ISSN 1000-0054 清华大学学报（自然科学版）2009 年 第 49 卷 第 S1 期 30/31
CN 11-2223/N J Tsinghua Univ (Sci & Tech), 2009, Vol. 49, No. S1 1413-1418

面向情感变化检测的汉语情感语音数据库

徐 露[1,2]， 徐明星[1,2]， 杨大利[3]

（1. 清华大学 计算机科学与技术系，智能技术与系统国家重点实验室，北京 100084；2. 清华信息科学与技术国家实验室，北京 100084；3. 北京信息科技大学 计算机科学与技术系，北京 100101）

摘 要：该文介绍了面向普通话情感变化检测的情感语音数据库CESD。该数据库的语音以对话形式录制，包括男女声情感对话语音1 200 段。以生气、着急、中性、愉悦、高兴为基本情感，共包含20 种情感变化模式。除语音文件外，还包含带有静音段/有效语音段、情感类别、情感变化段、情感质量等内容的标注文件。为了使更多的研究人员可以使用该数据库，利用Praat 工具提取出67 维常用声学特征，作为特征文件一同存储在该数据库中。对该数据库进行主观评价和情感变化检测的结果表明：语音情感状态自然、情感变化真实，能够满足语音情感识别和语音情感变化检测研究的双重需求。

关键词：语音识别；情感识别；汉语；数据库

中图分类号：TP 391 **文献标识码**：A

文章编号：1000-0054(2009)S1-1413-06

Chinese emotional speech database for the detection of emotion variations

XU Lu[1,2], XU Mingxing[1,2], YANG Dali[3]

(1. State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China; 2. Tsinghua National Laboratory of Information Science and Technology, Beijing 100084, China; 3. Department of Computer Science and Technology, Beijing Information Science & Technology University, Beijing 100101, China)

Abstract: This paper describes a database of emotional speech variations named CESD. The database contains 600 utterances in the form of dialogues with 20 emotional variation modes consisting of 3 different emotions including anger, impatience, neutral, joy, and happiness. Besides the utterances, the database also includes the corresponding label files which include silence or effective speech segments, emotional classes, emotional variation segments, and emotional quality. 67 normal acoustical features are extracted based on the Praat tool and stored in the database. Subjective assessments of the emotional variations demonstrate that the database is suitable for research on speech emotion recognition and emotional variations.

Key words: speech recognition; emotion recognition; Chinese; database

情感语音数据库可以为语音情感的建模提供分析数据，为带情感的语音识别提供训练和测试样本[1]。近年来，越来越多的研究人员致力于语音数据库的建立，充分证明了语音数据库在语音处理领域的重要性。国内外已经成立了多家从事语音数据的收集、整理、发布工作的组织，主要有LDC (Linguistic Data Consortium)，ELRA (European Language Resources Association)，国际中文语言资源联盟(CCC，Chinese Corpus Consortium)，以及我国的社科院语言所(CASS)等[2]。

Ververidis 等[3]对国际上现有的32 个情感语音数据库进行了归纳和总结。这些数据库涉及到十几个语种，其中英语和德语的数据库最多，其次是日语、西班牙语、荷兰语。相对来说，汉语的情感语音数

收稿日期：2009-03-13
基金项目：国家自然科学基金重点项目 (60433030)
作者简介：徐露(1981—)，男(汉)，辽宁，博士后。
通讯联系人：徐明星，副教授，E-mail: xumx@tsinghua.edu.cn

据库较少, Ververidis 只提到了微软建立的一个从电视剧截取录音片段的数据库。随着语音情感研究逐渐受到关注,国内的浙江大学、江苏大学也建立了汉语情感语音数据库。但是这些数据库的规模相对较小,也不能满足情感语音研究的要求。只有建立大规模的、自然真实的情感语音数据库,才能更好地促进情感语音合成、情感语音识别、语音情感识别等领域的研究工作。

为了进行更为深入的语音情感识别研究,清华大学建立了一个情感语音数据库[4]。该数据库采用离散的情感类别标签来描述情感,并选择了高兴、生气、恐惧、悲伤、中性这5 种情感状态。考虑了语音情感识别以及文本内容识别的双重需求,因此既适合进行情感识别又适合语音识别研究。包括25 个男生和 25 个女生的语音,每人 5 种情感,每种情感 200 句,共50 000 句情感语音,是目前国内较大的一个情感语音数据库。

