

## 基于 G.729A 的 CS-ACELP 语音编码算法的优化改进\*

宋钦梅, 黄冰, 杨召青

(桂林电子科技大学 通信与信息工程系, 广西 桂林 541004)

**摘要:** 介绍了 G.729A 的现状以及在发展过程中其语音编码的改进方向及具体方法。提出了一种结合 WD-LSP (Weighted Delta-LSP) 并采用次最优部分码本快速搜索的 CS-ACELP 语音编码算法, 同时采用基于声学心理模型的知觉加权滤波器, 使语音编码在不降低语音质量的情况下降低计算复杂度。

**关键词:** G.729A WD-LSP CS-ACELP 部分码本快速搜索 知觉加权滤波器

### 1 G.729A 现状

1996 年 3 月, ITU-T 制定了使用共轭代数码激励线性预测 (CS-ACELP) 的 8Kbps 语音编码标准 G.729, 并于同年 11 月制定了 G.729A<sup>[1]</sup>。G.729A 是 G.729 的简化版本, 改进后的算法复杂度降低很多, 而语音质量下降却很少, 且两者在编解码器的比特流一级完全兼容, 可以互通。因此, G.729A 被订为 ITU-T V.70 系列 DSVD 的语音编码标准。

由于 G.729A 和 G.729 的比特流完全兼容, 所以在降低复杂度应用中, G.729A 可以用来代替 G.729。G.729A (在 DSVD 中) 的应用主要有: 多方多媒体电话会议、协同计算、电话教学、交互式游戏、说话的同时传输文件、移动视听服务、远程通信、电话购物等。但是对于常用的快速 DSP 芯片的实时实现, G.729A 的复杂性仍然较高, 有待进一步地改进。

### 2 语音编码改进中的关键技术

语音编码 (或称语音压缩编码) 研究的主要问题有: (1) 在给定编码速率的条件下, 如何能得到尽量好的重

建语音质量(或称编码质量)。(2)在尽量减小编解码时延的同时也应尽量减小算法的复杂度。因此对整个 G.729A 声码器的改进也即是针对上述两个方面进行的。

### 3 改进方向及方法

#### 3.1 提高重建语音质量

ITU-T G.729A 对于纯语音信号来说,译码恢复的语音信号质量较好,但是对有背景声的语音信号或者只有噪音或乐音的信号,处理效果则不尽人意。在下面的一些改进方案中,虽然有的不是明确针对 G.729A 的,但其思想可以吸收进 G.729A 的改进方案中。

##### (1)广义基音预测的 CELP

对于噪声信号的优化处理,Paul Mermelstein 和 Yasheng Qian 提出了一种广义基音预测(Generalized Pitch Prediction)的 CELP<sup>[2]</sup>,与典型的 CELP 的不同在于:噪声部分是语音信号被分解成可用自适应码本预测的预测分量和不能预测的非周期分量,而不是 LPC 残差信号。该方案提高了基音预测增益和信噪比。

##### (2)自适应变换编码 ATC 和 CELP 相结合的算法 ATCELP

PierreCombescore Ürgen Schnitzler 等人提出的一种自适应变换编码 ATC 和 CELP 相结合的算法 ATCELP<sup>[3]</sup>,对于语音信号,主要运用 CELP 算法;而对于乐音和噪声信号,则采用 ATC 模式。该方案的实验结果显示 16、24、32Kbps 的 ATCELP 的性能接近 48、56、64Kbps 的 ITU-T G.722。

##### (3)波形匹配和能量匹配的综合

Roar Hagen 和 Erik Ekudden 针对应用在许多现存的声码器中的典型波形匹配 LPAS 结构对背景噪声处理欠佳的不足,提出了一种对语音信号和非语音信号处理都有较佳表现的、综合了波形匹配和能量匹配的 8Kbps ACELP 的声码器<sup>[4]</sup>。

