

文章编号:1004-9037(2012)01-0057-06

针对语音变换的语音篡改检测

丁 琦^{1,2} 平西建¹

(1. 解放军信息工程大学信息工程学院, 郑州, 450002; 2. 解放军信息工程大学电子技术学院, 郑州, 450004)

摘要:针对使用语音变换技术的语音篡改,提出一种自动检测方法。在分析语音变换基本模型和变换语音失真的基础上,提取语音信号的声道参数以及相关的信号统计量,并通过支持向量机递归特征消除法,选择出对语音变换比较敏感的特征作为分类特征,使用支持向量机进行语音变换检测和变换语音的说话人性别判别。对于一种语音变换软件的实验结果表明,该方法具有较高的检测准确率,其中语音变换检测的平均准确率为 94.90%,变换语音的说话人性别判别平均准确率为 92.09%。

关键词:语音变换;语音篡改检测;声道参数;信号统计量

中图分类号:TP309

文献标识码:A

Speech Tampering Detection for Voice Transformation

Ding Qi^{1,2}, Ping Xijian¹

(1. School of Information Engineering, Information Engineering University of PLA, Zhengzhou, 450002, China;

2. School of Electronic Technology, Information Engineering University of PLA, Zhengzhou, 450004, China)

Abstract: An automatic speech tampering detection method for voice transformation is proposed. Based on the analysis of the basic model of voice transformation and distortions introduced by it, vocal tract parameters and statistics of speech signals are extracted. Sensitive features are selected as classifying features by support vector machine recursive feature elimination method. Support vector machine classifiers are applied to get the decision of voice transformation detection and speaker gender of transformed speech. Experimental results on a kind of voice transformation software show that the proposed method has high detection accuracy, where the average accuracy is 94.90% for voice transformation and 92.09% for speaker gender detection of transformed speech.

Key words: voice transformation; speech tampering detection; vocal tract parameters, signal statistics

引 言

语音变换是指改变说话人的语音个性特征,使之听起来不像原说话人的声音,但语义信息保持不变。根据变换是否以模拟另一个人的声音为目的,语音变换又可以分为语音转换和语音更改^[1]。语音变换技术的研究引起了很多学者的关注,从1971年Atal研究通过控制线性预测编码(Linear predictive coding, LPC)中声学参数改变语音特性的方法以来^[2],已经有多种语音变换方法提出^[3-6]。这些方法分别从语音产生模型的源或者滤波器模型

入手,使用不同的分析-合成模型,通过各种映射和更改方法,达到改变语音的说话人特征的目的。

语音变换在文语转换、说话人识别、机器翻译、极低速率编码等领域起到重要作用,还可用于医学上声道受损说话人语音恢复、语音通信中说话人伪装等。随着语音变换的应用日渐普及,说话人伪装的功能除用在军事、媒体、娱乐等之外,也常被有不法企图的人利用,用于诈骗、绑架等犯罪行为。相应地,针对语音变换的语音篡改检测也有了更多需求,尤其是用于侦察、取证等。

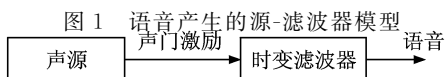
目前对语音变换检测的手段主要依靠人耳听觉,自动检测技术的研究较少。Perrot针对以手掩

口、提高或者降低音调、捏鼻子等声音伪装手段,提出用美尔频率倒谱系数(Mel frequency cepstral coefficients, MFCC)及其差分等特征鉴别声音伪装的方法^[7-8],其中文献[8]中使用 12 阶 MFCC 及其一阶差分作为分类特征,用支持向量机(Support vector machine, SVM)分类器的声音伪装检测方法得到了较好的效果,为针对语音变换的语音篡改检测提供了借鉴。

本文根据对自然语音产生和语音变换基本模型的分析,提取语音的声道参数以及相关的信号统计量,利用 SVM 分类器分别作语音变换检测以及变换语音的说话人性别判别,并通过 AV VCS^[9]软件的 4 种典型变换效果的检测及与文献[8]中的方法比较,验证方法的有效性。

