

宽带 ISF 参数的转换分类乘积码锥形矢量量化

李海婷, 鲍长春

(北京工业大学电子信息与控制工程学院, 北京 100022)

摘 要: 本文提出了一种新的应用于宽带导抗谱频率参数量化的转换分类乘积码锥形矢量量化方案. 该量化器基于转换分类与乘积码锥形矢量量化原理, 首先对待量化的 ISF 参数矢量进行分类, 然后按类进行乘积码锥形矢量量化. 该算法具有低存储量及低复杂度的特点. 实验表明, 该算法在每帧编码比特数为 46 时, 平均谱失真比乘积码锥形矢量量化低, 且达到了透明量化标准.

关键词: 语音编码; 矢量量化; 导抗谱频率; 锥形矢量量化

中图分类号: TN912.3 **文献标识码:** A **文章编号:** 0372-2112(2008)02-0362-05

Switched Product Code Pyramid Vector Quantizer for Wideband ISF Parameters

LI Hai-ting, BAO Chang-chun

(School of Electronic Information and Control Engineering Beijing University of Technology, Beijing 100022, China)

Abstract: A new Switched Product Code Pyramid Vector Quantizer (SPCPVQ) used for quantizing the ISF (Immittance Spectral Frequencies) parameters of wideband speech is proposed in this paper. Based on switch vector quantization and product code pyramid vector quantization, in this quantizer, the ISF parameters vector to be quantized is first switched to one of the possible directions and then quantized using the corresponding product code pyramid vector quantizer. The characteristics of this Algorithm are low complexity and low memory for the codebook storage. Experimental results show that this quantization scheme achieves lower spectral distortion with transparent coding at 46bits/frame.

Key words: speech coding; vector quantization; immittance spectral frequencies; pyramid VQ

1 引言

在语音编码领域, 通常用线性预测编码(Linear Predictive Coding, LPC)系数表征声道合成滤波器. 但是由于 LPC 系数的动态范围较大, 出于合成滤波器稳定性和量化效率的考虑, LPC 系数通常被转换成在数学上完全等价的其他参数后再进行量化, ISF(Immittance Spectral Frequencies)参数^[1]就是其中的一种. 与广泛采用的线谱频率(Linear Spectral Frequencies, LSF)参数^[2]不同的是, ISF 参数在保证了精确度的同时降低了计算复杂度, ITU-T G. 722. 2 标准中采用的就是这种转换方式.

随着无线通讯系统中高速率数据服务的发展, 宽带语音(50~7000Hz)编码已经越来越广泛地被采纳, 人们对宽带 LP 系数量化方面的研究也越来越深入. 对于宽带语音通常需要 16 阶线性预测系数才能较好的表示语音的谱包络. 因此, 我们需要用较多的比特数来进行矢量量化, 同时待量化矢量的维数也会变大. 这种情况下采用传统的矢量量化方法, 如分裂式矢量量化^[2]、多级矢量量化^[2]等会造成较大的码书存储量和计算复杂度. Fischer 提出的锥形矢量量化^[3], 以其低计算复杂度和无

需码书存储的特点引起了人们的关注. 这种量化方法恰好可以解决传统方法的缺陷, 但是在中等维数的矢量量化上, 量化精度却不如传统量化方法.

为了更好的解决宽带语音编码中 LPC 系数量化的问题, 本文提出了一种新的应用于宽带语音编码 ISF 参数量化的转换分类乘积码锥形矢量量化(Switched Product Code Pyramid Vector Quantization)方案. 该量化器基于转换分类矢量量化^[4]和 Fischer 提出的乘积码锥形矢量量化^[3]原理, 首先对待量化的 ISF 参数矢量进行分类, 然后按类进行乘积码锥形矢量量化, 在每帧编码比特数为 46 时, 达到了透明量化性能. 虽然计算复杂度和码书存储量比 Fischer 提出的锥形矢量量化器有所增加, 但是量化精度却提高了. 由此看出, 本文提出的量化方案利用转换分类和乘积码锥形矢量量化原理, 在计算复杂度、存储量和量化精度上取了折中, 实现了一种高效的低复杂度、低存储宽带 ISF 参数矢量量化器.

