

文章编号: 1001- 9081(2008) S2- 0360- 03

基于高斯混合模型的语音性别识别

张超琼^{1, 2}, 苗夺谦^{1, 2}, 岳晓冬^{1, 2}

(1 同济大学 计算机科学与技术系, 上海 201804 2 同济大学 嵌入式系统与服务计算教育部重点实验室, 上海 201804)
(zhangchaoqing1@126. com)

摘 要: 利用高斯混合模型 (GMM) 方法进行语音的性别识别。首先概述了特征提取、识别方法及性别识别的过程; 然后通过减少提取特征的语音帧数和降低高斯混合模型的混合阶数来提高性别识别速度; 最后, 将由 Mel 频率倒谱参数 (MFCC) 特征和基音频率特征两种方法得到的测试样本后验概率结合, 提出新的计算测试样本后验概率的方法。实验表明依据此后验概率能有效提高识别的正确率。

关键词: 基音频率; 高斯混合模型; 性别识别; Mel 频率倒谱参数

中图分类号: TP391. 42 **文献标志码:** A

Speech gender recognition based on Gauss mixture model

ZHANG Chao2qiong^{1, 2}, MIAO Duo2qian^{1, 2}, YUE Xiao2dong^{1, 2}

(1 Department of Computer Science and Technology, Tongji University, Shanghai 201804, China;
2 Key Laboratory of Embedded System and Service Computing of Ministry of Education, Tongji University, Shanghai 201804, China)

Abstract In this paper, the method of Gauss Mixture Model (GMM) was used to do the gender identification. First, the method of feature extraction, gender identification and the recognition process are introduced. Second, the number of the frame in the feature extraction is reduced and the mixture number of the GMM was modified to improve the rate of the recognition. At last, the posterior probability of the feature Mel Frequency Cepstral Coefficients (MFCC) was combined with the pitch result to form a new method for computing the posterior probability of the test example. The better correct rate of the recognition was improved by the experiment.

Key words pitch; Gauss Mixture Model (GMM); gender recognition; Mel Frequency Cepstral Coefficients (MFCC)

0 引言

性别识别是语音信号处理中一个很重要的课题, 他与语音识别、说话人识别、语音通信等都有很大的联系。基音频率是性别识别最重要的判别依据。他反映了说话人发浊音时的声带振动频率^[1]。一般而言, 男声的基音频率分布范围为 0~200 Hz, 女声的基音频率分布范围为 200~500 Hz。因此, 准确而可靠地估计基音周期对于说话人性别识别非常重要。

判断说话人的性别可以采用类似说话人识别的方法, 常用的性别识别方法有 VQ 算法、HMM 算法、支持向量机方法等。鉴于性别识别较说话人识别相对简单, 所以本文选用高斯混合模型 (Gauss Mixture Model, GMM) 进行语音性别识别。

1 方法分析

1.1 特征提取方法

基音频率提取包括基音频率候选估计和后处理两个必要步骤^[2-4]。基音频率候选估计法主要有两类: 时域估计法和变换域估计法。常用的时域估计方法有自相关函数法和平均幅度差函数法等; 变换域方法有频域法和倒谱域法等。

本文采用归一化幅度差平方和函数法 (Sum of Magnitude Difference Square Function, SMDSF)^[5]来进行基音周期候选估计, 并利用 viterb 算法进行后处理, 快速、准确地提取基音

频率。SMDSF 法定义为:

$$D_2(S) = \sum_{j=0}^{L-1} [s_2(j+S) - s_1(j)]^2 \tag{1}$$

其中: $s_1(j) = s(j)w_1(j)$, $s_2(j) = s(j)w_2(j)$, $S = 0, 1, \dots, L-1$ 。窗函数为: $w_1(j) = \begin{cases} 1, & j = 0, 1, \dots, L-1 \\ 0, & \text{其他} \end{cases}$ 和 $w_2(j) = \begin{cases} 1, & j = 0, 1, \dots, 2L-2 \\ 0, & \text{其他} \end{cases}$ 。

