

# 利用谱包络变换后 LPC 系数实现频谱搬移

张辉, 李波, 王宝良

(空军工程大学 电讯工程学院, 陕西 西安 710077)

**摘要:** 频谱包络转换是语音转换中的一项重要内容, 基于双线性转换函数进行频谱搬移的方法可以有效的进行谱包络转换, 并且基于小语料库训练时仍能得到稳定的转换谱包络。本文基于 LPC 分析, 在求解频谱搬移后的系统单位冲激响应的基础上, 提出求解频谱搬移后 LPC 系数的方法, 以此来实现频谱包络的变换, 取得较好的效果。

**关键词:** 语音转换; 频谱包络转换; LPC 分析

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1009-3516(2006)06-0062-03

语音转换 (VC, Voice Conversion 或 VT, Voice Transformation) 是指改变一个说话人 (源说话人, Source Speaker) 的语音个性特征, 使之具有另外一个说话人 (目标说话人, Target Speaker) 的语音个性特征<sup>[1-5,7-8]</sup>。语音包含很多有用信息, 其中最主要的是语义信息, 另外一个很重要的信息为语音的个性化信息。语音转换就是要保留原有语义信息不变, 而改变语音的个性化信息, 使一个人的语音经语音转换后听起来象是另一个人说的语音。

语音转换是一项比较复杂的技术, 谱包络的转换是语音转换的一个重点和难点, 语音工作者研究的谱包络转换的实现方法有多种, 如矢量量化法<sup>[1]</sup>、说话人插值法<sup>[2]</sup>、线性多变量回归法 (LMR, linear multivariate regression)、动态频率规整法 (DFW, dynamic frequency warping)<sup>[5]</sup>、神经网络法<sup>[4]</sup>、高斯混合模型法 (GMM, Gaussian Mixed Model)、HMM 法、基于双线性转换函数法<sup>[6]</sup>等, 以及 GMM 法与 DFW 法相结合的方法、GMM 法与 MAP 自适应相结合的方法。文献[6]采用基于双线性转换函数的方法来实现男声和女声之间的相互转换时的频谱搬移, 我们也研究采用基于双线性函数的方法来实现不同人之间的频谱搬移, 基于双线性函数进行谱包络变换的优点是在基于小语料库训练的情况下仍能得到稳定且比较准确的转换谱包络。

文献[6]采用的是基于求解频谱搬移后的系统单位冲激响应的方法来实现谱包络的搬移, 文献[9]给出了两种求解频谱搬移后系统单位冲激响应的有效方法, 研究发现, 该方法存在两个明显的缺点: 一是系统单位冲激响应一般只有取的较长时收敛特性才较好, 尤其基音周期较小的女声语音更是如此; 二是为计算方便在与 LPC 残差进行卷积时常要对系统单位冲激响应进行截断处理, 这样得到的系统单位冲激响应的收敛特性常不好, 与其对应的幅度谱常有较明显的振荡现象。为了克服这些问题, 我们提出基于频谱搬移后的系统单位冲激响应求 LPC 系数的方法, 实现频谱包络转换, 再合成出频谱搬移后的语音信号。

## 1 双线性转换函数

双线性转换函数 (BLT, bilinear transform) 可用下式表示:

$$\hat{z} = Q'(z) = \frac{z - \beta}{1 - \beta z} = \frac{1 - \beta z^{-1}}{z^{-1} - \beta} \quad \beta \in R \quad (1)$$

双线性转换函数的频率弯折关系为

$$\hat{w} = \arctan\left(\frac{\sin w - \beta^2 \sin w}{\cos w + \beta^2 \cos w - 2\beta}\right) = \arctan\left(\frac{(1 - \beta^2) \sin w}{(1 + \beta^2) \cos w - 2\beta}\right) \quad (2)$$

收稿日期: 2006-05-15

作者简介: 张辉 (1976-), 男, 江苏镇江人, 讲师, 主要从事无线通信研究。

对应不同的频率弯折系数  $\beta$  得到的频谱弯折曲线如图 1 所示

当  $\beta > 0$  时, 频谱被向高频处搬移, 而当  $\beta < 0$  时, 频谱被向低频处搬移。

在实际的频谱搬移过程中, 只需在声道传输函数中如式(3)进行替换就可实现频谱的搬移处理。

$$z = Q(\hat{z}) = (\beta + \hat{z}) / (1 + \beta \hat{z}) \quad (3)$$

### 2 求解谱搬移后的 LPC 系数

先将语音信号进行 LPC 分析, 基于 LPC 系数进行频谱搬移, 利用在文献[9] - [10]中提出的方法得到频谱搬移后的系统单位冲激响应, 再对系统单位冲激响应进行 LPC 分析, 即得到频谱搬移后的 LPC 系数, 将原语音信号的 LPC 残差与频谱搬移后的 LPC 系数进行线性预测合成, 得到频谱搬移后的语音信号。

