

宽带语音编码算法及实现*

Algorithms and Their Implement on Wideband Speech Coding

秦德兴, 覃团发, 陈琳

Qin Dexing, Qin Tuanfa, Chen Lin

(广西大学计算机与电子信息学院, 广西南宁 530004)

(Coll. of Comp. & Elec. Info., Guangxi Univ., Nanning, Guangxi, 530004, China)

摘要:介绍基于 G. 722 标准的嵌入式 CELP/ 矢量和激励线性预测 (VSELP) 编解码、子带码激励线性预测 (SB-CELP)、多带码激励线性预测 (MB-CELP)、G. 729、自适应变换编码 (ATC)、变换码激励 (TCX) 以及变换预测编码 (TPC) 等算法, 给出基于 C 语言和 MATLAB 语言的 10.4kbit/s 和 14.4kbit/s 两种不同比特率的宽带语音编码的实现方法。仿真结果表明, 主观意见评分 (MOS) 可达 3.1~3.6, 语音质量达到预定要求。

关键词: 宽带语音编码 码激励线性预测 变换码激励 算法

中图分类号: TN912.3

Abstract: Several algorithms about wideband speech coding are introduced. Base on G. 722 standard, different methods of CELP are discussed in detail, such as embedded CELP/VSELP codec, SB-CELP, MB-CELP and the algorithm of G. 729 standard. These algorithms can be used to conduct speech coding at very low bit rate. Based on SBC method, the implementation of wide speech coding by C and MATALAB simulation is obtained at the bit rates of 10.4kbit/s and 14.4kbit/s. The result of 3.1~3.6 was obtained in the hearing test.

Key words: wideband speech coding, CELP, TCX, algorithm

语音编码是实现语音数字化传输和存储的基础。数字语音在可靠性、抗干扰和保密性等方面优于模拟语音,近十年来语音编码技术得到了很大的发展。国际电联 (ITU-T) 制定出一系列具有代表性的低码率语音编码,如 ITU-G. 723, G. 729 等^[1,2]。语音压缩的对象主要是电话质量的语音,典型的带宽为 3.2kHz (200Hz~3.4 kHz),采用 8kHz 的采样率。但近几年来现代通信正在由窄带通信向宽带通信发展,语音带宽扩展为 7kHz (50Hz~7.5kHz),通常称为宽带语音^[3]。频带的扩展使得宽带语音在通信的透明度,语音质量,真实感,语音识别等方面均有很大的改善,因而宽带语音将代替窄带语音运用于视频会议,电话会议以及 ISDN 的多用户/多信道通信等领域。本文介绍了当前实现宽带语音编码所采用的几种算法,并基于子带编码的思想,给出了 C 语言和 MATLAB 语言仿真实现 10.4kbit/s 和

14.4kbit/s 的两种不同比特率的宽带语音编码的实现方法。

1 宽带语音编码

1.1 波形编码

宽带语音编码是信源编码的主要研究课题之一,它融合时域和频域的编码技术,从而得到高质量低速率的编码^[4]。1972 年 ITU 提出 64kbit/s 的 G. 711 PCM^[5]。由于语音信号中有大量的冗余,如果在量化和编码前去掉这些多余的信息,编码的比特数就可以减少,最简单的方法就是采用 DPCM。由于 DPCM 方法获得的预测增益是有限的,而且量化器同样不能适应语音的短时平稳性,由此造成的量化误差不能减小到足够小。在实际应用中,采用自适应差分脉冲编码 ADPCM。该算法针对语音信号的非平稳特点使用了自适应预测和自适应量化,从而使波形数字编码能够自适应地与语音信号幅度的变化相匹配,能够适应语音信号本身及其传输过程中的不平稳性。CCITT 于 1984 年公布了 32kbit/s 的 ADPCM 编码器标准 G. 721^[6],不久又公布了 G.