上述数据库虽然在数据规模、情感表现力等方面具有较高的实用性,但是其中的每段语音都只包含了一个人的语音和一种情感状态。在现实应用中,语音情感状态是变化的,并且情感变化往往发生在某个场景的两人或多人对话中,如争吵、辩论、电话聊天等。由于语音中包含的情感状态单一,因此该数据库只适合用于基本情感识别的训练和测试,而不适合于情感变化检测的研究。

本情感对话语音数据库(Chinese emotional speech database, CESD)面向汉语普通话情感变化检测的研究。语音由两个说话人共同录制,内容主要涉及电话服务系统中客户与接线员的对话。语音的采样率为8 kHz, 量化精度为8 bits, 全部数据约200 MB。

本文简要介绍CESD 的设计和建立工作。其中,第1 节介绍了采用的基本情感及其变化模式,第2 节介绍了如何设计录音剧本,第 3 节介绍了数据库的具体内容,第4 节对研究工作进行了总结,并对未来工作进行了展望。

1 情感状态及其变化模式

能否有效地描述情感状态是语音情感识别的前提,也是建立语音情感数据库的关键。因此,采用何种情感描述方法、以及采用哪些基本情感,是建立数据库之前必须解决的问题。

1.1 情感的描述方法

由于人类的情感复杂多样,而且属于学习经验的一个基本方面[5],因此目前还没有一种统一的情感描述方法。常用的方法主要有两种[6]: 一种方法用离散的标签描述情感; 另一种方法则用连续维度来描述。

很多现有的文献都采用若干个离散的情感类别来描述情感,例如生气、高兴、悲伤等。这些描述情感的类别称为情感标签[6],不同文献中定义的情感标签也不尽相同。例如,张石清等[7]对汉语普通话中的生气、高兴、悲伤、惊奇这4 种情感进行了情感识别实验; 而张立华等[8]采用了兴高采烈、愤怒、惊慌、悲伤这4 种情感状态; 罗毅[9]则将惊奇、愤怒、喜悦、悲伤、厌恶这5 种情感作为基本情感。

用情感维度空间描述情感的方法也称为情感维度论。维度论把不同的情感看作是逐渐的、平稳的转变,通过不同情感之间在维度空间中的距离来衡量彼此的相似性和差异性[6]。在实际应用中接受比较广泛的是Cowie[10]提出的基于激发维(Activation)和评价维(Evaluation)的二维情感空间。蒋丹宁等[11]在区分不同情感时,则依据了Pereira[12]提出的“激励- 评价- 控制”三维情感空间理论。

从语音情感识别的研究现状来看,应用更为广泛的是采用情感标签的方法。这种情感描述方法的好处在于,标签式的离散情感更接近现实生活,也更容易为大多数情感状态找到依据,可以获得更具有普遍性的情感语音数据。相比之下,情感维度论虽然可以更科学、更细致的划分人类的情感状态,但是其中的大部分情感很难出现和区分,因此在实际应用中很少使用。这也是大多数现有文献都采用基本情绪理论的原因之一。为了兼顾语音情感变化检测的研究与实际的工程应用,本文也采用离散标签来描述情感状态。

1.2 基本情感的选择

确定了情感的描述方法之后,接下来要解决的问题是如何确定基本情感的数量、以及如何从众多的情感状态中选择合适的基本情感。本文参考现有情感语音数据库中情感状态的使用情况,来确定基本情感状态的数量和种类。表1 记录了各种情感在现有数据库中的数量。

参考该表并结合应用需求,本文选择生气(Anger)、高兴(Happiness)、愉悦(Joy)、着急(Vexation)、以及中性(Neutral)这5 种基本情感建立情感对话语音数据库。其中,中性是指录音者不受任何情绪影响的情感状态。选择这5 种基本情感的

原因在于, 它们是人们在日常生活中最普遍的情感状态, 因此实际出现的概率很大。而且这5 种情感都可以通过简单的激发方式(例如观看带有某种情感的视频), 使普通录音者容易地表现出来。

表1 各种情感在数据库中的数量[3]

情感状态	数量
Anger	26
Sadness	22
Happiness	13
Fear	13
Disgust	10
Joy	9
Surprise	6
Boredom	5
Stress	3
Contempt	2
Dissatisfaction	2
Shame, pride, worry, startle, elation, despair, humor, etc	1