采取后处理的目的是使用基音周期全局的信息, 纠正基音周期的局部错误, 通过 Viterb 算法可以找到一个最优的基音周期序列, 使得发生基音周期误判错误的损失最小。

算法的总时间复杂度为 $O(TL \ln L)$, 其中 T 是语音信号的总帧数, L 是一帧语音中采样点的个数。与通用基音提取算法相比, 此方法在保证实时性的基础上提高了基音周期提取的正确率。

1.2 GMM 分类器分析

高斯混合模型本质上是一种多维概率密度函数^[6], 一个具有 M 个混合成分的 K 维高斯混合模型, 可以用 M 个高斯成员的加权和来表示, 即:

$$p(x|K) = \sum_{i=1}^M p(x|i|K) = \sum_{i=1}^M c_i p(x|i|K) \tag{2}$$

式中, K 为高斯混合模型的参数集; x 为 K 维的声学特征矢量;

收稿日期: 2008- 05- 05 修回日期: 2008- 07- 11。

基金项目: 国家自然科学基金资助项目 (60475019, 60775036); 2006 年博士学科专项科研基金资助项目 (20060247039)。

作者简介: 张超琼 (1983-), 女, 山西晋中人, 硕士研究生, 主要研究方向: 模式识别、粗糙集理论、粒度计算; 苗夺谦 (1964-), 男, 山西晋中人, 教授, 博士生导师, 主要研究方向: 人工智能、模式识别、数据挖掘、粗糙集理论、主曲线; 岳晓冬 (1980-), 男, 山西太原人, 博士研究生, 主要研究方向: 模式识别, 图像处理, 粗糙集理论, 粒度计算。

c_i 为第 i 个分量的混合权值。

GMM 参数估计从 K 的一个初值开始, 采用 EM 算法估计新的参数 K_c 。GMM 识别原理为: 使得待识别语音特征矢量组 X 具有最大后验概率 $p(K_i | X)$ 的模型 K_i 即为说话人所属的性别组。在男女识别中识别公式可简化为:

$$\hat{i} = \arg \max_i p(X | K_i); i = 1, 2 \tag{3}$$

1.3 系统流程描述

本文利用幅度差平方和函数方法提取训练及测试语音所有帧的基音频率, 分别基于男女训练集特征文件利用 EM 参数估计法建立男女两个高斯混合模型, 然后利用已训练好的两个模型分别对测试集中语音文件计算后验概率值, 后验概率值大的性别类别即为该测试语音文件的性别类型, 最后统计整个测试集的正确率。

2 方法改进

2.1 识别速度改进

鉴于性别识别只需要判断说话人是男性还是女性, 不需要知道说话的内容, 因此我们认为只需要提取较少的语音特征便可以进行识别。这样既可以减少特征提取的时间, 又可以降低训练和识别的数据集。实验表明只提取部分帧数的特征同样可以达到较好的识别效果。特征提取帧数为 100 时男性、女性及平均识别率为: 99.54%, 93.01% 和 96.08%; 帧数为 30 时的结果为: 99.31%, 94.29% 和 96.84%。帧数为 30 时语音持续时间不到 1 秒, 可以大大加快识别速度。

由于性别识别是两类问题, 所以模型的建立不需要像语音识别和说话人识别那样复杂。我们首次实验时选用 16 阶的 GMM。为了提高性别识别系统的训练及识别速度, 我们在修改输入特征文件的大小的基础上尝试降低高斯混合模型的混合阶数, 以使识别算法能达到实时识别的要求。在模型建立的过程中, 我们建立混合阶数不同的多个模型原型, 分别进行训练, 以找到阶数低且识别效果较好的高斯混合模型。

2.2 识别正确率改进

为了提高性别识别的正确率, 我们尝试了如下改进:

