能有效表达情感的语段长度,使基于该语段的统计特征既能除去文字相关性,又不会减弱重要的韵律特性的表达。所以一般考虑依据韵律词或韵律短语为情感特征语段,那耳语语音的韵律词或韵律短语怎么对应和确定呢?

(3) 根据心理学和韵律学的研究结果,说话者的情感在语音中最直观的表现就是韵律特征和语音质量特征的变化。韵律特征主要有音调、音强和语速等,一般用基频、能量和时长等声学参数描述;音质特征如呼吸声和喉化音等,一般用相关频谱参数描述。在正常语音的情感识别中,由于韵律特征提取较为方便所以使用较多。但在耳语语音中不能只着眼于韵律特征,音质特征所起的作用应该比正常语音要大,那耳语语音音质特征如呼吸声等,

怎样提取是值得研究的问题。

(4) 正常音和耳语音的情感识别都涉及特征提取和模式识别的问题。对于正常音的特征提取方面,依据“激励-评价-控制”情感空间理论,有必要考虑提取语音的音质特征参数作为情感识别用的新的特征参数。此外为了进一步提高系统识别性能,有必要对全局统计特征和时序特征进行融合。但由于耳语音特殊的发音方式,此前一些适用于正常音识别的特征参数就需要重新评估或者寻找新的参数。例如通常用于描述正常音的情感特征参数将无法提取(基音、过零率)或规律发生变化(共振峰)。那么,有什么特征参数可以有效地描述耳语音的情感特性呢?在模式识别方面,传统方法是否依然有效,如果效果不佳,如何改进?

3 数据库制作过程和一般规范

参考国际著名语料库及其相关规范,结合中科院语言研究所李爱军等人关于四大方言普通话语音库的建立过程^[21],本文设计了此次耳语音情感数据库建立的流程,如图1所示。本文首先制定了耳语音情感数据库的规范,包括发音人规范、语料设计规范、录音规范、数据存储技术规范、标注规范和法律声明等等,具体规范及其含义如表1所示。然后进行准备工作,包括软硬件平台的搭建和录音工作开始之前的预备。再进行预采集,对发音人进行初步的语料采集,判断整套采集平台工作状态是否正常,发音人情感表达是否合适,说话方式是否流利自然等等。所有工作准备就绪之后,开始正式采集。为了判断发音人情感表达是否准确,本文进行了隶属度评价,删除情感表达不到位的句子。最

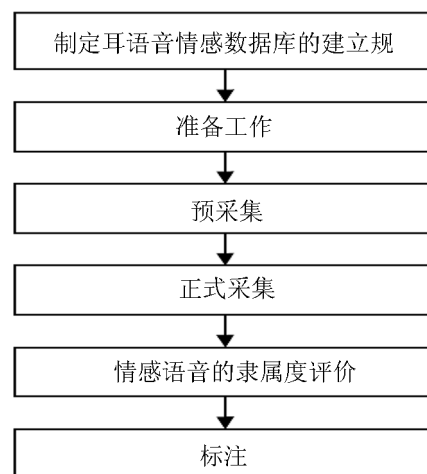


图1 耳语音情感数据库的制作过程
Fig.1 Construction flow of whispered speech emotion database

后对每条语句进行情感标注。

表 1 耳语音情感语音库的制作规范
Table 1 Construction standard of whispered speech emotion database

规范	详细说明
发音人规范	描述发音人的年龄、性别、教育背景
语料设计规范	描述语料的组织和设计内容。包括语料的选择、情感的选择、语料的涵盖范围等。
录音规范	描述录音设备、录音声学环境等技术指标, 录音软件
数据存储	描述采样率、语音文件的
技术规范	存储格式及其技术规范
语料库标注规范	标注内容和标注系统说明
法律声明	发音人录音之后签署的有关法律条文