更重要的是, 这5 种情感与实际应用可以紧密地结合起来。以电话服务系统为例, 如果将着急和生气看作负性情感, 愉悦和高兴看作正性情感, 就可以根据情感变化准确地判断出接线员的服务质量。例如, 当接线员的情感是着急或生气时, 可以认为接线员的态度“不好”或“很差”; 而当接线员表现出愉悦或高兴情感时, 可以认为其态度“好”或“很好”。从客户的角度来看, 如果其情感由负性变为正性, 说明接线员很好地解决了客户的问题; 相反, 如果客户的情感由正性变为负性, 说明接线员的工作是失败的。通过上述讨论可知, 本文选择的5 种基本情感符合实际情况、具有可行性。

1.3 情感的变化模式

在现实生活中, 语音情感通常不会长时间保持一种状态不变, 即使是中性情感也很可能因为受到某种刺激而变化为其他情感。这里所谓的“刺激”, 可能是说话者周围环境的改变、说话时间的延长、情绪的变化、他人的语言或动作、以及气味或声音等。例如, 当某人处于高兴状态时, 很可能因为听到一声巨响而变得害怕。

理论上各种情感状态是可以互相转换的, 但实际上很多情感之间不太可能发生直接的转变, 例如从生气变为高兴。虽然汉语中有“悲喜交加”、“喜极而泣”这样的词汇, 但是出现这种情感变化的概率是相当小的。相比之下, 更容易出现的情况是: 随着某个咨询者的问题得以解决, 其情感由着急变为中性; 进而由于得到了非常满意的答复, 其情感又从中性变为高兴。

上述讨论是为了说明: 在5 种基本情感之间, 存在着一定的变化规律, 只有找到这种规律, 才能设计出真实的录音剧本, 从而获得可靠的情感语音。通过对基本情感进行分析, 得到了如图1 所示的5 种情感状态的变化情况。

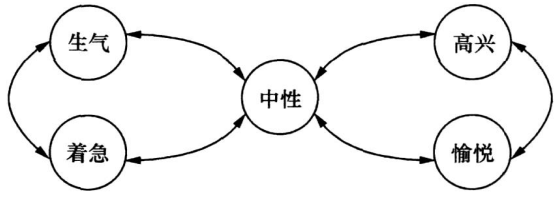

图1 不同情感之间的变化情况

下面要根据情感的变化规律, 确定合理的情感组合模式, 具体情况如下:

1) 如果语音中只包含一种情感状态, 其情感变化为“单一模式”, 即在说话过程中情感没有发生任何变化;

2) 语音中的情感由一种状态转移至另一种状态, 其情感变化为“二组合模式”, 例如着急- 中性、高兴- 愉悦等;

3) 语音中包含三种不同的情感状态, 其情感变化为“三组合模式”, 例如当情感状态出现的次序为高兴、中性、愉悦时, 将三组合模式记为中性- 着急- 生气;

4) 除上述几种情感变化模式外, 语音中还可能包含更多的情感状态, 形成“四组合”、“五组合”等模式, 例如高兴- 愉悦- 中性- 生气、生气- 着急- 中性- 高兴- 愉悦等。

需要说明的是, 单一模式只包含一种情感, 对于情感变化检测的研究没有意义; 包含很多(4 种以上) 情感的变化模式既不现实, 也不容易表演, 故不被采用; 三组合模式可以包含全部的二组合模式, 情感数量合理、情感变化真实, 因此本文采用三组合的情感变化模式。按照图1 所示的情感状态变化, 可以得到20 种符合实际的三组合模式, 包括生气- 着急- 中性, 着急- 生气- 中性等。

2 录音剧本的设计

为了录制对话形式的情感语音, 需要设计不同的场景, 并假定对话者的身份。本文针对电话服务系统, 设计了20 个不同的对话场景, 对话内容根据真实事件改编。每个场景有两个说话人, 其角色假定

为: 发起对话者为客户(Customer), 接受对话者为接线员(Operator)。

2.1 剧本的设计原则

考虑到语音情感识别和情感变化检测研究的双重要求, 剧本设计应该遵循以下几个原则:

1) 尽量用较少的剧本包含所有的情感变化模式, 即用最小的冗余度实现最大的覆盖率, 理想情况是每个剧本包含两种不同的模式, 10 个剧本即可包含所有的情感变化模式;

2) 剧本的场景和对话最好是生活中发生过的事, 情感状态的变化要真实、情感过渡要自然, 这样既容易表演, 又能保证语音数据中情感的真实性;

