

基于离散余弦变换的波形内插语音编码算法

刘靖宇, 鲍长春, 李如玮

(北京工业大学电子信息与控制工程学院语音与音频信号处理研究室, 北京 100124)

摘 要: 针对波形内插(Waveform Interpolation, WI)语音编码的特征波形分解问题, 本文首先提出了基于离散余弦变换(Discrete Cosine Transform, DCT)的特征波形分解方法, 避免了复杂的特征波形对齐运算; 其次, 针对WI的相位重建问题, 提出了清/浊音相位判决和浊音相位分类的方法, 提高了重建语音质量; 最后, 分别构建了速率为2.0kbps和1.6kbps的DCT-WI声码器。主观MOS分表明, 2.0kbps的DCT-WI声码器质量优于2.4kbps MELP声码器, 1.6kbps的DCT-WI声码器亦取得了良好的听觉效果。

关键词: 语音编码; 波形内插; 特征波形; 离散余弦变换; 特征波形对齐; 相位重建

中图分类号: TN912.3 **文献标识码:** A **文章编号:** 0372-2112 (2009) 07-1599-07

Waveform Interpolation Speech Coding Based on DCT

LIU Jing-yu, BAO Chang-chun, LI Ru-wei

(Speech and Audio Signal Processing Laboratory, School of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China)

Abstract: Characteristic waveform (CW) decomposition based on discrete cosine transform (DCT) is proposed firstly in waveform interpolation (WI) speech coding. The complicated CW alignment operation is cancelled by using this method. Secondly, a new unvoiced/voiced phase decision method and the classification method of voiced phase are proposed with respect to phase reconstruction. Finally, DCT-WI coders at 2.0kbps and 1.6kbps are designed, respectively. The subjective MOS test results indicate that the 2.0kbps DCT-WI coder outperforms 2.4kbps MELP coder and the 1.6kbps DCT-WI coder also has a good performance.

Key words: speech coding; waveform interpolation; characteristic waveform; discrete cosine transform; alignment of characteristic waveform; phase reconstruction

1 引言

WI语音编码是低速率语音编码中有代表性的一种方法^[1~4]。该方法基于语音信号的类周期特性, 对去除声道信息的残差域语音信号进行处理, 即根据基音周期, 在每帧残差信号中提取10个特征波形(Characteristic Waveform, CW), 并将每一个CW表示成离散傅立叶级数(Discrete Time Fourier Series, DTFS), 从而实现一维残差信号到二维CW的转换。当以一定的速率提取CW时, 可精确重建残差信号。

为了构建高质量的WI声码器, 必须考虑二维CW的高效量化问题。二维CW维数高, 数据量大, 在量化之前, 要对其进行分解处理。目前CW的分解方法有FIR低通滤波方法、小波变换方法、奇异值分解方法^[5,6]和非负矩阵分解方法^[7~10]。本文针对二维CW分解开展研究, 提出了基于离散余弦变换的二维CW分解方法, 该

方法在算法精确度、计算复杂度和时延等方面都具有明显的优势, 可以应用于WI声码器。本文的第二部分详细介绍了基于离散余弦变换的二维CW分解模型, 第三部分介绍了特征波形对齐及相位重建问题的改进, 第四部分介绍基于离散余弦变换的二维CW量化方法, 第五部分给出了2.0kbps和1.6kbps DCT-WI语音编码方案及性能评估, 第六部分为全文的总结。

2 基于离散余弦变换的二维CW分解方法

离散余弦变换(DCT)是N Ahmed和Rao在1974年提出的正交变换方法, DCT具有与信号较为匹配的独立变换核函数, 性能更接近于理想的K-L变换, 通常被认为是语音、图像进行变换的最佳方法, 目前DCT已广泛应用于图像压缩、视频压缩、音频编码、数字水印、信号检测、人脸识别等领域^[11~14]。

给定信号 $x(n)$, $n = 0, 1, \dots, N-1$, 其离散余弦变

收稿日期: 2008-05-05; 修回日期: 2009-03-02

基金项目: 北京市教委科技发展项目(No. KM200710005001); 国家自然科学基金(No. 60372063); 北京市自然科学基金(No. 4042009)