2 转换分类乘积码锥形矢量量化

2.1 锥形矢量量化(PVQ)编码算法

定义位于锥形表面的矢量 $\tilde{c}(n)$ 的集合为:

$$S(L_p, L_p/\lambda) = \left\{ \tilde{c}(n) : \sum_{i=1}^{L_p} |\tilde{c}_i(n)| = L_p/\lambda \right\} \quad (1)$$

其中 $\lambda = 1/E[\tilde{c}(n)]$, L_p 为待量化矢量的阶数. 对于给定的正整数 K , $S(L_p, K)$ 表示该锥形表面上的立方格点的集合. 给定编码速率 R , 如果定义 $S(L_p, K)$ 中所有格点的个数为 $N(L_p, K)$, 则 K 的选择必须满足 $N(L_p, K) \leq 2^R$. 锥形矢量量化编码的步骤如下:

- (1) 在锥形表面寻找与待量化的矢量 $c(n)$ 最近的码字 $\tilde{c}(n) \in S(L_p, L_p/\lambda)$.
- (2) 将矢量 $\tilde{c}(n)$ 乘以 $K/\|\tilde{c}(n)\|_1$ 得到矢量 $\hat{v}(n)$.
- (3) 寻找与矢量 $\hat{v}(n)$ 最近的格点 $\hat{v}(n) \in S(L_p, K)$, 并编码格点索引值.
- (4) 将 $\hat{v}(n)$ 乘以 $\|\tilde{c}(n)\|_1/K$ 得到锥形矢量量化输出矢量 $\hat{c}(n)$.

2.2 乘积码锥形矢量量化(PCPVQ)编码算法

对于中等维数的矢量, 普通的锥形矢量量化在将 $c(n)$ 映射到 $\tilde{c}(n) \in S(L_p, L_p/\lambda)$ 时会产生较大量化误差. 为提高量化的性能, Fischer 又提出了乘积码锥形矢量量化, 其主要思想是用 R_r 比特量化 $r = \|c(n)\|_1$, 以减少映射时产生的误差. 乘积码锥形矢量量化编码的步骤如下:

- (1) 计算并量化 $r = \|c(n)\|_1$, 得到 r 的码书索引值和量化后的值 $\hat{r} = Q(r)$.
- (2) 在锥形表面寻找与待量化的矢量 $c(n)$ 最近的码字 $\tilde{c}(n) \in S(L_p, \hat{r})$.
- (3) 将矢量 $\tilde{c}(n)$ 乘以 K/\hat{r} 得到矢量 $\hat{v}(n)$.
- (4) 寻找与矢量 $\hat{v}(n)$ 最近的格点 $\hat{v}(n) \in S(L_p, K)$, 并编码格点索引值.
- (5) 将 $\hat{v}(n)$ 乘以 \hat{r}/K 得到锥形矢量量化输出矢量 $\hat{c}(n)$.

2.3 转换分类乘积码锥形矢量量化(SCPVQ)编码算法

乘积码锥形矢量量化的量化误差主要产生于三个方面: (1) 量化 r 时产生的误差; (2) 将 $c(n)$ 映射到 $\tilde{c}(n)$ 时产生的误差; (3) 寻找与矢量 $\hat{v}(n)$ 最近的格点 $\hat{v}(n)$ 时产生的误差. 第一个误差的降低在一定程度上受到了编码比特数的限制. 第二、三个误差的降低与待量化矢量的一范数 $\|c(n)\|_1$ 的大小直接相关, 如果 $\|c(n)\|_1$ 大则误差大, $\|c(n)\|_1$ 小则误差小.

针对这样的问题, 本文提出了转换分类乘积码锥形矢量量化. 该方案是在进行乘积码锥形矢量量化之前, 先对待量化的矢量进行转换分类处理, 将待量化的矢量与所选类对应的分类码字相减, 然后根据其在类对应的量化 r 值的码书, 将这个差矢量用乘积码

形矢量量化器进行量化.

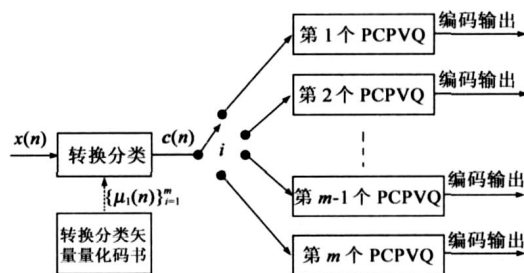


图1 转换分类乘积码锥形矢量量化器的编码示意图

转换分类乘积码锥形矢量量化器的编码原理如图1所示. 首先待量化的矢量 $x(n)$ 根据最近邻准则在转换分类矢量量化码书 $\{\mu_i(n)\}_{i=1}^m$ 中搜索得到最优码字并以此码字 $\mu_i(n)$ 在整个码书中的序号作为分类选择的标志, 得到分类号 $i \in \{1, \dots, m\}$ 和 PCPVQ 的输入矢量 $c(n) = x(n) - \mu_i(n)$; 然后根据分类号 i 的值选择对应的第 i 个 PCPVQ 量化器量化 $c(n)$.