3) 所选的语句应该能够比较容易地加入录音者的情感, 如果语句是强中性的或者很难强加情感的, 将给录音者的表演造成很大的困难;

4) 对话内容最好使用口语化的句子, 长度控制在15 个字左右, 因为语句太长容易发生情感状态的转移, 语句太短又很难加入不同的情感。

为设计出科学合理的对话剧本, 进行了大量调研, 主要途径有: 到服务行业的营业厅观察记录, 向电话服务系统拨打电话, 上网查找客户服务的常见问题等。然后设计出20 个剧本, 分两组, 每组10 个, 一组剧本即可覆盖20 种情感变化模式。这是经过筛选的结果, 每个对话的场景和内容都与特定的情感变化模式相对应, 情感变化真实且容易表演。

2.2 剧本的格式

剧本分为Page-I 和Page-II 两部分。Page-I 用于对稿或半脱稿表演, 其中包含了剧本的全部对话内容。两个录音者的对话在水平方向相互对应, 并以一定的空白间隔表示出语句的先后顺序。这样的格式可以使说话人在录音时互相对照说话内容, 有助于录音者熟悉和记忆情感状态的变化模式。Page-II 为剧本的大纲模式, 用于脱稿表演。录音者可以对照剧本大纲, 在限定的场景中自由发挥。保留了对话顺序和情感变化, 但不提供对话的具体内容。说话人可以根据排练时的记忆, 在空白处记录若干关键词作为提示, 同时可以互相对照情感变化, 更好地把握情感发生变化的时间(时机)。

剧本格式如表2 和表3 所示。其中, C_m 和O_n 表示说话人的角色和语句的编号, 例如C_3 表示客户说的第3 句话, O_5 表示接线员说的第5 句话。

表2 剧本Page-I

客户对话提示			接线员对话提示		
生气	…… ……	C_1 C_2	O_1 O_2	…… ……	中性
着急	…… ……	C_3 C_4	O_3 O_4	…… ……	愉悦
中性	…… ……	C_5 C_6	O_5 O_6	…… ……	高兴

表3 剧本Page-II

客户对话大纲		接线员对话大纲	
生气	C_1 C_2	O_1 O_2	中性
着急	C_3 C_4	O_3 O_4	愉悦
中性	C_5 C_6	O_5 O_6	高兴

3 CESD 的具体内容

CESD 的建立工作包括语音的采集和数据库的具体内容2 部分。

3.1 语音采集

2 名录音者分别在相隔一定距离的不同房间内使用固定电话进行对话, 录音设备为Samsung YV-120 录音笔。图2 给出了录音设备的连接示意图。参加录音的人员共50 人, 是从在校大学生中按照一定的标准挑选出来的, 要求具有较强的情感表达能力, 普通话尽量标准。

录音时两人一组, 表演2 个场景不同、但包含相同情感变化模式的剧本。每个剧本先表演6 次, 前2 次为对稿表演, 对照剧本Page-I, 同时看到自己和对方的说话内容; 中间2 次为半脱稿表演, 只能看到自己的对话和情感变化; 最后2 次为脱稿表演, 只提供剧本Page-II 以及对话关键词。然后2 人互换角色, 重新表演6 次。这种录音方式的好处是, 录音者

多次表演相同的模式, 可熟练地掌握情感变化, 使语音数据中的情感状态及其变化更真实可信。

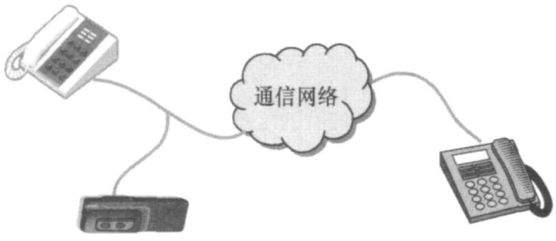

图2 录音设备的连接方法

3.2 情感确认

按照上述过程采集到的语音, 并非每段都适合作为训练或测试样本用于情感变化检测的研究, 原因主要有: 第一, 语音中的情感表达可能不够明确; 第二, 语音中的情感变化可能非常模糊; 第三, 语音中的情感可能过于生硬或做作。完成录音之后, 需要通过人工听取的方式对语音进行听取和确认。