换为

$$X_c(0) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \quad (1)$$

$$X_c(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \cos \frac{(2n+1)k}{2N}, \quad k = 1, 2, \dots, N-1 \quad (2)$$

其反变换为

$$x(n) = \frac{1}{\sqrt{N}} X_c(0) + \sum_{k=1}^{N-1} \frac{1}{\sqrt{N}} X_c(k) \cos \frac{(2n+1)k}{2N} \quad (3)$$

$$k = 1, 2, \dots, N-1; n = 0, 1, \dots, N-1$$

DCT是一种变换域处理方法,其作用是将时域或空域信号变换到另一个正交矢量空间.DCT又是信号的正交变换方法,使变换域各信号分量之间互不相关,且信号变换前后保持等维,不引入虚部成分^[15].经过DCT正交变换,信号具有明确的频率特性,信号能量更加集中.

本文置语音信号的分析帧长为25ms,采样率为8kHz,即每帧200个样点.一帧语音被均匀分成10个子帧,在每一子帧起始点提取一个CW,每个CW的长度为提取点处的基音周期,CW的提取速率为10个/帧.为了进行分解前后的特征波形对比,图1给出了一帧语音信号分解前的特征波形.其中,横轴代表时间轴,沿其方向代表不同时间点处提取的特征波形.纵轴为特征波形的DTFS展开方向,表达了其幅度谱的一个周期.竖轴表示幅度值.

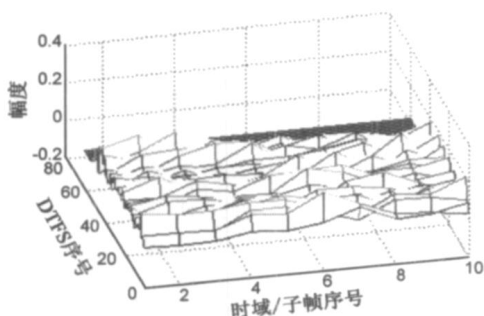


图1 一帧语音的特征波形示例

2.1 特征波形的等维处理

CW展开成DTFS形式后, $K(t)$ 是谐波个数,与基音周期 $P(t)$ 的关系为:

$$K(t_n) = \left\lfloor \frac{P(t_n)}{2} \right\rfloor + 1 \quad (4)$$

一般而言,一帧内的各个CW的DTFS长度不一致.在对特征波形进行分解之前,必须对各个CW的DTFS长度进行等维处理.常用的等维处理方法有截取与补零、带限内插以及多项式拟合等方法^[16],本文采取了补零的方法.

对变维矢量 x 的补零处理是指通过而在的最后一个

元素后添加 $p-K$ 个零值($p-K$)获得,即补零后的矢

量 z 为:

$$z_k = \begin{cases} x_k, & k \leq K \\ 0, & K < k \leq p \end{cases} \quad (5)$$

其中, x_k 和 z_k 分别代表矢量 x 和 z 的第 k 个元素.经过补零后,1帧内10个子帧CW具有相同的维数.

通过补零处理,特征波形的长度达到一致,下一步需要考虑将特征波形的DTFS调整为哪种长度来表示.在WI声码器中,语音信号的基音长度为20~120之间,即CW的DTFS长度为11~61.直观理解是将特征波形的长度补零到最大长度,即每一帧的特征波形构成一个10×61维的幅度谱矩阵.

2.2 特征波形的分类处理

对于语音帧而言,大部分CW的DTFS维数都远小于61,按照上述的固定模式进行等维处理,将CW长度全部调整到61维,会造成计算复杂度高和存储空间消耗大等问题.同时,为了提升量化精度,本文采用特征波形分类方法,按照基音长度对特征波形分类.基音周期是WI声码器中精确传递的信息,具有很高的精确性,且利用基音分类不引入额外的比特开销和计算量.

在WI模型的基音检测中,将不具备周期特性的类噪声语音帧的基音周期规定为80,不具备周期特性的语音段包括清音语音及静默音部分.采用40000帧语音进行统计,基音周期为80的情况所占的比重为40%以上,在分类时,需将该情况作为CW单独一类处理.图2为将基音周期为80的情况剔除的基音周期统计直方图.