1 语音变换检测算法

自然语音的产生是一个复杂的过程。人在发音时声道的形状和谐振特性发生改变,而声带的振动激发声道中空气产生振动并经口、鼻向外辐射,产生语音。声道的形状和声门处的激励决定了语音的音色。这个过程可以简化为一个源-滤波器模型,即把肺、气管和支气管组成的次声门系统作为产生语音的声源,把声道看作一个时变滤波器。语音则可看作是肺部气体通过声门产生的激励信号通过滤波器后的输出。语音的产生模型框图如图 1 所示^[10-11]。



在这个模型中,滤波器的频率响应取决于声道的形状和大小。声道可以看作是一个时变的不均匀截面的声管,该声管可以用时变的声道面积函数表示。对声道面积函数的精细测量是很困难的,但可以把声道简化为一个级联的无损声管模型,每节声管具有不同的长度和时变的横截面积。对应于语音产生的源-滤波器模型,语音变换技术可以分为 3 类,即源的更改,滤波器的更改,以及源和滤波器的联合更改^[1]。源的更改主要包括时间比例、基音周期和能量的更改等。滤波器更改是指对声道系统频率响应幅度谱的更改。把源和滤波器的更改相结合则可以提高语音变换系统的性能。通过对源和滤波器的更改,改变了与说话人个性特征相关的音段特征和超音段特征,从而使语音不再像原说话人的声音。语音变换的模型框图如图 2 所示。

语音变换技术的研究已经取得了很大进展,但目前的语音变换系统得到的变换语音自然度通常

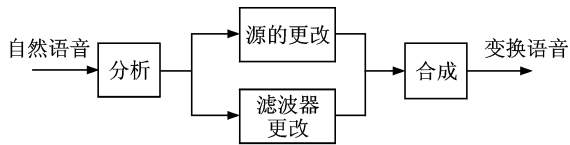


图 2 语音变换基本模型

会有一些程度的降低。造成变换语音失真的可能原因有变换过程中语音频谱细节信息的损失,语音生成模型的简化等;另一方面,自然语音的产生是各发音器官协同工作的结果,任一部分的改变会对其它部分产生相应的影响,对应的语音产生模型各参数之间相互作用相互依赖,任意参数的改变会引起其它参数同时发生非线性变化,通常使用的语音变换方法很难精确地模拟这种协调关系。

在语音产生的源-滤波器模型中,声道参数决定了滤波器的频率响应,对语音的音色起了关键性作用。对于不同性别和年龄段的人群,声道参数会有不同的统计特性;如果利用无损声管模型,则声道的统计特性可以通过声管的横截面积相关参数可以直接反映出来。声管横截面积相关参数包括无损声管模型中各节声管的横截面积,及其引申得到的声腔横截面积和它们在信号持续时间范围内的相关统计量^[12]。

此外,对语音信号的统计分析是估计语音信号特性的重要途径。本文考虑的语音信号统计量包括倒谱参数和线性预测残差信号的峰度、偏度等相关统计量。倒谱参数是一种重要的说话人个性特征参数,主要反映声道响应的特性,而线性预测残差信号中包含较多的激励源信息。

本文利用声管的横截面积相关参数,以及倒谱参数和线性预测残差信号的相关统计量,并通过支持向量机递归特征消除法(SVM recursive feature elimination, SVM-RFE)选择出 7 个对语音变换引起的失真比较敏感的特征,用作语音变换的检测分类特征^[13]。

2 语音变换检测分类特征

2.1 声道参数

在级联无损声管模型基础上进行声道分析,其中声道由 8 节声管组成。根据人的主声道的结构,把这些声管划分成三个声腔:后腔(即咽腔,包括第 1-3 节声管),中腔(包括第 4-6 节声管),前腔(即口腔,包括第 7、8 节声管)。通过声道分析提取声道参数的过程如图 3 所示^[12]。