2004-08-28 收稿。

* 广西科学研究与技术开发计划 (桂科攻 0330010-1) 和广西“新世纪十百千人才工程”专项资金 (2002222) 联合资助项目。

723 标准^[1],将编码速率扩充到 40kbit/s 和 24kbit/s;在 G. 726 标准中进一步将速率扩充到 16kbit/s,从而形成了 16kbit/s 至 40kbit/s 的一套完整速率的 ADPCM 算法。以上的标准是基于电话质量的语音,已经不能满足目前视频会议质量的要求。于是,CCITT 制定了宽带语音压缩标准 G. 722。G. 722 使用的编码技术为子带 ADPCM 编码^[3],其编解码框图如图 1 所示。

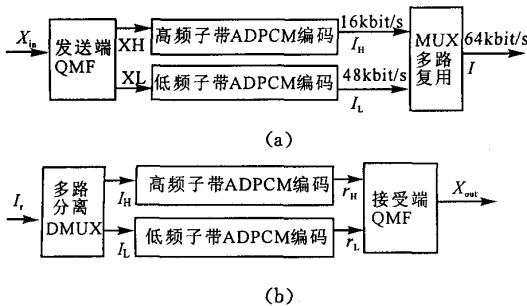


图 1 SB-ADPCM 编解码框图
(a) 编码框图; (b) 解码框图

G. 722 的音频信号采样频率为 16kHz, 音频信号被正交镜像滤波器(QMF)分割成高频(4kHz~7kHz)和低频(0~4 kHz)两个子带, 每个子带都采用 ADPCM 编码。最后原始语音被压缩成 64kbit/s、56kbit/s、48kbit/s 三种传输速率的码流, 以供不同的信道使用。随着多媒体技术的发展, 尤其是活动图像编码 MPEG 系列的发展, G. 722 标准的码流速率显得太高, 线路利用率不经济。此外, 窄带语音编码技术发展比较完善, 其中有许多高质量的算法如 LPAS、CELP、ACELP、VSELP、MBE 以及 MELP 等陆续被一些地区组织定为标准, 这为研究新的高效率的低比特率宽带语音编码算法奠定基础, 从而可将编码速率降至 24kbit/s, 16kbit/s, 甚至更低。

1.2 宽带中的 CELP 编码

CELP 算法在低比特率的窄带语音编码的应用中已经取得了很大成果, 标准的 CELP 算法为固定比特流, 而在实际应用如视频会议, 移动通信等, 都需要变速率。变速率语音编码的原理图如图 2 所示。

1.2.1 嵌入式 CELP/VSELP 编码算法^[7]

嵌入式的语音编码如图 2(b)所示。该编码算法的语音编码的码率范围是 16kbit/s~32kbit/s。一种改进的嵌入式 CELP 方法^[7]是使用最佳的正交化增益来获得最佳的嵌入式编码器, 并采用嵌入式 CELP/VSELP 编解码, 它的 CELP 码本采用了 1 个自适应码本和 2 个固定码本。第一个新的目标量化的计算来自加权振铃, 加权振铃是在 32bit/s 的比特

流中, 遵守基音计算以及从低比特码流速率的重建余项中搜索第一个波形时获得。其他的波形索引与增益是对该目标信号使用 CELP 差错准则来获得。经过测试初步的结果表明嵌入式 32kbit/s 与 24kbit/s 的 CELP/VSELP 编码可以与 G. 722 的 64kbit/s, 56kbit/s 相当。嵌入式的 VSELP 编码可以为宽带语音编码提供变比特率的编码, 并将比特率降低到 32kbit/s 与 24kbit/s。