##### (4)采用更好的帧检错纠错及丢失补偿处理方法

提高重建语音信号质量的另一有效途径是采用更好的帧检错纠错及丢失补偿的处理方法。Juan Carlos De Martin TakahiroUnno 和 VishuViswanathan 提出了两种方案<sup>[5]</sup>,一种是插值方法,它使用前一帧的信息和后一帧的信息来重建当前出错的帧;另一种是基于重复的方案,该方案不需要后一帧的信息,而是在语音特征的提取方面作进一步的改善,如进一步优化基音特征的检测。

#### 3.2 降低算法的复杂度及减小编解码时延

对于采用码本结构的算法,减小算法时延一般集中在改进码本搜索上,而改进的焦点一般集中在减少搜索空间。杨树堂、周敬利、余胜生等人采用聚类优化策略实现了对码本的快速搜索<sup>[6]</sup>。Miguel Arjona Ramirez 和 Gerken 提出了幅度和位置的联合搜索方案来搜索代数多脉冲码本<sup>[7]</sup>。为了减少计算复杂性,参考文献[8]采用自适应码本和随机码本两种搜索方法。在进行码本搜索时,使用端点矫正算法进行快速卷积和能量计算。基音延迟的搜索采用先整数后分数两层次搜索方法,在使用内插计

算非整数延迟所对应的码字时,使用 8 个点的内插进行搜索,为减少计算量,合成使用 40 个点的内插;对于每一个偶数子帧延迟,则使用差分搜索。

下面介绍 G.729A 声码器中代数码本的一种快速搜索方法<sup>[9]</sup>:

在有一个脉冲的码本中计算目标矢量和滤波后的自适应码本矢量的均方加权误差,根据该均方误差重排码本搜索顺序。由于在固定码本的搜索过程中,计算量主要集中在  $\Phi(i,j)$  ( $i=0, \dots, 39; j=i, \dots, 39$ ) 的计算上,根据  $\Phi(i,j)$  的计算公式  $\Phi=H^T H$  ( $H$  是下三角形 Toeplitz 卷积矩阵,由加权合成滤波器的冲激响应构造而成, $h(i)$  是卷积矩阵  $H$  中的元素),有

$$\Phi(i,j) = \sum_{t=0}^{39} h[t+(j-i)] \times h[t], \text{ 由 } (t=n-j) \text{ 得到下式:}$$

$$\Phi(i,j) = \Phi(i-1,j-1) - h[40-i]h[40-j]$$

根据该推导公式,在计算  $\Phi(i,j)$  时减少了较大的计算量。

本方案提出的方法由三个阶段的搜索过程构成:第一阶段,在重排序的代数码本中根据门限值从  $p_i$  ( $i=0, \dots, 4$ ) 中选取前  $n$  个脉冲位置;第二阶段,用第一阶段没有检测的  $p_2$  和  $p_3$  或  $p_4$  的脉冲对  $(i_2, i_3)$  检测以及  $p_0$  和  $p_1$  的脉冲对  $(i_0, i_1)$  检测,此过程同样可应用于其他组合;第二阶段分别对  $(i_0, i_1, i_2)$ 、 $(i_0, i_1, i_3)$ 、 $(i_0, i_2, i_3)$ 、 $(i_1, i_2, i_3)$  利用剩下的一个  $p_i$  计算  $\frac{X^T H C_k}{C_k^T H H C_k}$ 。

#### 4 一种结合 WD-LSP<sup>[10]</sup>采用次最优部分码本快速搜索的 CS-ACELP 语音编码算法

G.729A 编码器部分中的功能模块的改进重点在于自适应码本搜索(因为它们占用了声码器大部分的处理时间),包括自适应码本矢量的产生、自适应码本时延的码字编码、自适应码本增益的计算、固定码本搜索(包括固定码本矢量的产生、固定码本的码字编码、固定码本增益的量化),同时与自适应码本搜索相关的开环基音分析也有一定的改进余地。本文提出了一种新的降低算法复杂度的方法,即结合 WD-LSP 采用次最优部分码本快速搜索的 CS-ACELP 语音编码算法<sup>[11]</sup>,其原理图如图 1 所示。