图 2 为对一女声语音先求频谱搬移后的系统单位冲激响应, 其中图 (a) 为原语音分析信号; (b) 为原语音单位冲激响应; (c) 为原语音单位冲激响应幅度谱; (d) 为  $\beta = -0.2$  时系统单位冲激响应; (e) 为  $\beta = -0.2$  时单位冲激响应幅度谱; (f) 为  $\beta = -0.2$  时由单位冲激响应得到的 LPC 谱包络; (g) 为  $\beta = 0.2$  时系统单位冲激响应; (h) 为  $\beta = 0.2$  时单位冲激响应幅度谱; (i) 为  $\beta = 0.2$  时由单位冲激响应得到的 LPC 谱包络。从图中可以看出单位冲激响应的幅度谱有很明显的振荡, 从基于频谱搬移后的 LPC 谱包络, 可以有效的得到频谱搬移后的 LPC 系数。

图 3 是对一男声语音先求频谱搬移后系统单位冲激响应, 再求 LPC 参数的结果示意图。其中图 (a) 为原语音分析信号; (b) 为原语音单位冲激响应; (c) 为原语音单位冲激响应幅度谱; (d) 为  $\beta = -0.2$  时系统单位冲激响应; (e) 为  $\beta = -0.2$  时单位冲激响应幅度谱; (f) 为  $\beta = -0.2$  时由单位冲激响应得到的 LPC 谱包络; (g) 为  $\beta = 0.2$  时系统单位冲激响应; (h) 为  $\beta = 0.2$  时单位冲激响应幅度谱; (i) 为  $\beta = 0.2$  时由单位冲激响应得到的 LPC 谱包络。

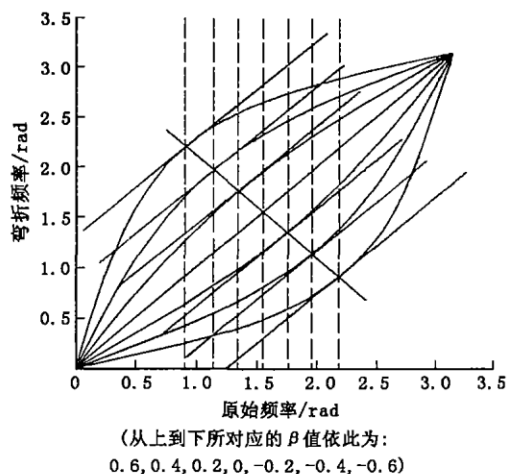


图 1 对应不同  $\beta$  值的频率弯折曲线图

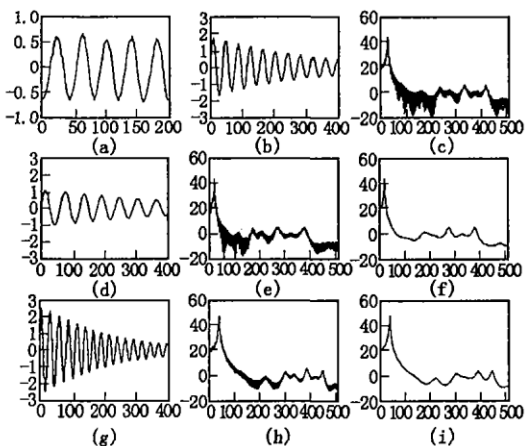


图 2 一女声频谱搬移效果图

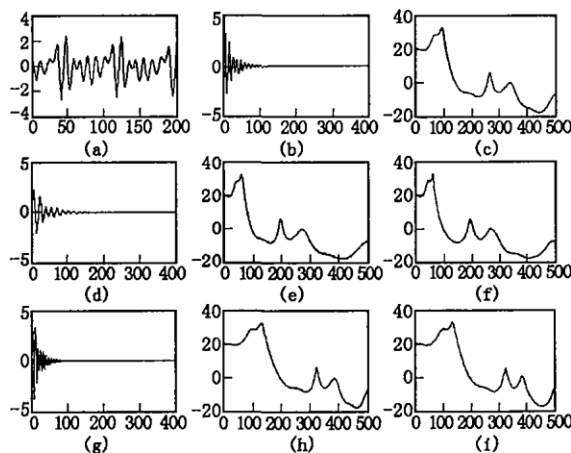


图 3 一男声频谱搬移效果图

### 3 频谱搬移实验

我们采用图 4 所示的方法进行频谱搬移试验, 实验语音为实验室中采集所得的一段男声语音和一段女声语音, 语音信号采样频率为 11 025 Hz, 量化精度为 16 bit。从试验结果可以看出, 当  $\beta > 0$  时, 频谱被有效的向高频处搬移, 而当  $\beta < 0$  时频谱被有效的向低频处搬移; 测听效果显示, 进频谱搬移后音质发生明显的