1) 修改提取出的基频特征文件, 增加特征文件中基频的特征信息。在原始的特征文件中, 只包含语音文件每一帧的基音频率值, 我们在文件中再加入每帧基频的高阶成分, 即各帧基频的二阶差分。这样每帧语音的特征矢量由一维变为三维。计算高阶基频的方法有很多种, 我们尝试了几几种:

$$\begin{aligned} 1) \quad & \Delta p_i = p_i - p_{i-1}; \\ 2) \quad & \Delta^2 p_i = p_i - 2p_{i-1} + p_{i-2}; \\ 3) \quad & \Delta^3 p_i = p_i - 3p_{i-1} + 3p_{i-2} - p_{i-3}; \\ 4) \quad & \Delta^4 p_i = p_i - 4p_{i-1} + 6p_{i-2} - 4p_{i-3} + p_{i-4}. \end{aligned}$$

上面给出求一阶差分的公式, 二阶差分在一阶差分的基础上利用上面的公式求出。

2) 将基于基频特征的识别结果与基于 Mel 频倒谱参数 (Mel Frequency Cepstral Coefficient, MFCC) 的识别结果相结合^[7]。这样既利用了男女基音频率的差异, 又利用了 MFCC 模拟人耳听觉模型进行识别, 两种特征结合起来共同决策。

MFCC 是目前使用最广泛的语音特征之一^[8-10], 具有计算简单、区分能力好等突出的优点, 因而常常成为许多实际识别系统的首选。我们选取 13 维的 Mel 倒谱系数加上 MFCC 的二阶差分 $\Delta^2 MFCC$ 和 $\Delta MFCC$ 共 39 维参数作为识别语音的特征向量, 同样建立男女两个高斯混合模型, 分别利用两个模型

计算测试集的后验概率值。然后将 MFCC 的两个结果值与基频的两个结果值线性结合起来。

我们设 $S_{p,m}$ 和 $S_{p,f}$ 分别为基于基频特征时利用男、女两个模型得到的测试语音文件后验概率值; $S_{M,m}$ 和 $S_{M,f}$ 分别为基于 MFCC 特征时, 利用男、女两个模型得到的测试文件后验概率值。由于两种特征的结果值之间有量差, 所以我们先利用如下计算将它们归一化为 0 ~ 1 之间的值:

$$S_p^m = S_{p,m} / S_{p,max} \tag{4}$$

$$S_p^f = S_{p,f} / S_{p,max} \tag{5}$$

$$S_M^m = S_{M,m} / S_{M,max} \tag{6}$$

$$S_M^f = S_{M,f} / S_{M,max} \tag{7}$$

其中, $S_{p,max}$ 为所有测试集采用基频特征时的最大后验概率值, $S_{M,max}$ 为采用 MFCC 特征时的最大后验概率值。由于我们判断男女时需要比较男女两个模型后验概率值得大小 (即比较 $S^m - S^f$ 与 0 值的大小), 所以两种特征男女后验概率值差的大小体现了两种特征在最后结果中的影响程度。所以我们计算下面两个差值: $d_1 = S_p^m - S_p^f$ 和 $d_2 = S_M^m - S_M^f$, 利用测试集的平均差值来表示两种方法的影响因子:

$$k_1 = \text{ave}(d_1) / (\text{ave}(d_1) + \text{ave}(d_2)) \tag{8}$$

$$k_2 = \text{ave}(d_2) / (\text{ave}(d_1) + \text{ave}(d_2)) \tag{9}$$

k_1 为基频特征结果的系数, k_2 为 MFCC 特征结果的系数, 且 $k_1 + k_2 = 1$ 。最后的男女后验概率值为两种特征结果值得线性结合:

$$S^m = k_1 S_p^m + k_2 S_M^m \tag{10}$$

$$S^f = k_1 S_p^f + k_2 S_M^f \tag{11}$$

最后, 如果有 $S^m > S^f$, 则该语音说话人为男性; 如果 $S^m < S^f$, 则该语音说话人为女性; 如果 $S^m = S^f$, 则无法判定说话人性别。