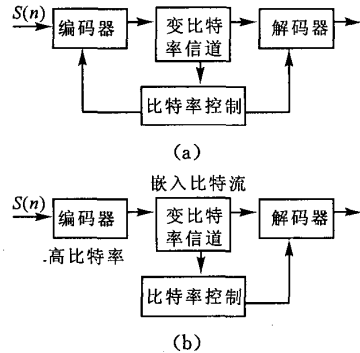


图 2 变速率语音编码原理

(a) 变比特率语音编码的原理及编解码同步工作框图; (b) 在传输时嵌入式语音编码比特流

1.2.2 SB-CELP 编码算法^[8]

G. 722 标准是最早提出的宽带语音编码标准, 并可以给新的宽带语音编码算法研究提供参考。基于 G. 722 的编码思想, 对于低比特率的宽带语音编码可以采用 2 个子带, 分割频率为 6 kHz。其中, 高频子带部分(6kHz~7kHz)在语音信号中所占比例较小, 而引入噪声占总噪声的 60%, 因而可用简单的模型和编码来替换, 产生低比特率编码数据, 或不传输任何信息而在解码时重新合成高频子带部分的语音。这样整个比特流的数据主要集中在低频子带部分(0~6kHz)。由于输入信号的减少, 从而在整个宽带内增加了比特率的效率。13kbit/s 宽带 SB-CELP(子带激励线性预测)语音编码在文献[8]中有详细描述: 它的解码部分只使用低频带解码信号, 并采用了 HFR(高频重建)方法产生高于 6kHz 的高频部分。对于它的编码结构, 使用比特率转化模块从输入的宽带信号(7kHz)中提取 0~6kHz 低频子带信号, 并使用线性相位分析滤波器使采样速率从 16kHz 降低到 12kHz。对于解码结构, 将接收到的比特流使用 ACELP 解码器合成低频子带信号。为了提高感知质量, 采用后置滤波器对信号进行处理。收到信号经过比率转化模块与后置滤波器输出内差到原始的 16kHz 的采样速率处。最后经过高频重建(HFR)模块生成高频子带信号。

0~6 kHz 带宽的 ACELP 的编码主要包括短时线性预测(LP)分析、长时预测(LTP)分析、固定码本(FCB)激励、感知加权、自适应后置滤波器。其中,长时预测分析是基于将一个自适应码本(ACB)的开环和闭环的长时分析相结合而进行的。自适应后置滤波器由一个级联的共振峰后置滤波器、一个谐波滤波器和倾斜补偿滤波器组成。

在文献[9]16kbit/s 和 24kbit/s 的 SB-CELP 算法所用的频带分裂点为 5 kHz,它采用不相等的两个子带,即 0~5kHz 和 5kHz~7kHz,低频子带使用 ACELP。SB-CELP 算法可以基本实现 ITU-T Q.20/16 标准所要求的语音质量。另一种与上述不同的方法是将子带与 LPAS(线性预测综合分析)编码技术相结合,即 MB-CELP(多频带的 CELP)^[10]实现 24kbit/s 的低延时的宽带语音,并在 CELP 基础上采用前置预测 LP 分析和编码。在 MB-CELP 算法中的 FCB(固定码本)激励由 9 个滤波器的随机码本组成,与不相等划分频带空间相对应。结合 FA/BA-LP 分析和验证,该算法的效果比 G. 722 编码的 56kbit/s 好。

1.2.3 采用 G. 729 算法^[2]

G. 729 是 ITU-T 组织制定的适用于 PSTN 的第四代语音编码标准,该协议采用共轭结构——算术码本激励线性预测(CS-ACELP)算法,其带宽约 4kHz,工作速率为 8kbit/s, MOS 分在 4.0 左右,达到长途电话质量的语音编码国际标准。由于 G. 729 编解码器有较高的带宽利用率,所以被广泛应用于 DSVD、数字卫星系统、电视会议、可视电话、IP 电话及各种多媒体通信系统中。文献[11]实现的宽带语音编码中,采用 G. 729 标准作为基础编码,而根据信道的情况,解码器输出的语音可选择窄带或宽带。该方案由基础编码器和增强编码器两层构成。其中,基础编码器使用 G. 729 算法对窄带语音解码。增强编码器是基于窄带的 CELP 模型,它利用可得到的基层信息对宽带语音解码。可伸缩带宽编码器的设计采用 8kbit/s 的 G. 729 基础编码器和 6.4kbit/s 的 G. 729(附录 D)基础编码器。