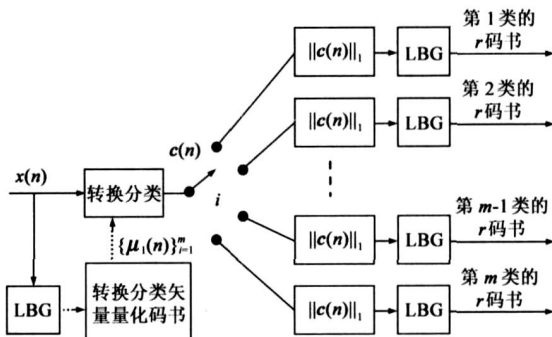


图2 转换分类乘积码锥形矢量量化器的码书训练示意图

2.4 转换分类乘积码锥形矢量量化的码书训练

转换分类乘积码锥形矢量量化器的码书训练数据为包括男女声在内的多说话人汉语数据库, 长度约为2个小时, 采样率为16kHz, 精度为16位线性PCM. 具体训练步骤如图2所示: 首先利用 LBG 算法^[5] 由全部训练数据得到 m 个质心矢量 $\{\mu_i(n)\}_{i=1}^m$, 并以此作为转换分类码书. 然后, 根据分类信息由全部训练数据得到各类的差矢量训练集. 最后, 根据各类的差矢量训练集中各矢量的一范数, 利用 LBG 算法训练各类用于乘积码锥形矢量量化的 r 码书.

3 一种宽带 ISF 参数的转换分类乘积码锥形矢量量化方案

在宽带 ACELP 语音编码器中, 首先, 对预处理后的 12.8kHz 采样的输入语音信号进行线性预测分析. 分析帧长为 30ms, 其中中间的 20ms (256 个样点) 来自当前帧, 前面的 5ms (64 个样点) 来自上一帧, 后面的 5ms (64 个样点) 来自下一帧. 窗函数采用集中于当前帧第四子帧处的非对称窗, 该窗由两部分组成: (1) 半个汉明窗 (256 点); (2) 四分之一余弦函数 (128 点). 其次对加

窗后的语音进行自相关估计,用滞后窗乘以自相关系数使其具有 60Hz 的带宽扩展;(3)采用莱文逊-杜宾递归算法获取 $L_p=16$ 阶 LP 系数;(4)将 LP 系数 $\alpha_i(n)$, $i=0,1,\dots,L_p-1$ 转化为 ISF 参数 $x_i(n)$, $i=0,1,\dots,L_p-1$;(5)应用本文提出的 46bits/frame 转换分类乘积码锥形矢量量化器对 ISF 参数矢量进行量化,选用的格为 $S(L_p,K)$, $L_p=16$ $K=15$. 其中 L_p 的值等于 ISF 参数的阶数, K 值的选取同锥形矢量量化中 K 值选取的基本原则一致. 具体的比特数分配如表 1:

表 1 46bits/frame 转换分类乘积码锥形矢量量化器的比特分配				
	转换分类标志	r 值的量化	格点编码	合计
比特分配 (bits/frame)	8	2	36	46

最后,将得到的转换分类码书索引值、 r 码书量化索引值和格点索引值编码后写入码流.

在宽带 ACELP 语音解码器中,当接收到来自编码器的码流后,根据解码出的转换分类码书索引值 i 得到量化后的转换分类码矢量 $\mu_i(n)$. 根据 i 值选择相应的 PCPVQ 解码器. 根据 r 码书量化索引值和格点索引值,解码得到 \hat{r} 和格点 $\tilde{v}(n)$, 则重构 ISF 参数矢量为 $\hat{x}(n) = \mu_i(n) + \frac{\hat{r}}{K} \tilde{v}(n)$. 将重构的 ISF 参数矢量 $\hat{x}_i(n)$, $i=0,1,\dots,L_p-1$ 转换成 ISP 矢量. 内插 ISP 矢量,得到四个子帧的 ISF 参数 $\hat{x}_{ki}(n)$, $i=0,1,\dots,L_p-1$, $k=0,1,2,3$ 转化为 LP 系数 $\alpha_{ki}(n)$, $i=0,1,\dots,L_p-1$, $k=0,1,2,3$.

4 实验结果与分析

为了验证本文所提方法的有效性,本文进行了 5 个仿真试验,并进行了性能分析,具体如下所述.