2.3 改进算法描述

1) 特征提取。对每个语音文件利用 SMDSF 方法及 Viterbi 算法求得最优基频序列 p_1, p_2, \dots, p_n 。(改进 2.1 节中根据给定的 n 值只提取前 n 帧基频特征)

2) GMM 建立。建立两个高斯混合模型原型, 模型各分量初始均值为 0, 方差为 1。通过 EM 参数估计法求得分量的均值 L_i , 方差 R_i 及混合权值 c_i (改进 2.1 节中采用不同混合阶数的高斯混合原型进行训练)。

3) 根据已建立的男女 GMM 分别计算测试集的后验概率值 $S_{p,m}$ 和 $S_{p,f}$ 。

4) 提取语音文件的 MFCC 特征, 建立相应的男女高斯混合模型, 并计算后验概率值 $S_{M,m}$ 和 $S_{M,f}$ 。

5) 将 3) 和 4) 中计算出的四个后验概率值归一化, 并利用改进方法 2.2 节计算出结合后验概率值: $S^m = k_1 S_p^m + k_2 S_M^m$, $S^f = k_1 S_p^f + k_2 S_M^f$ 。比较 S^m 和 S^f 值, 确定说话人性别。

6) 将识别结果与已标记的真实结果对比, 统计测试集的识别正确率。

3 实验及分析

本实验基于美国宾州大学提供的数库华尔街日报 (Wall Street Journal WSJ) 进行训练和识别。我们选取的男训练集包含 9737 条语音, 女训练集包含 10000 条语音, 测试集也包含 9796 条语音 (其中的男性语音 5051 条, 女性语音 4745 条), 语音的采样频率为 8000 Hz。我们首先利用 C++ 编程

语言^[11]编写提取特征、模型训练及识别的模块,然后在 Linux 环境下利用 perl及 shell脚本进行实验。具体实验如下:

实验 1 我们分别提取语音所有帧特征和前 100 帧、50 帧、30 帧、22 帧及 15 帧的特征进行实验。每次实验高斯混合模型的混合阶数均为 16。实验结果列表如表 1 所示。

表 1 不同帧数的识别结果

| 帧数 | 男 | 女 | 平均 |
|-----|-------|-------|-------|
| 所有 | 99.49 | 93.66 | 96.66 |
| 100 | 99.54 | 93.01 | 96.08 |
| 50 | 99.49 | 94.23 | 96.87 |
| 30 | 99.31 | 94.29 | 96.84 |
| 22 | 99.06 | 93.92 | 96.53 |
| 15 | 98.97 | 93.13 | 96.12 |

从表 1 可以看出,只选取前端部分帧数的语音特征时,我们的识别方法仍然有较好的识别效果。尤其帧数为 30 时,提取特征所用时间少而且识别效果好,识别的速度可以提高。

进一步的实验中我们固定语音的特征提取帧数为 30,分别选取混合阶数为 16、12、9、6、4、2 及 1 的模型实验,实验结果如表 2 所示。

表 2 不同高斯混合阶数识别结果

| 阶数 / 正确率 | 男 / % | 女 / % |
|----------|-------|-------|
| GMM1 | 99.21 | 94.27 |
| GMM2 | 99.27 | 94.76 |
| GMM4 | 99.27 | 94.61 |
| GMM6 | 99.29 | 94.29 |
| GMM9 | 99.36 | 94.23 |
| GMM12 | 99.31 | 94.23 |
| GMM16 | 99.31 | 94.29 |

为了更直观地观察识别结果随混合阶数改变的变化情况我们将实验结果画图表示。

图 1 中横轴表示模型混合阶数,纵轴表示识别正确率,上下两线分别表示男性及女性的识别正确率随混合次数增加的改变。从图 1 我们可以看出随着混合阶数的增大,识别正确率的改变较小,因此我们只需要选取较低的混合阶数。这样的模型训练速度快可以满足一些对时间要求高的实时系统。