CS-ACELP 结构框图如图 3 所示,对于基础层,使用低通滤波器,将输入的宽带语音降低采样率,以便适应基础层编码器。基础层编码器是将窄带语音编码生成基层比特流。增强层编码器是直接将输入的宽带语音编码产生增强层的比特流。解码器输出的语音可根据信道情况选择窄带或宽带。考虑到宽带语音信号的特点,语音的主要信息包含在 0

~4kHz 的频带中,语音信号采用 CELP 算法模型,结果就会在一些编码参数(如线性谱对和基音)中存在冗余,因而提出了 4 种增强编码器结构。其中,增强编码器结构 D(图 4)最好。系统将结构 C 与 B 的观念结合,产生 2 种类型的 ACBs,通过开关选择其中的一个。如果只选中 ACB2,该系统提供与结构 B 相同的输出。如果选中 ACB1,该系统的输出与结构 C 的输出相同,对于码本的搜索过程与系统 C 的相似。不同之处在于通过开关选择两个 ACBs 中码本的失真较低的一个。经过客观的听音评价,该算法实现 16kbit/s 编码的宽带语音质量要比 16kbit/s 的 MPEG-4 CELP 的算法效果好。

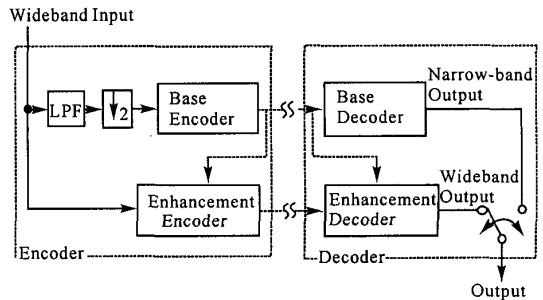


图 3 CS-ACELP 的结构框图

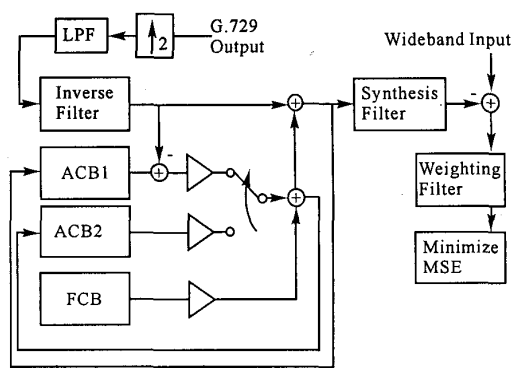


图 4 增强编码器结构 D 框图

1.3 其它的编码算法

宽带语音通信中除了上述可以采用的波形编码算法以及具有潜力的 CELP 算法外,还有以下的几种算法。

1.3.1 变换编码算法

变换编码(Adaptive Transform Coding,简称 ATC)即自适应的变换编码,它是设计低比特率宽带语音编码的另一种主要方法^[4,9]。ATC 算法已经广泛的应用在音频编码上。其缺点是编码延迟通常较大。对于视频会议,采用 ATC^[12]作为基础的编码算法,是考虑到视频会议实际的情况,通常有多个讲话者同时讲话,存在背景噪声,并且一般情况下视频

的编码延迟要比话音延迟大,因而可以不考虑该算法的编码延迟。ATC编码器是基于MDCT(modified discrete cosine transform)即修正离散余弦变换,MDCT的系数由频谱的包络校正,使用自适应比特分配算法进行量化,然后发送。ATC对频谱包络和功率预测的残留也进行量化与传输。

1.3.2 变换码激励和变换预测编码算法

变换码激励 TCX (Transform Coded Excitation) 和变换预测编码 TPC (Transform Predictive Coding) 算法^[4]是最有希望解决语音和音乐同时实现低比特率编码的方法,其基本思想是:在变换域中对线性预测编码器激励信号编码。TCX和TPC将时域的LPAS编码与ATC方法的优点结合在一起。线性预测允许应用效率高的频谱包络或共振峰结构,使用LSF表现编码参数。在时域中,基音预测通过闭环的ACB查找,这对低比特率的语音输入非常重要。在变换域中,对额外的残余或“固定”的激励信号进行量化,其优点是在误差判据方面具有适应性。例如,通过较好感知模型,使整个编解码器具有适中的复杂度和简单的可分级性。TCX的算法中感知加权滤波器和闭环基音预测被应用在时域中,产生的目标信号通过DFT变换被转换到频率域中。通过对TCX的感知加权滤波器的存储器进行适当的处理,结构上仍保持它的AbS特性。