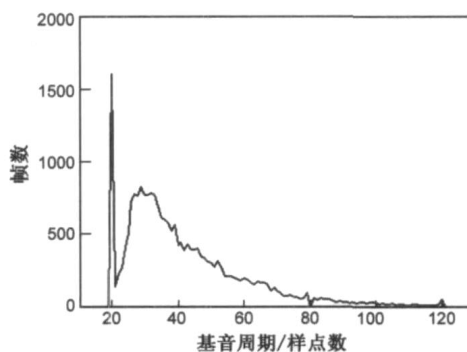


图2 去掉基音为80情况后的基音统计直方图

如何按基音周期对CW进行分类?从分类带来的量化精度提升和存储空间增大两个方面折中考虑,分类情况取4~10之间都是可以接受的,本文将CW分成4类,具体的分类范围如表1所示.

表1 基音分类表

类别	第一类	第二类	第三类	第四类
基音范围	20 ≤ P ≤ 32	32 < P ≤ 60	P = 80	60 < P ≤ 120 且 P ≠ 80

经过分类和补零处理后,不同类CW的频域长度为当前类最大基音周期CW所对应的DTFS长度,即分别为17维,31维,41维和61维.

2.3 CW 的离散余弦变换分解方法

本文采用一维离散余弦变换方法对二维 CW 进行分解,待分解的二维 CW 如图 1 所示。提取一帧内 10 个子帧 CW 的 DTFS 相同次谐波幅度构成 10 维矢量,并对该矢量进行一维离散余弦变换,实现由时域到频域的变换。该 DCT 变换从 DTFS 直流分量开始,至 DTFS 最高次谐波截止。图 3 为二维 CW 经 DCT 处理后频谱。

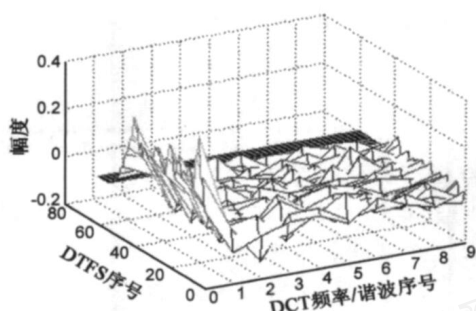


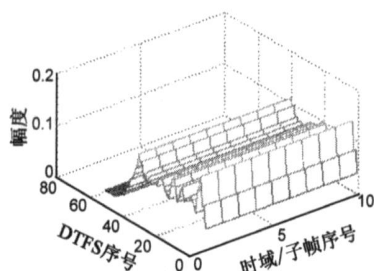
图3 DCT处理后的特征波形

离散余弦变换是一种频域分析方法,经过变换后,原来具有时间意义的 10 维矢量变换后,具有 DCT 域的频率特性,且按照频率由低到高的顺序依次为直流、基频、二次谐波、三次谐波、……、九次谐波。信号的能量更加集中,具有了明确的频率特性,各个频率成分相互正交。对语音信号而言,能量主要集中到了直流和低频部分,高频成分能量较弱。二维 CW 经过 DCT 处理后,

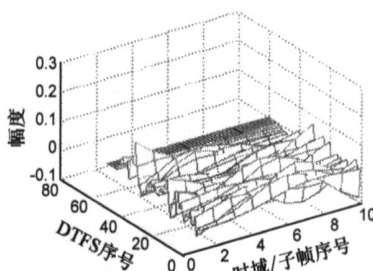
横轴为 DCT 域频率轴,纵轴仍为 DTFS 轴,竖轴为表示幅值。沿 DCT 域频率轴方向,将直流成分构成的矢量定义为直流成分(Direct Component, DC),将 DCT 域的基波、二次谐波和三次谐波构成的矩阵定义为低频成分(Low-Frequency Component, LFC),将 DCT 域的四次谐波、五次谐波、……、九次谐波构成的矩阵定义为高频成分(High-Frequency Component, HFC)。直流成分、低频成分和高频成分彼此独立,具有明确的频率特性。在 WI 声码器中,待量化的二维 CW 即为这三种成分。经过逆离散余弦变换后,可恢复三种成分对应的时域二维 CW,如图 4 所示。将这三部分求和,可精确地重建原始二维 CW。