为了更精确地模拟语音产生过程,图 3 中的线性预测分析是基于基音同步帧的,相邻帧有 50%

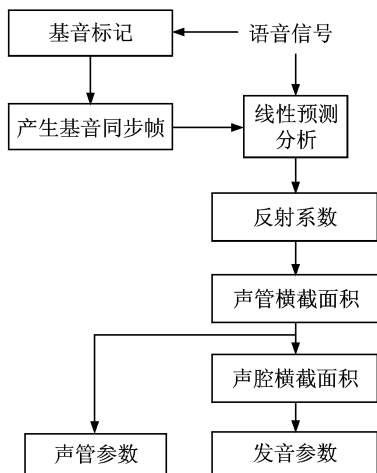


图3 声道参数提取过程

的叠加。得到基音同步帧的方法是先通过自相关法作基音检测得到基音标记,每基音同步帧包含两个连续的基音标记,帧长度为这两个基音标记之间距离的两倍。假设 A, B, C 是三个连续的基音标记位置,则 A, B 所决定的基音同步帧起始位置为 $A - (B - A)/2$,结束位置是 $B + (B - A)/2$,同理 B, C 所决定的基音同步帧起始位置为 $B - (C - B)/2$,结束位置是 $C + (C - B)/2$ 。这两个相邻的基音同步帧之间重叠部分长度为 $(C - A)/2$ 。

假设一段语音信号的基音同步帧数为 N ,对第 n 个基音同步帧 ($n=1, 2, \dots, N$),通过 8 阶线性预测得到各节声管的反射系数 $\mu_i(n)$ ($i=1, 2, \dots, 8$),然后计算各节声管的横截面积

$$A_i(n) = \frac{1 + \mu_i(n)}{1 - \mu_i(n)} A_{i+1}(n), \quad i = 8, 7, \dots, 1 \quad (1)$$

其中 $A_9(n) = 1$ 。定义声腔横截面积为组成该声腔的所有声管横截面积的和,则 3 个声腔横截面积 $S_1(n), S_2(n), S_3(n)$ 可由下式直接得到

$$\begin{cases} S_1(n) = A_1(n) + A_2(n) + A_3(n) \\ S_2(n) = A_4(n) + A_5(n) + A_6(n) \\ S_3(n) = A_7(n) + A_8(n) \end{cases} \quad (2)$$

对语音信号,使用以下声管参数作为特征:

(1) f_1 第一个声管横截面积 $A_1(n)$ 的最大值,反映声门附近振动波幅情况。

$$f_1 = \max_n(A_1(n)) \quad (3)$$

通常成人男性比女性声带明显要长,振动波幅大,因而 f_1 比较大。相对地,老人声带弹性变小,而儿童声门比较窄,因而振动波幅都比较小, f_1 小于成人。

(2) f_2 最后一个声管横截面积 $A_8(n)$ 的均值,反映嘴唇附近横截面积情况。

$$f_2 = \frac{1}{N} \sum_{n=1}^N A_8(n) \quad (4)$$

通常成人女性发音时舌位比男性靠前,开口度也比较大,因而成人女性 f_2 较大。老人和儿童倾向不明显。

(3) f_3 最大横截面积声管位置转移度,反映的是声道形状变化的剧烈程度。

首先对每基音帧找出横截面积最大的声管,并记该声管序号为 $B(n)$,然后由下式得到

$$D_1(n) = |B(n-1) - B(n+1)|/2 \quad (5)$$

其中 $B(0) = B(N+1) = 0$ 。语音信号各帧 $D_1(n)$ 的均值即为最大横截面积声管位置转移度

$$f_3 = \frac{1}{N} \sum_{n=1}^N D_1(n) \quad (6)$$

通常女性声音比男性柔和,发音时声道形状变化没有男性剧烈, f_3 小于男性。老人和儿童倾向不明显。

使用的发音特征参数包括:

(1) f_4 后腔的平均截面积,即 $S_1(n)$ 的均值。

$$f_4 = \frac{1}{N} \sum_{n=1}^N S_1(n) \quad (7)$$

通常成人男性后腔前后径大于女性,所以 f_4 比较大,老人与成人差别不大。儿童后腔较狭窄, f_4 比较小。

(2) f_5 中、后腔横截面积连续相关度,反映声道中的后腔形状变化的短时平稳性。

首先对于每基音帧,计算中、后腔的截面积的差,记为 $E(n)$,然后由下式得到

$$D_2(n) = |E(n-1) - E(n+1)|/2 \quad (8)$$

其中 $E(0) = E(N+1) = 0$ 。然后在语音信号各帧的 $D_2(n)$ 组成的数组中计算连续部分的长度。如果 $D_2(n) \leq 0.25$,则认为第 n 帧是属于连续部分的。假设语音信号中一共有 G 个连续部分,且所有连续部分的帧数之和为 H ,则中、后腔的横截面积连续相关度为