4 耳语音情感数据库的设计与建设

4.1 发音人的选择和语料的设计

选择 10 个发音人, 五个男声, 五个女声, 年龄在 20-30 岁之间, 都具有本科及以上学历, 对语句的情感理解不存在任何问题。由于语料是具有情感的, 选择的发音人都具有一定的表演才能, 能够充分表现语料所要求的情感。

选择喜、怒、惊、悲、平静五种情感。语音材料分为单词、短句和长句三种类型, 如表 2 所示。对于每种情感, 单词 20 个, 短句 25 句, 段落 6 个, 每位发音者耳语音重复 3 遍, 正常音重复 1 遍(用于后期比较), 共计 9600 条语句。表 3 是耳语音情感数据库的材料组成。

表 2 中文耳语音情感库材料的语料类型
Table 2 Corpora types of mandarin whispered speech emotion database

类型	分类	数目
单词	名词	10
	动词	10
	陈述句	19
短句	疑问句	2
	祈使句	2
	感叹句	2
	平静语气	6
	高兴语气	6
段落	愤怒语气	6
	惊讶语气	6
	悲伤语气	6

表 3 耳语音情感数据库的材料组成
Table 3 Components of whispered speech emotion database

类型	一种情感的数量	情感种类	重复次数	正常音复次数	人数	总计	合计
单词	20	5	3	1	10	4000	
短句	25	5	3	1	10	5000	9600
段落	6	5	1	1	10	600	

对于单词和短句的选择, 主要考虑以下两个方面。第一, 所选择的单词和语句必须不包含某一方面的情感倾向; 第二, 必须具有较高的情感自由度, 对同一个单词或语句能施加各种感情进行分析比较。为了兼顾日常生活中常用的单词, 着重考虑了名词、动词, 名词 10 个, 动词 10 个。形容词附带感情色彩, 所以没有加以考虑。短句的选择按不同句型分类, 着重考虑日常生活中使用频率较高的陈述句, 挑选少量的疑问句、祈使句和感叹句。语料例句中平均句长为 7.8 个音节, 其中最短的短句长 4 个音节, 最长句长 11 个音节, 基本符合平时人们的说话习惯。为了便于比较不同情感各自独有的特征, 对相同语料(单词、短语)设计了 5 种不同的情感语境, 减少由于语料的不同带来的特征差异。

考虑到对一段长时语句进行 5 种情感的表达, 对语料的要求太高。因此选择具有单一情感倾向的段落作为样本, 让表演人员用对应的情感进行表演。表 4 是一个短句“他的总分是 340 分”的四种情感场景设计。

表 4 短句的四种情感场景设计
Table 4 Design of four emotional scenes for one sentence

场景	场景
喜	妈妈兴奋的对同事说: “我家小孩今年中考考得很好, 他的总分是 340 分, 几乎满分了。”
怒	爸爸怒斥小明: “看看你, 就考了 200 分, 你同桌他的总分是 340 分, 你竟然比他差了那么多。你怎么学的?”
惊	“天哪! 他的总分是 340 分! 要知道, 满分才 350 分啊。”
悲	妈妈非常难过, 跟同事说: “小孩考得很差, 他的总分是 340 分, 可能连个大专都上不了。哎, 多年的辛苦付之东流了。”

4.2 耳语音录制的软硬件设备

录音软硬件设备包括: 笔记本电脑 1 台、M-audio mobilePre USB 声卡 1 台、M-audio Nova 大振膜电容话筒 1 部、监听耳机一副等。录音软件采用 Cool Edit pro 2.0, 录音时采用单声道、16 位采样精度、16kHz 采样频率, 录制的语音保存为 wav 格式。

由于耳语音发音时的能量比较小、信噪比低, 周围的环境噪声和回声对录音的影响较大, 因此选择在专业的消音室里进行录音。为了尽可能减少机器噪声和人员的影响, 只有发音人手持麦克风在消音室内录音, 其他一切设备和录音人员都在消音室外工作。