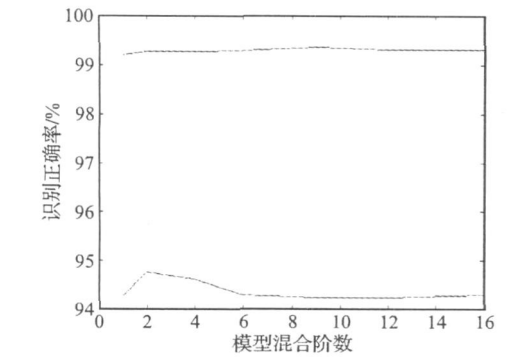


图 1 男女识别率随混合阶数变化图

实验 2 为了进一步证明我们识别方法的泛化能力及实验 1 中改进的可行性,我们利用新的数据库 (English name database)来进行测试,数据库包含近 4000 条语音。我们利用 WSJ 数据库的训练集,每个语音样本只提取前 30 帧特征,分别建立 2、4、6、9、12 及 16 阶的混合模型对新的测试集进行测试,结果如表 3 所示。

表 3 不同阶数模型在新测试集上结果

| 阶数 / 正确率 | 男 / % | 女 / % |
|----------|-------|-------|
| GMM2 | 98.94 | 97.63 |
| GMM4 | 98.80 | 97.12 |
| GMM6 | 98.80 | 97.46 |
| GMM9 | 98.80 | 97.46 |
| GMM12 | 98.94 | 97.46 |
| GMM16 | 98.94 | 97.63 |

表 3 结果表明:我们利用 WSJ 数据库训练集语音前 30 帧特征所建立的模型具有较好的泛化能力,同时表明模型混合阶数从 16 降低到 2 时,对结果的影响很小,说明对于性别识别建立的高斯混合模型不需要很高的混合阶数,2-4 阶即可。

实验 3 根据 2.2 节的改进方法,我们使用增加了高阶特征的基频进行实验,但实验结果提高很小,说明高阶基频对于性别识别的影响很小。

我们将 MFCC 结果与基频结果结合的方法进行试验。选用 39 维的 MFCC 做特征,建立 16 阶 GMM,将 MFCC 实验结果与相同条件下的基频结果线性结合,我们尝试了不同的结合系数。结果如表 4 所示。

表 4 MFCC 和基频结合方法结果

| 结合系数 | 男 / % | 女 / % |
|-------|-------|-------|
| 0.500 | 99.54 | 94.42 |
| 0.300 | 99.56 | 95.43 |
| 0.100 | 99.64 | 98.34 |
| 0.050 | 99.60 | 99.45 |
| 0.030 | 99.58 | 99.68 |
| 0.036 | 99.62 | 99.62 |

只使用基频特征时男、女的识别正确率分别为:99.49% 和 93.66%;只使用 MFCC 特征是结果分别为:93.53% 和 99.37%,从表 4 可以看出,将两种特征结合决策确实提高了识别的正确率,而且根据改进 2.2 利用结果信息选择的结合系数 0.036 的结果较好。实验 3 证明了用两种特征方法共同决策语音文件的说话人性别比只选用一种效果更好。

4 结语

本文利用基音频率作为语音的特征,基于高斯混合模型进行男女识别,并通过改变语音特征的提取帧数及模型的混合阶数来提高识别的速率。最后将基频特征方法和 MFCC 特征方法的结果结合利用投票方式得到最终的识别结果,达到更高的正确率。两种特征的结合只采用了简单的结果线性结合,其他的结合方法还有待进一步的研究。

参考文献:

[1] 杨行俊,迟惠生. 语音信号数字处理 [M]. 北京:电子工业出版社,1995

[2] ROSS M, SHAFFER H. Average magnitude difference function pitch extractor [J]. IEEE Transactions on Acoustics Speech and Signal Processing 1974 22(5): 353- 362