(1)在 G.729A 语音编码阶段,为了减少搜索最佳自适应码本延迟必需的计算量,开环基音延迟  $T_{op}$  在每帧(10ms)都做一次分析,因而需要较多的 CPU 时间用于开环基音分析。为此可采用 WD-LSP 函数来减少用于开环基音分析的计算量。WD-LSP 函数主要用于区分 UV-V (UnVoice-Voice)/S-V (Silence-Voice) 的边界。其原理是:如果函数值大于给定的极限值  $\eta$ , 则开环基音延迟  $T_{op}$  重新估计;否则,开环基音延迟  $T_{op}$  用前一帧的自适应码本延迟来更新。在第  $i$  帧的 WD-LSP 函数  $F_i$  和用于确定开环基音延迟  $T_{op}$  的算法描述如下:

$$F_i = \sum_{k=0}^{10} w_k \times (LSP_i(k) - LSP_{i-1}(k))^2$$

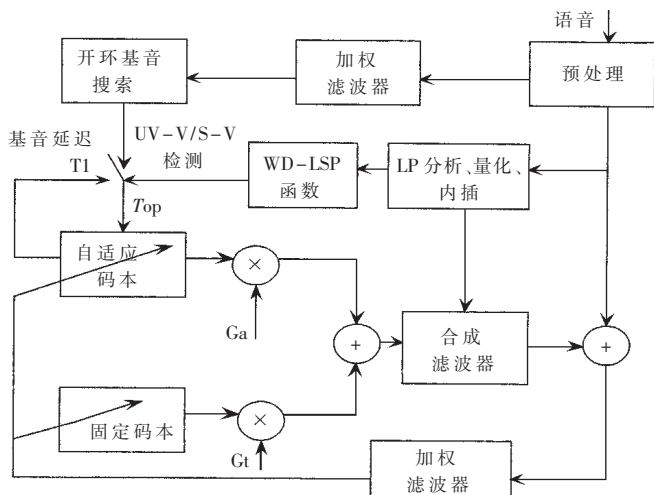


图1 改进的编码原理图

如果  $F_i > \eta$ , 重新估计  $T_{op}$ ; 否则  $T_{op}$  = 自适应码本延迟  
 其中,  $LSP_i(k)$  是在第  $i$  帧中的  $k$  阶 LSP 系数;  $w_i$  是加权系数, 用于增强 UV-V/S-V 边界的 WD-LSP 函数。为了获取  $w_i$ , 一个包含 23 014 个 UV-V 边界和 9 519 个 S-V 边界的大型数据库用于估计 delta-LSP 在 UV-V/S-V 边界的平方根值(RMS)。因此, WD-LSP 用于检测 VU-V/S-V 边界非常敏感;  $\eta$  是一个设为 0.01 的极限值。整个计算可节省 21% 的计算量, 而合成语音质量仍然较好。

(2) 本语音编码方法的码本搜索采用次最优部分码本快速搜索法并采用基于声学心理模型的知觉加权滤波器。次最优搜索方法是参考次最优顺序优化方法提出的, 因为不是全码本搜索, 在最坏的情况下, 搜索到的不是最优的激励矢量, 它把四层循环分成三部分, 第一部分和第二部分分别为 1 个单层循环, 分别对应第 1 和第 2 个脉冲的位置和符号; 第三部分为一个包含 2 个循环的多层循环, 这个多层循环的 2 个循环分别确定后 2 个对应脉冲的位置和符号。部分搜索方法只在一个低的码书百分比范围。该方法利用了准最优顺序优化的思想, 但不完全是对单个脉冲顺序优化, 而采用了一次进行 2 个脉冲的位置幅度搜索, 对另外 2 个脉冲则仍然进行单个脉冲定位和定幅。实验表明, 在次最优部分码本搜索法中, 即使仅仅是一小部分的码本被搜索, 也会在一定程度上降低重建语音的质量, 不过其再生语音质量的降低几乎不能被听觉感觉出来。