4.3 耳语音情感语料库的命名规则及其标注

为了便于后期的加工和使用, 确定了语音文件的命名规则。一个耳语音文件的名称由三个部分组

成。详细命名规则如表 5 所示。具体示例如图 2 所示,其代表编号为 2 号的女声朗读第 12 个短句,情感为“怒”,第二次重复。

每个语音文件都带有一个标注文件,用于记录语音文件各方面的信息。包括录音时间、地点、环境噪声、发音人的姓名、年龄、性别、学历等信息。

表 5 耳语音库文件命名规则
Table 5 Document name regulation of whispered speech emotion database

第一部分		第二部分		第三部分		
语句信息		录音人信息		情感信息		
句子类型	编号	性别	编号	情感类别	重复次数	
s 单词	1-20	m 男声	1-5	喜	Hap	1-3
m 短句	1-25			怒	Ang	1-3
		惊	Sur	1-3		
l 段落	1-6	f 女声	1-5	悲	Sad	1-3
				平静	Neu	1-3

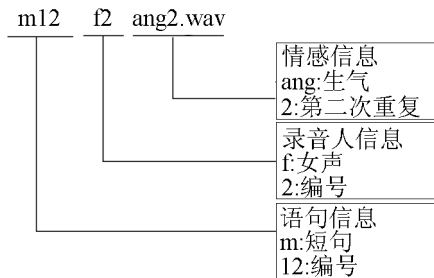


图 2 文件名称规则示例

Fig.2 An example for document name regulation

5 情感耳语音数据有效性的检验

为了检验所录制的情感语音数据的有效性,进行了一个听取实验。

选取 10 名非录制人员作为测试用户,随机听取语音材料,根据自己的主观判断对材料中的情感成分进行评判。然而这些主观评价是模糊性的,为了更加客观地反映主观评价的科学性,本文利用模糊数学中的隶属度函数和层次分析法^[22](Analytic Hierarchy Process, 简称 AHP 法)来进行语料情感的判断。

5.1 隶属度函数的引入

根据心理学家提出的“人区分信息等级的极限能力为 7 ± 2 ”的研究结论,引入九分位的比例标度。本文标度 1、3、5、7、9 表示情感的五种强度,分别对应极弱、较弱、一般、较强、极强五个等级,分别标识为 x_1 、 x_2 、 x_3 、 x_4 、 x_5 , μ_A 为集合 A 的隶属函数, $\mu_A(x_i)$ 表示第 i 个评价等级对模糊集的隶属度。

5.2 层次分析法

AHP 法是美国运筹学家 T.L.Saaty 教授在 70 年

代提出的一种定量与定性分析相结合的多目标决策分析方法。运用 AHP 方法,通过指标之间的两两比较,确定各评价等级的相对重要性,计算不同情感各评价等级的权重。

根据本文标度情感的 5 种强度,两两比较,五种情感强度的重要性依次递增。由此构造了一个判断矩阵 P 。

$$P = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{5} & \frac{1}{7} & \frac{1}{9} \\ 3 & 1 & \frac{1}{3} & \frac{1}{5} & \frac{1}{7} \\ 5 & 3 & 1 & \frac{1}{3} & \frac{1}{5} \\ 7 & 5 & 3 & 1 & \frac{1}{7} \\ 9 & 7 & 5 & 3 & 1 \end{bmatrix}$$

计算特征值和特征向量,得:

$$\lambda_{\max} = 5.2375$$

所对应的特征向量,即权向量为:

$$W = [0.0561 \ 0.1067 \ 0.2170 \ 0.4401 \ 0.8630]^T$$

归一化后得:

$$W_1 = [0.0333 \ 0.0634 \ 0.1289 \ 0.2615 \ 0.5128]^T$$

5.3 一致性检验

计算一致性指标 CI :

$$CI = \frac{\lambda_{\max} - n}{n - 1} = \frac{5.2375 - 5}{5 - 1} = 0.06$$

查表确定相应的平均随机一致性指标 RI :

$$RI = 1.12$$

计算一致性比例 CR :

$$CR = \frac{CI}{RI} = \frac{0.06}{1.12} = 0.053$$

因为 $CR < 0.1$, 说明判断矩阵具有满意的一致性。所以把 $W_1 = [0.0333 \ 0.0634 \ 0.1289 \ 0.2615 \ 0.5128]^T$ 作为情感评价等级的隶属度是合理的。