$$f_5 = \frac{G^2}{HN} \quad (9)$$

通过实验统计发现,男性和女性 f_5 比较接近,但语音变换会使 f_5 有较大变化,尤其在变换为男性声音时 f_5 变大比较明显。

2.2 信号统计特征

在信号统计分析之前,首先要对语音信号作预处理和话音激活检测。统计分析仅针对语音帧,每帧长度为预设固定值。通过统计分析,得到以下两个特征参数:

(1) f_6 线性预测残差平均偏度。

对各语音帧作线性预测,并进行线性预测分析

滤波,得到残差信号。假设一段语音信号共有 K 个语音帧,第 k 个语音帧残差信号为 $r_p(k)$ ($k=1,2,\dots,K, p=1,2,\dots,P, P$ 是每帧线性预测残差信号的样点数),则其标准差为

$$\sigma_{\text{LPR}}(k) = \sqrt{\frac{1}{P} \left[\sum_{p=1}^P r_p^2(k) - \frac{1}{P} \left(\sum_{p=1}^P r_p(k) \right)^2 \right]} \quad (10)$$

第 k 个语音帧的线性预测残差偏度为

$$\zeta_{\text{LPR}}(k) = \frac{1}{P} \sum_{p=1}^P \left[\frac{r_p(k) - \frac{1}{P} \sum_{p=1}^P r_p(k)}{\sigma_{\text{LPR}}(k)} \right]^3 \quad (11)$$

整个语音信号的线性预测残差平均偏度由所有语音帧的 $\zeta_{\text{LPR}}(k)$ 平均得到,并取绝对值

$$f_6 = \left| \frac{1}{N} \sum_{k=1}^K \zeta_{\text{LPR}}(k) \right| \quad (12)$$

通常男性 f_6 大于女性,原因与男性声带的紧张程度较弱有关。

(2) f_7 平均倒谱偏度。

平均倒谱偏度由每个语音帧的倒谱系数得到。对于第 k 个语音帧,首先计算倒谱系数 $c_m(k)$, ($m=1,2,\dots,M, M$ 为倒谱系数个数)。 $c_m(k)$ 的标准差为

$$\sigma_{\text{cep}}(k) = \sqrt{\frac{1}{M} \left[\sum_{m=1}^M c_m^2(k) - \frac{1}{M} \left(\sum_{m=1}^M c_m(k) \right)^2 \right]} \quad (13)$$

该帧的倒谱偏度为

$$\zeta_{\text{cep}}(k) = \frac{1}{M} \sum_{m=1}^M \left[\frac{c_m(k) - \frac{1}{M} \sum_{m=1}^M c_m(k)}{\sigma_{\text{cep}}(k)} \right]^3 \quad (14)$$

整段语音信号的平均倒谱偏度即为 $\zeta_{\text{cep}}(k)$ 的均值

$$f_7 = \frac{1}{K} \sum_{k=1}^K \zeta_{\text{cep}}(k) \quad (15)$$

平均倒谱偏度可以反映语音清晰度的失真程度。实验统计发现,男性和女性 f_7 比较接近,但语音变换会使 f_7 有明显变化,尤其在变换为男性声音时 f_7 明显变大。

3 实验结果

3.1 实验设置

实验中的自然语音数据来自“863”语音识别语料库。实验中使用了 100 个说话人(50 男 50 女)的语音数据,对每个说话人分别取 20 句话,共得到 2 000 个自然语音样本。语音数据为 16 kHz 采样, 16 bit 线性编码。