最终建立的数据库CESD 共包含600 段情感对话语音, 每段语音的时间大约为40～60 秒, 其存储格式为*.wav。CESD 中包含3 种类型的文件, 分别是原始的情感对话语音、语音标注文件、以及声学特征文件。图3 说明了CESD 中的文件是如何命名的。

图3 文件的命名方式

3.3 语音标注

语音标注文件以*.xml 格式存储, 标注内容包括以下几个部分:

1) 采用结合短时能量(energy)与过零率(zero-crossing-rate)的端点检测算法[13], 对情感变化语音中的有效语音段和静音段进行划分, 并用起始时间标注静音段和有效语音段;

2) 利用基于局部频谱特征和支持向量机的方法[13], 对有效语音段进行情感识别, 并用情感识别的结果标注有效语音段的情感类别;

3) 参考录音剧本中的情感变化时间范围, 利用自行设计的情感变化检测算法, 检测有效语音段中情感发生变化的范围, 并用其起始时间标注情感变化段;

4) 进行听取确认, 对情感表达不明确、情感变化模糊的语音进行情感质量标注, 利用质量标注, 可以对情感及其变化质量对情感变化检测结果的影响进行分析。

最后人工对情感类别和情感变化段的标注内容进行修正, 以避免情感识别和情感变化检测过程造成的错误。

3.4 声学特征提取

为了使更多的研究人员可以使用CESD, 还编写了Praat[14]脚本程序, 从每段语音中提取出语音识别常用的声学特征, 并将其以*.txt 格式存储在数据库中。声学特征文件可以为不熟悉Praat 工具的研究人员提供方便, 使他们可以容易地利用CESD 进行实验。Praat 是一个功能强大的语音学软件, 具有操作简单、通用性强、可移植性好等优点。如果使用者具有一定的语音学(特别是实验语音学)基础, 就可以方便地利用Praat 工具实现语音处理、语音识别、语音合成等科研工作。

本文借助Praat 工具, 从有效的情感对话语音中提取出以下一些声学特征。

1) 基音频率(pitch): 发浊音时声带的震动频率, 是语音情感识别中的一个重要声学特征。

2) 强度(intensity): 也称为短时平均能量, 语音信号的强度E_n 定义为

$$E_n = \sum_{m=-\infty}^{\infty} x^2(m)h(n-m) = x^2(n) \times h(n), \tag{1}$$

其中: $x(n)$表示语音信号, $h(n)$表示窗函数。

3) 共振峰(formant): 语音信号的共振峰频率可以反映发声时声道的特点。

4) 语音停顿(voice break): 语音中停顿的驻留时间, 描述了语音的清晰程度。

5) 信噪比(HNR): 表示语音信号的周期特性, 可以用于描述语音的声音质量。

6) 基频微扰(jitter): 描述了基频的周期在相邻语音帧之间的微小变化, 计算方法如下:

$$\text{jitter} = \frac{\left|T_i - T_{i+1}\right|}{\frac{1}{N}\sum_{i=1}^{N} T_i}. \tag{2}$$

其中: T_i 为第i 帧的基音周期, N 表示帧数。

7) 振幅微扰(shimmer): 描述了基频的振幅在相邻语音帧之间的微小变化。计算方法如下:

$$\text{shimmer} = \frac{\left|A_i - A_{i+1}\right|}{\frac{1}{N}\sum_{i=1}^{N} A_i}. \tag{3}$$

其中: A_i 为第i 帧的振幅峰值, N 表示帧数。

本文利用不同的统计计量方法,计算了上述7种声学特征的统计量,包括平均值、中位值、标准差、最小值、最大值,以及变化范围。特征的具体内容如表4所示。

表4 用Praat提取的67维特征

特征类别	相关统计量	数量
基音频率	平均值、中位值、标准差、最大值、最小值、变化范围、及其一阶差分的统计量	18
强度	平均值、中位值、标准差、最大值、最小值、变化范围、及其一阶差分的统计量	16
共振峰	前4个共振峰频率的平均值、中位值、标准差、最大值、最小值、变化范围	24
语音停顿	平均值、标准差	2
信噪比	平均值、标准差、变化范围	3
基频微扰	平均值、标准差	2
振幅微扰	平均值、标准差	2

4 结束语

本文介绍了一个情感变化语音数据库CESD。该数据库面向汉语普通话情感变化检测的研究。通过分析5种基本情感之间可能的变化情况,确定了20种真实可行的情感变化模式。针对这些情感变化模式,设计了20个不同的对话场景,语音内容主要涉及电话服务系统中客户与接线员的对话。该数据库共包含1 200段语音,以及相应的标注文件和声学特征文件。