与传统分解方法相比,离散余弦变换特征波形分解方法处理只需要当前帧语音信息,避免了额外延时,计算复杂度低,分解出的直流、低频、高频成分彼此独立,避免了 FIR 低通滤波方法分解后高、低频不独立的问题。离散余弦变换特征波形分解方法不引入分解偏差,可以精确重建原始信息。

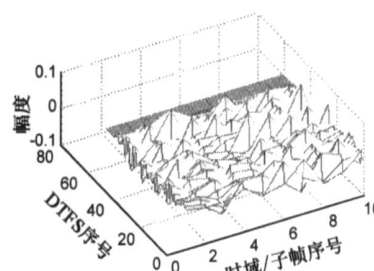
同时,该方法可以将特征波形分解成正交的 10 个频带,按频率升高逐渐精确地描述二维 CW 的细节信息。低频带携带最重要的语音信息,具有核心地位,仅由低频带即可重建良好听觉的语音信号。高频带是重建语音在细节上的补充。因此,基于 DCT 的分解方法具有频域的可升级性,可以应用于嵌入式语音编码。



(a) 直流成分重建的特征波形



(b) 低频成分重建的特征波形



(c) 高频成分重建的特征波形

图4 二维特征波形分解图

2.4 CW 分解方法的性能比较

上面介绍了离散余弦变换特征波形分解方法,现将该方法与传统的 FIR 低通滤波分解方法(LPF)、小波分解方法(Wavelet)、奇异值分解方法(SVD)、非负矩阵分解方法(NMF)进行综合性能比较。采用加权百万次操作每秒测度(wMOPS)^[17]对特征波形分解模块进行计算复杂度统计,如表 2 所示。

表 2 五种分解方法的性能比较

分解方法	额外延时(ms)	是否精确重建	计算复杂度(wMOPS)
FIR 低通滤波	25	是	9.697
小波变换	100	是	-----
奇异值分解	0	是	-----
非负矩阵分解	0	否	18.878
离散余弦变换	0	是	1.869

基于小波变换和奇异值分解的二维 CW 分解方法的计算复杂度高^[7],这里未做统计。由表 2,基于离散余弦变换的特征波形分解方法在算法延时、重建精度和计算复杂度等方面都具有明显的优势。

3 特征波形对齐、相位问题研究

3.1 特征波形对齐问题研究及改进

在传统的 WI 模型中,各个子帧 CW 的主特征(主脉冲)在时间上没有对齐,在二维 CW 分解之前,需要对二维 CW 进行对齐处理。该对齐是在每个子帧 CW 基础上进行的,以前一子帧 CW 作为参考,对当前 CW 进行圆周移位,达到最大互相关时所对应的循环移位 T 为最佳移位。根据 DFT 圆周移位性质,时域圆周移位对幅度谱无影响,仅在相位谱中引入了一个线性相位:

$$= \frac{2T}{P} \quad (6)$$

其中 P 为当前子帧 CW 的基音周期. 在 WI 声码器的编码端, 仅对 CW 的幅度谱分解和量化, 不对相位信息进行量化处理, 只是在解码端用固定相位和随机相位来恢复相位. 因此, 从理论分析的角度, 传统的 WI 声码器编码端的 CW 对齐操作是可以去掉的. 针对特征波形对齐问题, 构造了特征波形对齐和不对齐的两套声码器进行主观 A/B 测试, 测试结果如表 3 所示.

表 3 特征波形对齐问题的主观 A/B 测试结果

测试语音	偏爱对齐	偏爱不对齐	无偏爱
女声语音	21.9 %	15.6 %	62.5 %
男声语音	18.7 %	15.6 %	65.7 %
所有语音	20.3 %	15.6 %	64.1 %

从有无特征波形对齐的两套声码器听觉测试结果看出, 对齐操作对重建语音效果提升的作用很小, 但对齐操作却造成计算复杂度大幅提升. 因此, 建议去除特征波形对齐操作.