[3] 顾良,刘润生. 高性能汉语语音基音周期估计 [J]. 电子学报, 1999 27(1): 8- 11

[4] 张文耀,许刚,王裕国. 循环 AMDF 及其语音基音周期估计算法 [J]. 电子学报, 2003 31(6): 896- 890

对听阈的值。绝对听阈定义如下:

$$T_{abs}(i) = 3\ 64f^{-0.8} - 6\ 5\exp(f-3\ 3)^2 + 10^{-3}f^4 \quad (21)$$

所以, 最终的掩蔽阈值为:

$$Tc(i) = \max(T(i), T_{abs}(i)) \quad (22)$$

将由式 (22)得到的 Bark 域上的掩蔽阈值 $Tc(i)$ 映射到频域上得到频域掩蔽阈值 $T(w_i)$ 。求出掩蔽阈值 $T(w_i)$ 后由式 (23)可得到进一步增强后的语音谱 $|\hat{S}(w)|$ 。其中 $S^{\wedge}(w_i)$ 是通过卡尔曼滤波增强后的语音谱。

$$|\hat{S}(w_i)| = \begin{cases} |S^{\wedge}(w_i)| @ (A @ (1 + \frac{p_e(w_i)}{A^2(w_i)} + \overline{SNR}^2), & \text{其他} \\ |S^{\wedge}(w_i)| @ (\frac{p_e(w_i)}{A^2(w_i)} + \overline{SNR}^2), & p_e(w_i) < T(w_i) \end{cases} \quad (23)$$

其中: $\overline{SNR} = \frac{\hat{p}_s(w)}{\hat{p}_e(w)}$ 。

4 实验仿真结果

原始语音采用文件 / a favourite topic.wav, 噪声取白噪声, 将他们合成含噪语音信号。采样频率为 16KHz 系统对含噪语音信号进行分帧处理, 帧长为 25ms, 窗函数选择为矩形窗, 帧间重叠度为 40%。在处理中, 含噪语音首先经高通滤波去直流和预加重处理后再进行卡尔曼滤波和感知加权处理。

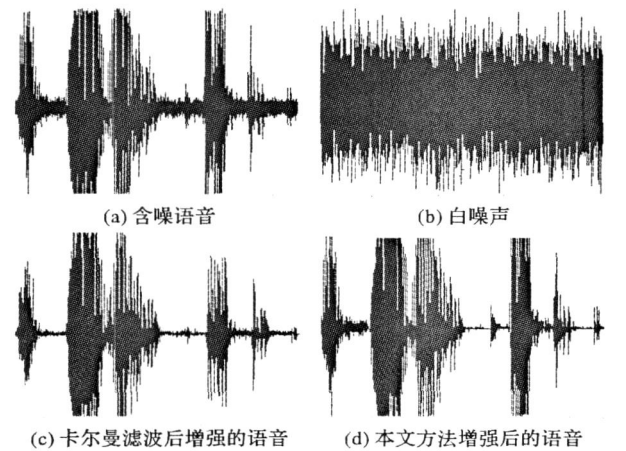


图 1 去噪后的语音时域波形对比

文章给出了白噪声下的去噪对比如图 1, 并将二者与纯净语音相比, 发现本文方法的语音畸变度比原卡尔曼滤波法小。采用单纯的卡尔曼滤波增强法时滤除噪声的同时把一部分语音信号成分也滤除了, 这样导致了语音的失真。结合谱减法 and 听觉掩蔽效应的卡尔曼算法优于传统的卡尔曼滤波法。

5 结语

传统的卡尔曼滤波法只注重提高语音信号的信噪比, 忽略了由此引起的语音失真, 人类听觉的感知特性既能在语音不失真的情况下衰减噪声, 又可以有效地消除残留的音乐噪声。因此结合人耳掩蔽特性的卡尔曼滤波语音增强法将力图找到感知意义 (其标准是基于与人耳感知相关的特性最小化残余噪声) 上的语音失真和噪声去除的最佳均衡。在非线性时域的卡尔曼参数更新过程中, 将参数更新与人耳噪声掩蔽阈值的计算相结合, 从而在较好的去噪的同时减少语音失真。

参考文献:

[1] BEROULT M, SCHWARTZ R, MAKHOUL J. Enhancement of speech corrupted by acoustic noise [J]. Proceedings of IEEE International Conference on Acoustic Speech Signal Processing 1979, 208- 211.