5.4 评判矩阵的计算

对于某条情感语句,由 10 名测试人员独立给出其所认为的情感类别 i 和相应的情感强度 j , 获得每种情感强度的投票数 n_{ij} , 计算每种情感强度的概率($r_{ij} = n_{ij}/10$), 构成评判矩阵。

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} \\ r_{51} & r_{52} & r_{53} & r_{54} & r_{55} \end{bmatrix}$$

正规化后得 R^* 。

使用加权平均型综合评判函数 f_1 , 得综合评判为:

$$Y_i = R^* \times W_i$$

由此可以得到此语句隶属5种情感的得分,把得分最大项的情感作为最终评判的结果。如果评判结果和实际相符合,则保留;否则,认为此语句的情感特征不明显,剔除。

经过语音情感有效性检验,最终获得符合设计要求的耳语音情感语音数据库。

6 耳语音情感识别

6.1 正常音与耳语音情感的对比

利用耳语音情感数据库的语料,进行初步的耳语音情感识别的研究。

图3是正常音和耳语音的波形对比,内容“大风”,女声,悲伤情感;图4是其相应的语谱图的对比。从图4可见,正常音元音部分的基频非常清楚,而在耳语音中基频已经不可见;正常音的能量主要集中在1kHz以下,而耳语音的能量主要集中在1kHz以上;正常音耳语音的共振峰频率, F_1 、 F_2 和 F_3 分别在1kHz、3kHz和4kHz左右。

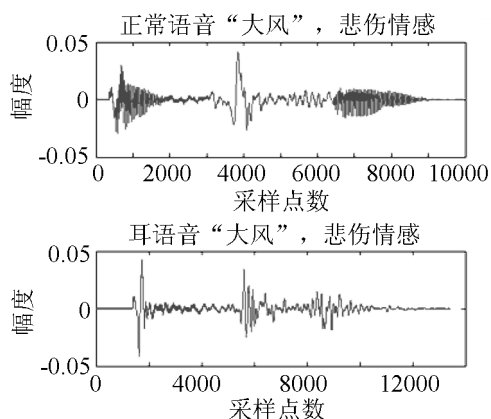


图3 正常音和耳语音波形对比
Fig.3 Waveform comparison between normal speech and whispered speech

6.2 耳语音多种情感的对比

图5是相同语句用不同情感表达的语谱图,内容“今天是晴天”,男声,喜、怒、惊、悲四种情感。从图5可见,相同的语句,“喜”的音长大约是0.75s,“怒”的音长大约是0.9s,“惊”的音长大约是0.95s,而“悲”的音长大约是1.3s。也就是说,“悲”的音长相比其它3种情感是最长的,这点与正常音的特点相符。从能量上看,“悲”的能量最低,而“喜”“怒”“惊”的能量相对而言比较高。因此可以利用音长和能量对耳语音情感进行初步的判断。音长和能量是正常音情感识别的2个重要参数,而这2个参数在耳语音中也依然适用。但由