3.2 相位问题的研究及改进

现在重写 CW 的 DIFS 展开, 表达式为

$$s(n, m) = \sum_{k=1}^M \left[A_k(n) \cos\left(\frac{2\pi km}{P}\right) + B_k(n) \sin\left(\frac{2\pi km}{P}\right) \right] \quad (7)$$

将其表示成极坐标形式, 可得其幅度谱 $Magn_k(n)$ 和相位谱 $\phi_k(n)$ 分别为

$$\left\{ \begin{aligned} Magn_k(n) &= \sqrt{[A_k(n)]^2 + [B_k(n)]^2} \quad (8) \\ \phi_k(n) &= \begin{cases} \arctan \frac{B_k(n)}{A_k(n)}; & A_k(n) > 0, B_k(n) \geq 0 \text{ 或 } A_k(n) < 0, B_k(n) < 0 \\ \arctan \frac{B_k(n)}{A_k(n)} - \frac{\pi}{2}; & A_k(n) < 0, B_k(n) < 0 \\ \arctan \frac{B_k(n)}{A_k(n)} + \frac{\pi}{2}; & A_k(n) < 0, B_k(n) \geq 0 \\ -\frac{\pi}{2}; & A_k(n) = 0, B_k(n) > 0 \\ \frac{\pi}{2}; & A_k(n) = 0, B_k(n) < 0 \end{cases} \end{aligned} \right. \quad (9)$$

传统的 WI 不对相位谱进行量化, 也有一些学者提出了对相位进行量化, 采用了建模、预测^[18,19]等方法. 本文考虑相位信息的感知次要性, 不对相位进行量化. 与传统 WI 声码器不同, 本文提出将语音进行清/浊音判决处理, 判决函数为基音周期, 清音帧和浊音帧添加不同的相位信息.

3.2.1 浊音帧相位

通过清/浊判决判定当前帧为浊音时, 特征波形幅

度谱的直流成分、低频成分和高频成分分别加入不同的相位.

特征波形的直流成分和低频成分代表着语音信号的频率较低的部分, 对应于语音的类周期部分, 周期特性明显, 加入固定相位重建低频语音的实部 $A_{k_low}(n)$ 和虚部 $B_{k_low}(n)$, 即

$$\begin{aligned} A_{k_low}(n) &= Magn_{k_low}(n) * \cos fixed_{\phi_k}(n) \\ B_{k_low}(n) &= Magn_{k_low}(n) * \sin fixed_{\phi_k}(n) \end{aligned} \quad (10)$$

其中, $fixed_{\phi_k}(n)$ 表示固定相位. 为了克服传统固定相位添加模式单一问题, 本文提出固定相位分类方法, 按照当前的基音周期长度对固定相位分类, 具体的分类如表 4 所示, 每一类具有不同的固定相位特性, 更加适用于相应基音周期的语音重建.

表 4 固定相位的基音分类表

类别	第一类	第二类	第三类	第四类
基音周期 P	20 $P < 40$	40 $P < 60$	60 $P < 80$	80 $P < 120$

特征波形的高频部分对应于语音信号中变化较快的部分, 具有类噪声特性, 周期特性不明显, 因此, 在重建实部 $A_{k_high}(n)$ 和虚部 $B_{k_high}(n)$ 时, 采用加入随机相位的方式, 即

$$\begin{aligned} A_{k_high}(n) &= Magn_{k_high}(n) * \cos random_{\phi_k}(n) \\ B_{k_high}(n) &= Magn_{k_high}(n) * \sin random_{\phi_k}(n) \end{aligned} \quad (11)$$

重建二维 CW 的实部 $A_k(n)$ 和虚部 $B_k(n)$ 为:

$$\begin{aligned} A_k(n) &= A_{k_low}(n) + A_{k_high}(n) \\ B_k(n) &= B_{k_low}(n) + B_{k_high}(n) \end{aligned} \quad (12)$$

3.2.2 清音帧相位

清/浊判决判定为清音帧时, 信号的周期特性不明显, 在加入相位时, 与浊音帧语音不同, 直流、低频和高频成分都采用随机相位重建二维 CW 的实部 $A_k(n)$ 和虚部 $B_k(n)$.

采用清/浊音分类技术, 明显降低了重建语音清音部分的周期性, 听觉上机械性得到了明显抑制, 重建语音听起来更加清晰自然, 取得了很好的改进效果.

4 特征波形的量化

CW 经过离散余弦变换分解后, 被分解成为直流成分、低频成分和高频成分, 如图 5 所示.