语音数据库的建设未来的研究工作将集中在以下几个方面: 1) 提取更多语音处理的常用特征,例如频谱特征,以合适的存储格式加入CESD; 2) 进一步研究语音情感识别方法,提高情感识别的正确率; 3) 深入分析语音信号中情感变化处的规律和特征,研究有效的情感变化检测方法。

参考文献 (References)

[1] 王青. 基于神经网络的汉语语音情感识别的研究[D]. 浙江: 浙江大学, 2004
WANG Qing. Research on emotion recognition in Chinese speech based on neural network [D]. Zhejiang: Zhejiang University, 1987. (in Chinese)

[2] 潘玉春. 带情感的汉语语音识别研究[D]. 北京: 清华大学, 2006.
PAN Yuchun. Research on emotional speech recognition for Chinese [D]. Beijing: Tsinghua University, 2006. (in Chinese)

[3] Ververidis D, Kotropoulos C. A state of the art review on emotional speech databases [C]//Proc 1st Richmedia Conference. Lausanne, Switzerland, 2003: 109 - 119.

[4] PAN Yuchun, XU Mangling, LIU Linquan, et al. Emotion-detecting based model selection for emotional speech recognition [C]//Proc IMACS Multiconference on Computational Engineering in Systems Applications (CESA). Beijing, China, 2006: 2169 - 2172.

[5] 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, **05**: 58 - 62, 67.
HAN Jiqing, SHAO Yanqiu. Research on progress of emotion processing based on speech signal [J]. *Audio Engineering*, 2006, **05**: 58 - 62, 67. (in Chinese)

[6] 王志良. 人工心理[M]. 北京: 机械工业出版社, 2007.
WANG Zhiliang. Artificial Psychology [M]. Beijing: China Machine Press, 2007. (in Chinese)

[7] 张石清, 赵知劲, 戴育良, 等. 支持向量机应用于语音情感识别的研究[J]. 声学技术, 2008, **27**(1): 87 - 90.
ZHANG Shiqing, ZHAO Zhijin, DAI Yuliang, et al. A study of support vector machine for speech emotion recognition [J]. *Technical Acoustics*, 2008, **27**(1): 87 - 90. (in Chinese)

[8] 张立华, 杨莹春. 情感语音变化规律的特征分析[J]. 清华大学学报(自然科学版), 2008, **48**(S1): 652 - 657.
ZHANG Lihua, YANG Yingchun. Emotional speech characteristics [J]. *J Tsinghua Univ (Sci & Tech)*, 2008, **48**(S1): 652 - 657. (in Chinese)

[9] 罗毅. 一种基于HMM和ANN的语音情感识别分类器[J]. 微计算机信息, 2007, **23**(12 - 1): 218 - 219, 296.
LUO Yi. A human speech emotion recognition classifier based on hidden Markov model and artificial neural network [J]. *Microcomputer Information*, 2007, **23**(12 - 1): 218 - 219, 296. (in Chinese)

[10] Cowie R, Dauglas-Cowie E, Savidou S, et al. FEELTRACE: An instrument for recording perceived emotion in real time [C]//Proc ISCA ITEW on Speech and Emotion: Developing a Conceptual Framework. Newcastle, Northern Ireland, 2000: 19 - 24.

[11] 蒋丹宁, 蔡莲红. 基于语音声学特征的情感信息识别[J]. 清华大学学报(自然科学版), 2006, **46**(1): 86 - 89.
JIANG Danning, CAI Lianhong. Emotion information recognition based on the acoustic features of speech [J]. *J Tsinghua Univ (Sci & Tech)*, 2006, **46**(1): 86 - 89. (in Chinese)

[12] Pereira C. Dimensions of emotional meaning in speech [C]//ISCA Workshop on Speech and Emotion. Belfast, Northern Ireland, 2000: 25 - 28.

[13] HU Hao, XU Mingxing, WU Wei. GMM supervector based SVM with spectral features for speech emotion recognition [C]//Proc 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Honolulu, USA, 2007: 413 - 416.

[14] Boersma P, Weenink D. Praat: Doing phonetics by computer (Version 4.4.20) [EB/OL]. [2008-10-15]. http://www.praat.org.