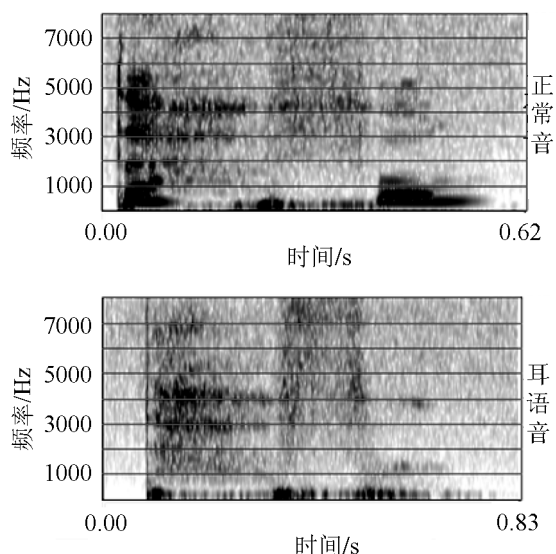


图4 正常音和耳语音语谱图对比
Fig.4 Spectrogram comparison between normal speech and whispered speech

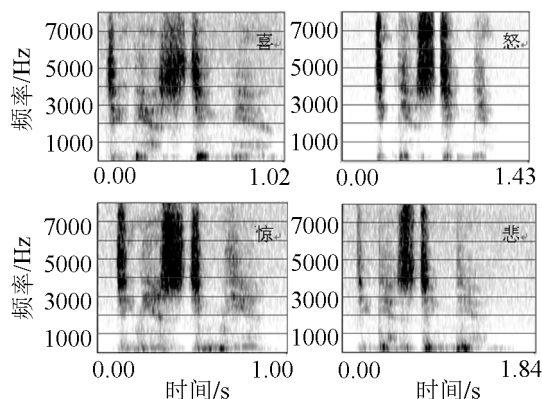


图5 相同语句四种情感语谱图对比
Fig.5 Four emotion spectrogram comparison with same sentence

于耳语音丢失了基频信号,正常音情感识别中的其它重要参数,比如基音的均值、基音的变化范围、基音频率变化、第一共振峰均值、第一共振峰变化范围等,无法直接使用,需要改进或者寻找其它一些替代参数。这些工作需要进一步研究。

7 结论

本文按照语音库制定规范,设计并建立了包括喜、怒、惊、悲、平静五种情感的耳语音情感数据库,包括4000个单词、5000条语句和600个段落。利用模糊数学中的隶属度函数和层次分析法对录制的语音进行有效性分析,得到一个满足规范要求的耳语音情感数据库。