变化,当 $\beta > 0$ 时,声音变的尖锐,而当 $\beta < 0$ 时,声音变的低沉,比较成功的实现频谱包络的搬移。

同时通过测听显示,对男声处理后的音质仍保持的较好,这与基于系统单位冲激响应<sup>[9-10]</sup>的方法基本相似;而对女声处理后,音质较基于系统单位冲激响应的方法有所提高,但仍有一定的失真现象。我们认为这是由于女声语音得到的LPC谱包络非常尖锐,不能很好的表示谱包络,LPC残差与频谱搬移后的LPC系数进行线性预测合成时就不可避免的引入失真所致。

## 4 结束语

针对语音转换中频谱包络搬移的问题,本文在以前提出的基于系统单位冲激响应进行谱包络搬移的基础上,提出基于频谱搬移后系统单位冲激响应求解频谱搬移后LPC系数的方法。该方法克服了系统单位冲激响应方法存在的两个缺点,实验表明,该方法可以比较成功的实现频谱包络的转换,语音质量有所提高。

### 参考文献:

- [1] Arslan L M. Speaker Transformation Algorithm Using Segmental Codebooks (STASC) [J]. *Speech Communication*, 1999, 28(3): 211 - 226.
- [2] Iwahashi N, Sagisaka Y. Speech Spectrum Conversion Based on Speaker Interpolation and Multi - Functional Representation With Weighting by Radial Bases Function Networks [J]. *Speech Communication*, 1995, 16(2): 319 - 326.
- [3] Mizuno H, Abe M, Voice Conversion Algorithm Based on Piecewise Linear Conversion Rules for Formant Frequencies and Spectrum Tilt [J]. *Speech Communication*, 1995, 16(2): 125 - 126.
- [4] Narendranath M., Murthy H A, Rajendran S. Transformation of Formants for Voice Conversion Using Artificial Neural Networks [J]. *Speech Communication*, 1995, 16(2): 207 - 216.
- [5] Valbret H, Moulines E, Tubach J P. Voice Transformation Using PSOLA Technique [J]. *Speech Communication*, 1992, 11(7): 175 - 187.
- [6] 王聪修. 语音转换及其相关问题的研究[D]. 北京:中国科学院声学研究所,2001.
- [7] 李波,王成友,蔡宣平,等. 语音转换及相关技术综述[J]. *通信学报*, 2004, 25(5): 109 - 118.
- [8] 李波. 语音转换的关键技术研究[D]. 长沙:国防科学技术大学,2005.
- [9] 李元良,李波,王成友. 语音转换中基于系统单位冲激响应的频谱搬移方法[J]. *矿业研究与开发*, 2005, 25(5): 59 - 61.

(编辑:门向生)

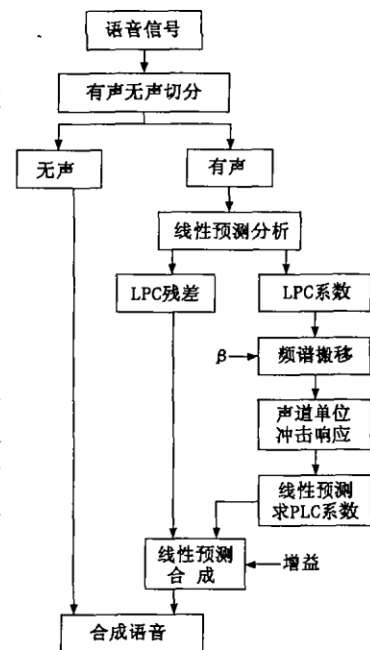


图4 频谱搬移方法流程图

## Spectrum Modification Based on Spectral Envelope Converted LPC Coefficients

ZHANG Hui, LI Bo, WANG Bao-liang

(The Telecommunication Engineering Institute, Air Force Engineering University, Xi'an 710077, Shaanxi, China)

**Abstract:** Spectral envelope conversion is an important content of voice conversion. Based on bilinear transform function, spectral envelope can be efficiently converted even with small training speech database. In this paper, based on LPC analysis and calculating the spectrum converted system unit impulse function, a method for calculating spectrum converted LPC coefficients is proposed, by which the spectral envelope is converted efficiently.

**Key words:** voice conversion; spectral envelope conversion; LPC analysis