各种成分的数据大小也是不同的, 按照 CW 的 DIFS 为 61 维的情况考虑, 直流成分、低频成分和高频成分的空间维数分别为 1 × 61 维、3 × 61 维和 6 × 61 维. 根据人的发声特性, 各个成分对应的语音信息也是不同的. 为了节省码书存储开销, 需要对三种成分进行压

表 7 MOS 分测试结果

算法	2.4kbps MELP	2.0kbps DCT-WI	1.6kbps DCT-W
MOS 分	3.23	3.24	3.01

DCT-WI 声码器的重建语音具有很好的清晰度和自然度,高频噪声得到了明显抑制.从 MOS 分的测试结果可以看出,2.0kbps DCT-WI 声码器取得了略高于 2.4kbps MELP 的听觉效果.1.6kbps DCT-WI 声码器是一个简化版本,比特数较低,亦取得了良好的听觉效果.

6 结论

本文提出了基于离散余弦变换的二维 CW 分解方法,将当前帧的特征波形分解为直流成分、低频成分和高频成分.与 FIR 低通滤波分解方法相比,该分解方法避免了传统方法存在的一帧额外延时,克服了分解后成分的不独立问题,保持了不引入分解偏差的优点,计算复杂度大幅降低.在特征波形对齐问题上,去除了传统 WI 声码器的对齐操作,去掉对齐操作不影响重建语音质量,且使得编解码计算复杂度大幅下降,降低了硬件开销.在相位问题上,本文提出清/浊音相位判决方法和浊音相位分类方法.实验表明,该方法能明显提升重建语音的听觉效果.最后,提出了速率分别为 2.0kbps 和 1.6kbps 的 DCT-WI 声码器,并分别对其进行了 MOS 分测试,主观听觉表明,2.0kbps DCT-WI 声码器取得了略高于 2.4kbps MELP 的听觉效果,1.6kbps DCT-WI 声码器亦取得了良好的听觉效果.

参考文献:

- [1] 鲍长春. 数字语音编码原理[M]. 西安:西安电子科技大学出版社,2007. 220 - 262.
- [2] 李靓,鲍长春,王贵平. 波形内插语音编码技术回顾与展望[J]. 通信学报,2005,26(1):95 - 103.
LI Jing,BAO Chang-chun,WANG Gui-ping. Review of waveform interpolation speech coding[J]. Journal on Communications,2005,26(1):95 - 103. (in Chinese)
- [3] W B Kleijn,J Haagen. Transformation and decomposition of the speech signal for coding[J]. IEEE Signal Processing Letters,1994,1(9):136 - 139.
- [4] W B Kleijn,J Haagen. A speech coder based on decomposition and characteristic waveform[A]. Proceedings of IEEE International Conference on Acoustics,Speech,Signal Processing[C]. Detroit:IEEE Press,1995. 508 - 511.
- [5] 王贵平,鲍长春,张鹏. 基于奇异值分解的低速率波形内插语音编码算法[J]. 电子学报,2006,34(1):135 - 140.
WANG Gui-ping,BAO Chang-chun,ZHANG Peng. Low bit rates waveform interpolation speech coding based on singular value decomposition[J]. Acta Electronica Sinica,2006,34(1):135 - 140. (in Chinese)
- [6] 王贵平,鲍长春. 低速率 WI 语音编码器中 LP 残差信号的 SVD 分解及量化[J]. 信号处理,2005,21(4A):156 - 159.
WANG Gui-ping,BAO Chang-chun. The decomposition and quantization for LP residual signal in low bit rates WI speech coding based on singular value decomposition[J]. Signal Processing,2005,21(4A):156 - 159. (in Chinese)
- [7] 刘靖宇,鲍长春. 基于 NMF 的语音特征波形矩阵量化方法[J]. 信号处理,2007,23(4A):103 - 106.
LIU Jing-yu,BAO Chang-chun. Matrix quantization of characteristic waveform based on NMF[J]. Signal Processing,2007,23(4A):103 - 106. (in Chinese)
- [8] 郭莉莉,鲍长春. 基于非负矩阵分解的语音特征波形分解方法[J]. 信号处理,2007,23(4A):107 - 110.
GUO Li-li,BAO Chang-chun. Decomposition of speech characteristic waveform based on NMF[J]. Signal Processing,2007,23(4A):107 - 110. (in Chinese)
- [9] Peng Zhang,Changchun Bao. A novel 2kbps waveform interpolation speech coder based on non-negative matrix factorization [A]. Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech) [C]. Antwerp,2007. 1661 - 1664.
- [10] 张鹏,鲍长春,郭莉莉. 基于非负矩阵分解的 2kb/s 波形内插语音编码算法[J]. 电子学报,2008,36(4):632 - 638.
ZHANG Peng,BAO Chang-chun,GUO Li-li. 2kbps waveform interpolation speech coding based on non-negative matrix factorization[J]. Acta Electronica Sinica,2008,36(4):632 - 638. (in Chinese)
- [11] 罗亚飞,鲍长春. 基于 DCT 分带谱熵与信号分解的高精度基音检测算法[J]. 电子学报,2007,35(1):13 - 22.
LUO Ya-fei,BAO Chang-chun. Super resolution pitch detection based on band-partitioning spectral entropy and signal decomposition in DCT domain [J]. Acta Electronica Sinica,2007,35(1):13 - 22. (in Chinese)
- [12] 鲁业频,李凤亭等. 离散余弦变换编码的现状与发展研究[J]. 通信学报,2004,25(2):106 - 118.
LU Ye-pin,LI Feng-ting,et al. Research on discrete cosine transformation coding and its present and feature [J]. Journal on Communications,2004,25(2):106 - 118. (in Chinese)
- [13] 陈军,吴成柯. 基于 DCT 变换的内嵌静止图像压缩算法[J]. 电子学报,2002,30(10):1570 - 1573.
CHEN Jun,WU Cheng-ke. Still image compression using embedded DCT approach[J]. Acta Electronica Sinica,2002,30(10):1570 - 1573. (in Chinese)
- [14] Sikora T,Makai B. Shape-adaptive DCT for generic encoding of video [J]. IEEE Transaction on Circuits system,1995,5(2):59 - 62.
- [15] 胡广书. 数字信号处理[M]. 北京:清华大学出版社,2003,363 - 408.