[2] YOU CHANG HUI, KOH NQEE SOO, RAHARDIA SUSANTO. Subband Kalman filtering incorporating masking properties for noisy speech signal [J]. Proc IEEE, 2007, 49: 559- 560.

[3] 姚天任, 孙洪. 现代数字信号处理 [M]. 武汉: 华中科技大学出版社, 1999.

[4] WEI JIANG QING, DULIMIN, YAN ZHAOLI, ZENGHUI. Improved Kalman filter-based speech enhancement with perceptual postfiltering [J]. Chinese Journal of Electronics, 2004, 13(2): 300- 304.

[5] 邓克岩. 基于谱减法的语音增强在 DSP 环境下实时实现的研究 [D]. 兰州: 兰州交通大学, 2006.

[6] VIRAG N. Single channel speech enhancement based on masking properties of the human auditory system [J]. IEEE Transactions on Speech Audio Processing, 1999, 7(2): 126- 137.

[7] M. GABREA. Adaptive Kalman filter-based speech enhancement algorithm [J]. Proc Canadian Conf Electrical and Computer Engineering, 2001, 1: 521- 526.

[8] JAN KYBIC. Kalman Filtering and Speech Enhancement [D]. Master thesis of Czech Technical University (CTU): 45- 46.

[9] 张明友, 吕明. 近代信号处理理论与方法 [M]. 北京: 国防工业出版社, 2005.

[10] K. PALWAL, A. BASU. A speech enhancement method based on Kalman filtering [J]. in Proc. ICASSP'87, 177 - 180.

[11] HASAN M K, ZILANY M S A, KHAN M R. DCT speech enhancement with hard and soft thresholding criteria [J]. Electronics Letters, 2002, 38(13): 669- 670.

[12] 杨行峻, 迟惠生. 语音信号数字处理 [M]. 北京: 电子工业出版社, 1995.

[13] MA. NING, BOUCHARD, MARTIN, GOUBRAN, RAFIK A. Speech enhancement using a masking threshold constrained kalman filter and its heuristic implementations [J]. IEEE Transactions on Audio Speech and Language Processing, 2006, 4: 19- 21.

(上接第 362 页)

[5] 刘建, 郑方, 吴文虎. 基于幅度差平方和函数的基音周期提取算法 [J]. 清华大学学报, 2006, 46(1): 74- 77.

[6] 王永琦, 邓琛, 李世超, 等. 噪声环境中基于 GMM 汉语说话人识别 [J]. 微计算机信息, 2005, 21(11): 177- 179.

[7] HUANG TING, YANG YINGZHU, WU ZHAOHUI. Combining MFCC and pitch to enhance the performance of the gender recognition [C] // Proceedings of the 8th International Conference on Signal Processing. Washington DC: IEEE Computer Society, 2006, 787- 790.

[8] 王让定, 柴佩琪. 语音倒谱特征的研究 [J]. 计算机工程, 2003, 29(13): 31- 33.

[9] 王炳钢, 屈丹, 彭焱. 实用语音识别基础 [M]. 北京: 国防工业出版社, 2005.

[10] MAK B. A mathematical relationship between fullband and multiband mel-frequency cepstral coefficients [J]. IEEE Signal Processing Letters, 2002, 9(8): 241- 244.

[11] LIPPMAN S B, LAJOE J, MOO B E. C++ Primer 中文版 [M]. 李师贤, 蒋爱军, 梅晓勇, 等, 译. 北京: 人民邮电出版社, 2006.