(3) 在该算法中, 为不降低语音质量, 考虑到人耳对不同频率信号的不同敏感度, 可采用一种基于声学心理模型的知觉加权滤波器, 其设计原理利用了听觉系统生理结构和人耳主观感知特性的数学抽象模型<sup>[12]</sup>。实验表明, 这种滤波器能使重建语音的 MOS 分改善 0.1~0.3。

(4) 在解码端的后处理中, 解码端采用 Julin-Hwey Chen 建议的实用自适应滤波算法, 该算法主要用于增强编码语音的主观质量, 算法的后滤波器由长时滤波器和短时滤波器级联组成, 包括谱的倾斜补偿和自动增益控

制。长时滤波器用来衰减基音谐波之间的谱的谷值<sup>[13]</sup>; 短时滤波器则用来衰减共振峰之间的谱的谷值。长时和短时滤波器都是零极点滤波器。谱倾斜补偿可以减小合成语音的削波现象。

根据当前国内外有关学者对于语音编码算法的研究, 本文在介绍 G.729A 现状的基础上, 对 G.729A 语音编码的改进指明了思路: 提高重建语音的音质, 降低算法复杂度和减少算法延迟。并给出了一种新的降低语音编码复杂度的方法, 这种方法对今后 G.729A 语音编码的改进具有一定的意义。

参考文献

- [1] ITU-T Recommendation G.729 Annex A[S]. International Telecommunication Union.1996.
- [2] MERMELSTEIN P, QIAN Y. Analysis by synthesis speech coding with generalized pitch prediction[D]. IEEE International Conference on Acoustics, Speech and Signal Processing(Volume 1), 1999,1-4.
- [3] COMBESCURE P, SCHNITZLER J. A 16, 24, 32 kbit/s wideband speech codec based on ATCELP[C]. IEEE International Conference on Acoustics, Speech and Signal Processing(Volume 1), 1999:5-8.
- [4] HAGEN R, EKUDDEN E. An 8 kbit/s ACELP coder with improved background noise performance[C]. IEEE International Conference on Acoustics, Speech and Signal Processing(Volume 1), 1999:25-28.
- [5] MARTIN J C D, UNNO T, VISWANATHAN V. Improved frame erasure concealment for CELP-based Coders[J]. Proceedings of IEEE ICASSP, In: Istanbul Turkey, 2000, (3). 1483-1486.
- [6] 杨树堂, 周敬利, 余胜生. G.723.1 语音编码器算法的聚类优化策略及其应用[J]. 通信学报, 2001, (2): 113-117.
- [7] RAMIREZ M A, Gerken M. A multistage search of algebraic CELP Codebooks[D]. IEEE International Conference on Acoustics, Speech and Signal Processing(Volume 1), 1999: 17~20.
- [8] Liyi. CELP 语音压缩系统的实现. <http://www.cs.washington.edu/homes/yi/projects/CELP.ppt>. 2004.
- [9] 徐正伟. ITU-T G.729A 协议的研究及其实现[D]. 中国优秀博硕士学位论文全文数据库, 2002:69-72.
- [10] HWANG S H. Computational improvement for G.729 standard[J]. ELECTRONIC LETTERS. 22nd June, 2000:1163-1164.
- [11] 杨国芳, 高飞. 一种基于 G[1].729 的 CS-ACELP 新算法[J]. 语音技术, 2004, (1):56-59.
- [12] 李文广, 陈健. 低码率语音编码器中感知加权滤波器的一种新型算法[J]. 上海交通大学学报, 1998, 32(6):38-43.
- [13] 鲍长春. 低比特率数字语音编码基础[M]. 北京: 北京工业大学出版社, 2001.

(收稿日期: 2006-10-04)