在仿真实验 1 中,用训练语音外的 345 秒(17250 帧)汉语语音产生 17250 个 16 阶的 ISF 矢量测试本文提出的 46bits/frame 宽带 ISF 参数的转换分类乘积码锥形矢量量化器的量化性能. 按国际通用的方法计算平均谱失真,实验结果如表 2 所示. 从表 2 中可以看出,本文提出的 46bits/frame 的转换分类乘积码锥形矢量量化方案得到的平均谱失真约为 1dB,谱失真介于 2dB 和 4dB 之间的帧的百分比仅为 1.65% (< 2%),而平均谱失真大于 4dB 的帧的百分比为零,其量化性能达到了透明量化的标准.

表 2 本文提出的 46bits/frame SPCPVQ 量化器的量化性能			
测试项	平均谱失真	> 2dB < 4dB	> 4dB
本文提出的算法	1.009275dB	286 帧(1.65%)	0 帧(0.0%)

在仿真实验 2 中,用同样的测试数据分别测试了本文提出的算法和 G. 722. 2 中 ISF 参数量化算法的存储复杂度

表 3 存储复杂度和计算复杂度的性能对比		
	G. 722. 2	本文提出的算法
所用比特数(比特/帧)	46	46
有无用到预测	有	无
码书存储量(float)	5280	5120
加法运算量(次/帧)	17760	8523
乘法运算量(次/帧)	8848	4290

从表 3 中可以看出,与 G. 722. 2 中 ISF 参数量化方案相比,本文提出的量化方案在 46 比特每帧时同样达到了透明量化的标准,但是其码书存储量比 G. 722. 2 小,计算复杂度约为是 G. 722. 2 的 50%. 而且,本文提出的量化方案没有用前一帧的值进行预测,在丢失帧时可以有更好的恢复出 ISF 参数,避免错误繁殖.

在仿真实验 3 中,对比乘积码锥形矢量量化器和本文提出的转换分类乘积码锥形矢量量化器,用同样的测试数据分别测试其量化性能,实验结果如表 4 所示. 从表 4 中可以看出,在宽带 ISF 参数的量化上,本文提出的转换分类乘积码锥形矢量量化器的量化性能要远远优于相同编码速率下乘积码锥形矢量量化器的量化性能. 这主要是由于本文提出的量化方案运用了转换分类技术,降低了进入 PCPVQ 量化器的矢量的一范数 $\|c(n)\|_1$ 的大小,并通过对于 r 码书的分类训练提高了 r 值量化的精细程度.

表 4 PCPVQ 与 SPCPVQ 的性能对比					
(a) 比特分配列表					
	选用的格 $S(16,K)$	编码格所需比特数 (bits/frame)	量化 r 所需比特数 (bits/frame)	转换分类所需比特数 (bits/frame)	所需比特数总和 (bits/frame)
PCPVQ 方案 1	$K=20$	42	3	0	45
PCPVQ 方案 2	$K=20$	42	4	0	46
PCPVQ 方案 3	$K=20$	42	5	0	47
PCPVQ 方案 4	$K=25$	47	4	0	51
PCPVQ 方案 5	$K=30$	50	4	0	54
本文提出的算法	$K=15$	36	2	8	46

(b) 性能对比				
	码书存储量 (float)	平均谱失真 (dB)	> 2dB < 4dB	> 4dB
PCPVQ 方案 1	8	1.391	20.65%	0.185%
PCPVQ 方案 2	16	1.359	19.55%	0.08%
PCPVQ 方案 3	32	1.350	19.35%	1.19%
PCPVQ 方案 4	16	1.103	7.84%	0%
PCPVQ 方案 5	16	0.926	2.65%	0%
本文提出的算法	5120	1.009	1.65%	0%

在仿真实验 4 中, 改变转换分类乘积码锥形矢量量化器的编码比特数, 用同样的测试数据分别测试其量化性能, 实验结果如表 5 所示.

表 5 不同编码比特数下 SPCPVQ 量化器的性能对比

	SPCPVQ 方案 1	SPCPVQ 方案 2	SPCPVQ 方案 3	SPCPVQ 方案 4
所需比特数的总和 (bits/ frame)	49	50	48	46
选用的格 $S(16, K)$	$K = 20$	$K = 20$	$K = 20$	$K = 18$
编码格所需的比特数 (bits/ frame)	42	42	42	40
量化 r 所需的比特数 (bits/ frame)	4	4	2	2
转换分类所需的比特数 (bits/ frame)	3	4	4	4
平均谱失真 (dB)	0.86198	0.797628	0.908138	1.13841
$> 2\text{dB} \< 4\text{dB}$	1.5%	0.915%	1.63%	4.56%
$> 4\text{dB}$	0%	0%	0%	0%

从表 5 中可以看出, 当量化 r 所需比特数和格的 K 值一定时, 转换分类乘积码锥形矢量量化器随着转换分类个数的增加, 其量化性能相应提高; 当转换分类个数和格的 K 值一定时, 随着量化 r 所需比特数的增加, 量化器的量化性能相应的提高; 当转换分类个数和量化 r 所需比特数一定时, 随着选取格的 K 值的增加, 量化器的量化性能相应提高. 因此, 在编码比特数一定的情况下, 只有合理地选择转换分类个数、量化 r 所需的比特数和格的 K 值, 才能得到性能优越的转换分类乘积码锥形矢量量化器.