语音变换工具采用的是 AV VCS 软件,实验

中选取了软件提供的 4 种有代表性的变换效果,包括男性、女性、女孩、老人(男性)。不同变换效果通过软件中不同的参数设置得到,其中这 4 种变换效果的主要参数设置如表 1 所示,表中数字均以百分比表示,指的是该参数在原始语音基础上的加权系数。同一种参数设置得到的输出语音的听觉效果会随原始语音的说话人不同而有差异,但说话人的语音特性例如性别、年龄特征已经完全改变,而且具有预设的变换效果。该软件的语音变换方法属于源和滤波器联合更改类型。

表 1 4 种变换效果的主要参数设置 %

| 变换效果 | 参数设置 | | |
|--------|------|-----|-----|
| | 基音 | 音色 | 高级 |
| 女性 | 128 | 100 | 0 |
| 男性 | 59 | 147 | 100 |
| 女孩 | 161 | 72 | 55 |
| 老人(男性) | 55 | 110 | 100 |

对所有的自然语音通过上述语音变换工具分别得到 4 种变换效果的变换语音,共 8 000 个变换语音样本。

本文使用 SVM 分类器^[14],其中核函数为径向基函数(Radial basis function, RBF),参数的最优化使用网格搜索的方法。

在进行语音变换检测实验时,把自然语音样本中 60 个说话人(30 男 30 女)的语音样本(共 1 200 个)作为训练样本中的第一类,对应的变换语音样本(共 4 800 个)作为训练样本中的第二类,对每个样本分别如第 3 部分所述提取 7 个特征,作为 SVM 分类器的训练数据。另 40 个说话人(20 男 20 女)的自然语音样本(共 800 个)和变换语音样本(共 3 200 个)则作为测试样本,使用 SVM 分类器对测试样本的分类特征进行检测。

在语音篡改检测中,对于变换语音常常还需要确定说话人的真正性别,因此我们进一步作了变换语音的说话人性别判别。由于在声道参数,以及倒谱和线性预测残差信号的高阶统计量中,也包含了性别在内的说话人重要信息,因而在变换语音的说话人性别判别实验中仍使用了第 3 部分所述分类特征。在进行变换语音的说话人性别判别实验时,把 30 个男性说话人的自然语音样本和变换语音样本作为训练样本中的第一类(共 3 000 个),30 个女性说话人的自然语音样本和变换语音样本作为训练样本中的第二类(共 3 000 个),对每个样本分别提取分类特征,作为 SVM 分类器的训练数据。另 40 个说话人(20 男 20 女)的变换语音样本(共 3 200 个)则作为测试样本,使用 SVM 分类器对测

试样本的特征进行检测。

作为比较,还使用上述训练样本和测度样本检验了文献[8]中使用12阶MFCC及其一阶差分作为分类特征的声音伪装检测方法,其中分类器为SVM分类器。

3.2 特征矢量分布

图4所示为训练数据中男性自然语音变换为女性声音时,变换语音与自然语音(包括男声和女声)的分类特征矢量分布,其中星形表示自然语音,圆圈表示变换语音。图4(a)中显示的是 $\{f_1, f_2, f_3\}$,4(b)中显示的是 $\{f_4, f_5\}$,4(c)中显示的是 $\{f_6, f_7\}$ 。由图可见,利用所提特征矢量可以较好地区分自然语音和变换语音。