本文比较了正常音和耳语音波形和语谱图的不同之处,也比较了相同的语句用四种不同情感表达的语谱图,得知:部分正常音情感识别的参数依

旧可以作为耳语音情感识别的参数,然而大部分涉及基音的参数可能无法适用于耳语音,需要进行改进或者提取新的替代参数,而这也是下一步工作的重点。

参 考 文 献

- [1] Schwartz M F, Rine M F. Identification of speaker sex from isolated, whispered vowels[J]. Journal of Acoustical Society of America, 1968, 44(6): 1736-1737.
- [2] Tartter V C. What's in a whisper?[J]. Journal of Acoustical Society of American, 1989, 86(5): 1678-1683.
- [3] Tartter V C. Identifiability of vowels and speakers from whispered syllables[J]. Perception and Psychophysics, 1991, 49(4): 365-372.
- [4] Yu H. The whisper is not helpful for treating hoarseness and recovering voice[J]. Journal of the Central University for Nationalities, 1996, 5(2): 163-166.
- [5] Matsuda M, Kasuya H. Acoustic nature of the whisper. EUROSPEECH, 1999: 137-140.
- [6] Tlitch K, Takeda, Itakura F. Acoustic analysis and recognition of whispered speech[A]. Proc. ICASSP[C]. Orlando, Florida, USA, 2002: 389-392.
- [7] 杨莉莉, 李燕, 徐柏龄. 汉语耳语音库的建立与听觉实验研究[J]. 南京大学学报(自然科学版), 2005, 41(3): 311-317.
YANG Lili, LI Yan, XU Boling. The establishment of a Chinese whispered database and perceptual experiment[J]. Journal of Nanjing University (Natural Science), 2005, 41(3): 311-317
- [8] 尹辉, 茹婷婷, 谢湘. 汉语耳语音数字串识别研究[R]. NCMMSC2007, 中国黄山.
YIN Hui, RU Tingting, XIE Xiang. Research on number series recognition over mandarin whispered speech[R]. NCMMSC2007, Mountain Huang, China.
- [9] 吴西愉, 孔江平. 基于小波变换的汉语耳语分析[R]. 第六届全国现代语音学学术会议, 2003 中国天津, 85-90.
WU Xiyu, KONG Jiangping. Mandarin whispered speech analysis based on wavelet transformation[R]. The 6th Conference on Modern Phonetics, 2003, Tianjin, China.
- [10] Gao M. Tones in Whispered Chinese: Articulatory and Perceptual Cues[D]. University of Victoria, Canada, 2002.
- [11] Morris R W, Clements M A. Reconstruction of speech from whispers[J]. Medical Engineering & Physics, 2002, 24(8): 515-520.
- [12] 樊星, 卢晶, 徐柏龄. 汉语耳语音转换为正常音的研究[J]. 电声技术, 2005, (12): 44-47.
FAN Xing, LU Jing, XU Boling. Reconstruction of Normal Speech from Chinese Whispers[J]. Audio Engineering, 2005, (12): 44-47.
- [13] Morris R W. Enhancement and recognition of whispered speech[D]. Georgia Institute of Technology, USA, 2002.
- [14] 沙丹青, 栗学丽, 徐柏龄. 耳语音声调特征的研究[J]. 电声技术, 2003, (11): 4-7.
SHA Danqing, SU Xueli, XU Boling. Study on the characteristics of the tones in whispered Chinese[J]. Audio Engineering, 2003, (11): 4-7.
- [15] GONG Chenghui, ZHAO Heming, LU Gang, LIU Jianxin. An algorithm for formant estimation of whispered speech[A]. The 8th International Conference on Signal Processing[C]. 2006.
- [16] CHEN Xueqin, ZHAO Heming. The research of endpoint detection and initial/final segmentation for Chinese whispered speech[A]. The 8th International Conference on Signal Processing Proceedings[C]. 2006.
- [17] 林玮, 杨莉莉, 徐柏龄. 基于修正 MFCC 参数汉语耳语音的话者识别[J]. 南京大学学报(自然科学版), 2006, (1): 54-62.
LIN Wei, YANG Lili, XU Boling. Speaker identification in Chinese whispered speech based on modified-MFCC[J]. Journal of Nanjing University (Natural Science), 2006, (1): 54-62.
- [18] 韩纪庆, 邵艳秋. 基于语音信号的情感处理研究进展[J]. 电声技术 2006, (5): 58-67.
HAN Jiqing, SHAO Yanqiu. Research progress of emotion processing based on speech signal[J]. Audio Engineering, 2006, (5): 58-67.
- [19] 周洁, 赵力, 邹采荣. 情感语音合成的研究[J]. 电声技术, 2005, (10): 57-59, 73.
ZHOU Jie, ZHAO Li, ZOU Cairong. Study on emotional speech synthesis[J]. Audio Engineering, 2005, (10): 57-59, 73.
- [20] JIANG Danning, CAI Lianhong. Speech emotion classification with the combination of statistic features and temporal features [A]. Proc IEEE Multimedia and Expo[C]. 2004: 1967-1970.
- [21] 李爱军, 王天庆, 殷治刚. 863 语音识别语音语料库 RASC863——四大方言普通话语音库[R]. Report of Phonetic Research 2003.
LI Aijun, WANG Tianqing, YIN Zhigang. 863 speech recognition corpus RASC863--Annotated 4 regional accent speech corpus[R]. Report of Phonetic Research 2003.
- [22] 彭祖赠, 孙韫玉. 模糊数学及其应用[M]. 武汉大学出版社. 2002, 142-162.
PENG Zuzeng, SUN Yunyu. Fuzzy math and its application[M]. Wuhan University Press. 2002, 142-16.