表 6 不同参数设置和比特分配的 46bits/ frame SPCPVQ 量化器的性能对比 (a) 比特分配列表

	选用的格 $S(16, K)$	编码格所需比特数 (bits/ frame)	量化 r 所需比特数 (bits/ frame)	转换分类所需比特数 (bits/ frame)	所需比特数总和 (bits/ frame)
PCPVQ 方案 1	$K = 18$	40	2	4	
PCPVQ 方案 2	$K = 17$	39	3	4	
PCPVQ 方案 3	$K = 17$	39	2	5	
PCPVQ 方案 4	$K = 16$	38	4	4	
PCPVQ 方案 5	$K = 16$	38	3	5	
PCPVQ 方案 6	$K = 16$	38	2	6	
PCPVQ 方案 7	$K = 15$	36	3	7	
本文提出的算法	$K = 15$	36	2	8	

(b) 性能对比

	码书存储量 (float)	平均谱失真 (dB)	$> 2\text{dB} \< 4\text{dB}$	$> 4\text{dB}$
PCPVQ 方案 1	320	1.13841	4.56%	0%
PCPVQ 方案 2	384	1.11921	5.42%	0%
PCPVQ 方案 3	640	1.0853	3.38%	0%
PCPVQ 方案 4	512	1.16622	7.38%	0%
PCPVQ 方案 5	768	1.08759	4.33%	0%
PCPVQ 方案 6	1280	1.07673	2.97%	0%
PCPVQ 方案 7	3072	1.02336	2.78%	0%
本文提出的算法	5120	1.009	1.65%	0%

在仿真实验 5 中, 固定编码比特数为 46bits/ frame, 改变转换分类乘积码锥形矢量量化器中的参数设置和比特分配, 用同样的测试数据分别测试其量化性能, 实验结果如表 6 所示.

从表 6 中可以看出, 当转换分类乘积码锥形矢量量化器的编码比特数固定在 46bit/ frame 时, 只有转换分类个数为 256 类(用 8bits 进行编码), 选取的格为 $S(16, 15)$ (用 36bits 编码) 并用 2bits 量化 r 值的方案, 即本文提出的方案, 满足透明量化的全部标准.

5 结论

本文提出了一种应用于宽带语音编码 ISF 参数量化的转换分类乘积码锥形矢量量化 (Switched Product Code Pyramid Vector Quantization) 方案. 该量化方案基于转换分类矢量量化和 Fischer 提出的乘积码锥形矢量量化原理, 在计算复杂度、存储量和量化精度上取得了平衡, 实现了一种高效的低复杂度、低存储宽带 ISF 参数矢量量化器. 该量化方案在每帧编码编码比特数为 46bits 时, 达到了透明量化的效果. 虽然计算复杂度和码书存储量比 Fischer 提出的锥形矢量量化器有所增加, 但是量化精度却提高很多.

参考文献:

[1] Y Bistriz and S Peller. Immittance spectral pairs (ISP) for speech encoding [A]. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP' 93) [C]. Minneapolis, Minnesota, 1993, Vol. 2, 9- 12.

[2] 鲍长春. 低比特率数字语音编码基础 [M]. 北京: 工业大学出版社, 2001. 96- 148.

[3] Fischer T R. A pyramid vector quantizer [J]. IEEE Trans Inf Theory, 1986, 32(4): 568- 583.

[4] Stephen So, Kuldip K Paliwal. Switched Split Vector Quantisation of Line Spectral Frequencies for Wideband Speech Coding [C]. Interspeech2005, Lisbon, 2005. 2705- 2708.

[5] 孙圣和, 陆哲明. 矢量量化技术及应用[M]. 北京: 科学出

版社, 2002. 221-231.

作者简介:

李海婷 女, 1982年6月出生于北京, 硕士研究生. 研究方向语音信号处理、宽带变速率语音编码等. E-mail: lihaiting@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》《信号处理》和《数据采集与处理》学报编委. 主要研究领域为语音与音频信号处理及编码. E-mail: chcbao@bjut.edu.cn