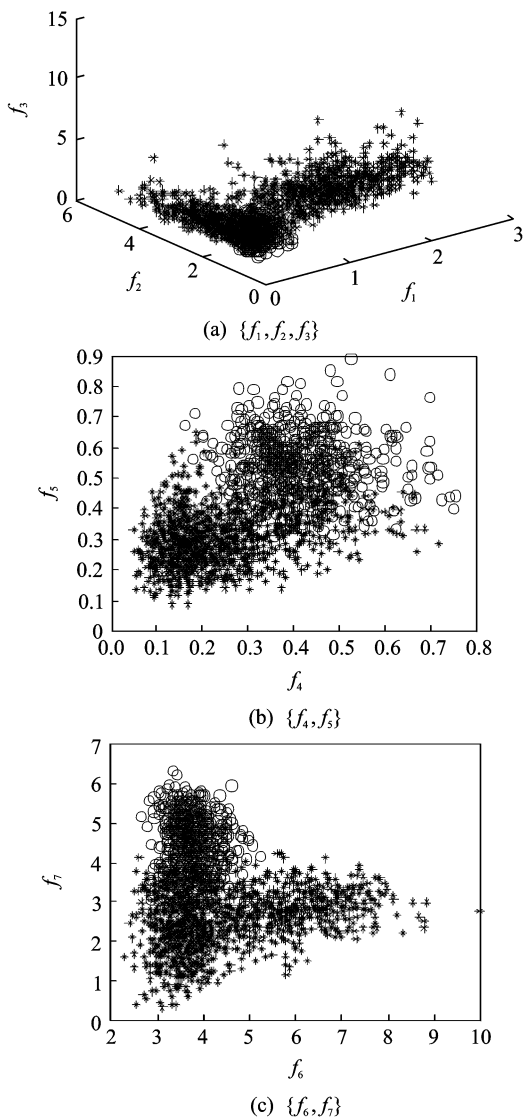


图4 变换语音(女声变换为男声)与自然语音(包括男声和女声)的特征矢量分布

其他几种变换效果的变换语音特征矢量分别分布在矢量空间中的不同位置范围,除了男声变换

为女孩声音的情况外,均与自然语音有很好的区分度。在男声变换为女孩声音情况下,部分人发音的一些特征会变得与自然语音的女声比较接近。

3.3 检测结果

3.3.1 语音变换检测

语音变换检测的结果如表2所示。由表2可见,本文方法除了在男声变换为女孩声音的情况下检测准确率只有70.5%之外,在其它情况下检测准确率均在90%以上。在男声变换为女孩声音情况下,部分人发音的一些特征变得与自然语音的女声比较接近,在矢量空间中有一定程度的混淆,导致误判,因而准确率较低。

表2 语音变换检测准确率 %

| 变换效果 | 原始语音类别 | 本文方法 | | 文献[8]方法 | |
|------|--------|-------|---------|---------|---------|
| | | 分类检测 | 按变换效果平均 | 分类检测 | 按变换效果平均 |
| 女性 | 女性 | 98.5 | 94.5 | 91.25 | 56.75 |
| | 男性 | 90.5 | | 22.25 | |
| 男性 | 女性 | 99.5 | 99.625 | 77 | 88.5 |
| | 男性 | 99.75 | | 100 | |
| 女孩 | 女性 | 100 | 85.25 | 100 | 87.625 |
| | 男性 | 70.5 | | 75.25 | |
| 老人 | 女性 | 99.75 | 99.875 | 100 | 100 |
| | 男性 | 100 | | 100 | |
| 自然语音 | 女性 | 92.5 | 95.25 | 100 | 99.875 |
| | 男性 | 98 | | 99.75 | |
| 总平均 | | | 94.90 | | 86.55 |

比较而言,文献[8]方法在一些情况下得到非常高的准确率,而在另一些情况下尤其是男声变换为女声时准确率偏低,造成这种情况的原因是MFCC及其差分参数对于说话人特性等因素比较敏感,当各种说话人类别和变换效果的语音特征混在一起训练时,会造成部分变换语音和自然语音混淆而难以区分。

从各种变换效果的平均准确率来看,本文方法对各种效果的语音变换检测准确率均大于85%,表现出比较稳定的性能。文献[8]方法虽然对女孩、老人这两种变换效果以及自然语音的检测准确率略高于本文方法,但整体性能不够稳定,特别是对于女性这种变换效果的平均检测准确率偏低。

3.3.2 变换语音的说话人性别判别

变换语音的说话人性别判别结果如表3所示。

由表3中数据可见,本文方法在变换语音的说话人性别判别中仍然表现出比较稳定的性能,对

表 3 变换语音的说话人性别判别准确率 %

| 变换效果 | 原始语音类别 | 本文方法 | | 文献[8]方法 | |
|------|--------|-------|---------|---------|---------|
| | | 分类检测 | 按变换效果平均 | 分类检测 | 按变换效果平均 |
| 女性 | 女性 | 100 | 98.75 | 100 | 98.5 |
| | 男性 | 97.5 | | 97 | |
| 男性 | 女性 | 88.5 | 89 | 54.25 | 77.125 |
| | 男性 | 89.5 | | 100 | |
| 女孩 | 女性 | 97.5 | 89.125 | 100 | 98.5 |
| | 男性 | 80.75 | | 97 | |
| 老人 | 女性 | 86 | 91.5 | 81.75 | 90.875 |
| | 男性 | 97 | | 100 | |
| 平均 | | 92.09 | | 91.25 | |