- [16] 朱娜娜,鲍长春,李靓. 一种新的 2kbit/s 波形内插语音编码器[J]. 通信学报, 2004, 25(11): 70 - 76.
ZHU Na-na, BAO Chang-chun, LI Jing. A new waveform interpolation speech coder at 2kbit/s[J]. Journal on Communications, 2004, 25(11): 70 - 76. (in Chinese)
- [17] ITU-T Software Tool Library User's Manual. ITU-T Basic Operators[S]. ITU-T, 2005.
- [18] O Gottesman, A Gersho. Enhanced waveform interpolative coding at 4kbps [A]. Proceedings of IEEE Workshop on Speech Coding Proceedings [C]. Porvoo: IEEE Press, 1999. 90 - 92.
- [19] 陈悦,鲍长春. 一种用于 WI 语音编码的相位预测式矢量量化方法[J]. 电子与信息学报, 2007, 29(11): 2672 - 2675.
CHENG Yue, BAO Chang-chun. A predictive phase vector quantization method in WI speech coding[J]. Journal of Electronic & Information Technology, 2007, 29(11): 2672 - 2675. (in Chinese)
- [20] 王贵平,鲍长春,李靓. 语音特征波形的分解与量化[J]. 电声技术, 2005, (1): 50 - 54.
WANG Gui-ping, BAO Chang-chun, LI Jing. Decomposition and quantization of the characteristic waveform of speech signal[J]. Audio Engineering, 2005, (1): 50 - 54. (in Chinese)

作者简介:



刘靖宇 男, 1982 年 8 月出生于北京, 硕士研究生, 研究方向为语音信号处理、窄带语音编码等.

E-mail: liujingyu @emails. bjut. edu. cn



鲍长春 男, 1965 年 6 月出生于内蒙古赤峰市. 博士, 教授、博士生导师, 国际语音通信学会 (ISCA) 会员, 中国电子学会理事, 信号处理学会委员, 《通信学报》编委会副主任委员, 《信号处理》和《数据采集与处理》编委. 主要研究领域为语音与音频信号处理及编码. (本文通信作者)

E-mail: chchbao @bjut. edu. cn



李如玮 女, 1972 年 1 月出生于四川眉山, 在读博士, 副教授, 主要研究方向: 数字语音信号处理和小波变换.

E-mail: liruwei @bjut. edu. cn