种变换效果的平均准确率均在 89% 以上。文献[8]方法的平均准确率对女孩这种变换效果高于本文方法,但由于 MFCC 及其差分参数对于说话人特性等多种因素的敏感性,当其用于这样多种说话人类别和变换效果混合的情况时,仍出现了性能不够稳定的情况,特别是对于女声变换为男声的情况检测准确率偏低,只有 54.25%。

总体来说,本文所用声道参数和信号统计量比较适合针对语音变换的语音篡改检测,在语音变换检测和变换语音的说话人性别判别中均有比较稳定的性能。

4 结束语

本文总结了语音变换技术的基本模型,分析了目前的语音变换系统引起变换语音自然度下降的原因,在语音的声道分析和信号统计分析基础上,提取出对语音变换的失真比较敏感的特征,并通过 SVM 分类器进行语音变换检测,以及变换语音的说话人性别判别。实验结果表明语音的声道参数,以及线性预测残差平均偏度和平均倒谱偏度可以用作针对语音变换的语音篡改检测,具有比较稳定的性能。下一步的工作将是进一步完善特征集,提高检测方法的准确性和适用性,实现适用于各种语音变换软件和硬件的检测方法,满足实际应用的需求。

参考文献:

[1] Stylianou Y. Voice transformation: a survey[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. Taipei, Taiwan: IEEE Press, 2009:3585-3588.

[2] Atal B S, Hanauer S L. Speech analysis and synthesis by linear prediction of the speech wave[J]. The Journal of the Acoustical Society of America, 1971,

50(2):637-655.

- [3] Desai S, Raghavendra E V, Yegnanarayana B, et al. Voice conversion using artificial neural networks [C]//IEEE International Conference on Acoustics, Speech and Signal Processing. Taipei, Taiwan: IEEE Press, 2009:3893-3896.
- [4] Erro D, Moreno A, Bonafonte A. Voice conversion based on weighted frequency warping [J]. IEEE Transactions on Audio, Speech, and Language Processing, 2010,18(5):922-931.
- [5] Qiao Y, Saito D, Minematsu N. HMM-based sequence-to-frame mapping for voice conversion[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. Dallas, Texas, USA: IEEE Press, 2010:4830-4833.
- [6] Kawahara H, Nisimura R, Irino T, et al. High-quality and light-weight voice transformation enabling extrapolation without perceptual and objective breakdown[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. Dallas, Texas, USA: IEEE Press, 2010:4818-4821.
- [7] Perrot P, Morel M, Razik J, et al. Vocal forgery in forensic sciences [J]. Forensics in Telecommunications, Information and Multimedia, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2009,8(1):179-185.
- [8] Perrot P, Chollet G. The question of disguised voice [J]. The Journal of the Acoustical Society of America, 2008,123(5):3878.
- [9] AV voice change software[EB/OL]. <http://www.audio4fun.com/voice-changer.htm>, 2005-01-01/2009-06-08.
- [10] Rabiner L R, Schafer R W. Digital processing of speech signals[M]. Englewood Cliffs, New Jersey: Prentice Hall, 1978.
- [11] Ismail M A. Vocal tract area function estimation using particle swarm[J]. Journal of Computers, 2008, 3(6):32-38.
- [12] Single-ended method for objective speech quality assessment in narrow-band telephony applications[S]. ITU-T Rec P563, 2004.
- [13] Guyon I, Weston J, Barnhill S, et al. Gene selection for cancer classification using support vector machines[J]. Machine Learning, 2002, 46(1-3):389-422.
- [14] Chang C C, Lin C J. LIBSVM-A library for support vector machines [EB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2000-01-01/2009-05-08.

作者简介:丁琦(1976-),女,博士研究生,研究方向:语音处理、音频取证与信息隐藏,E-mail:gogodq@126.com;平西建(1953-),男,教授,研究方向:图像处理、模式识别、信息隐藏。