2 两种低比特率宽带语音编码的实现

基于以上的子带编码思想,我们通过C和MATLAB仿真实现了2种不同比特率的宽带语音编码。

2.1 G.729A、MELP算法的语音实现

仿真语音(采样频率16kHz,16bit的PCM)通过一组由12阶的巴特沃思低通和高通滤波器组成的低时延的IIR带通滤波器,将输入音频信号的频带划分为等频带的高低频带,即0~4kHz、4kHz~8kHz。然后采用G.729A^[2,15]作为低频部分的编码;高频部分采用编码速率较低的MELP^[13,14]算法。在信道上传送时,将每个子带的代码复合起来。在接收端译码时,将每个子带的代码单独译码,然后把它们合成起来,形成合成语音输出。

以上方式实现了10.4kbit/s比特率的宽带语音编码,仿真结果如图5所示。

2.2 G.729A、G.729D算法的语音实现

采用G.729算法;使用QMF技术将宽带语音

的频带划分为0~4kHz、4kHz~8kHz,采用G.729A标准作为低频部分的编码;而高频部分则采用G.729D算法,实现14.4kbit/s的宽带语音编码,仿真结果如图6所示。

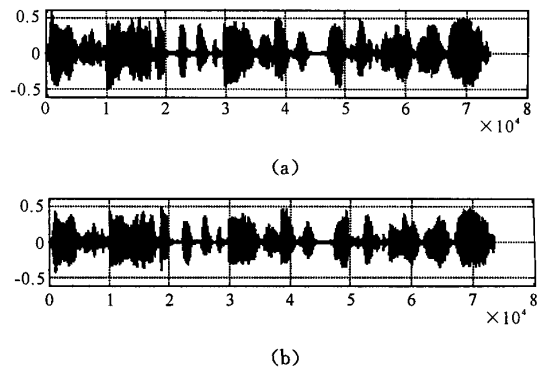


图5 G.729A、MELP算法实现的语音波形对比
(a)原始语音波形;(b)合成语音波形

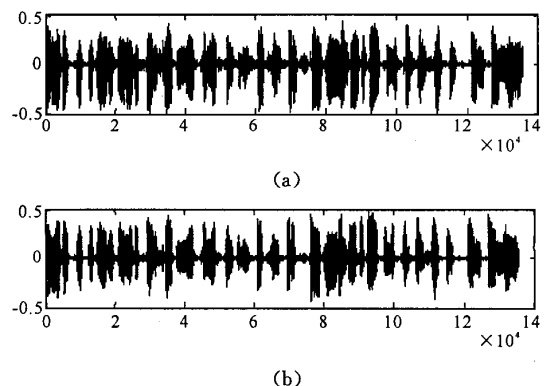


图6 G.729A、G.729D算法实现的语音波形对比
(a)原始语音波形;(b)合成语音波形

从语音输出波形图的对比可以发现,这两种算法的2个波形都很接近,经过听力试验MOS分可达3.1~3.6,表明输出合成语音的质量达到预定的要求。

3 结束语

宽带语音编码技术可广泛应用于各种通信系统和平台,如GSM,3G移动通信系统,VoIP系统等。由于宽带语音编码技术在不同的平台具有广泛的应用,因而需要在比特率、带宽以及算法的复杂度方面要具有可伸缩性。随着宽带语音编码在人们日常生活的应用不断深入,对于宽带的需求日益增大,使得国际电信联盟不断对一些标准进行改进。与此同时,

(下转第219页)

3.4 B 子块的传递

由上分析可知, B 的列块子矩阵可能列数不同, 可通过发送方将待发送子块列数先行发送来了了解接收子块的处理器接收的列数. 另外, 子块的列数应与待发送子块捆绑发送, 减少一次发送启动, 降低发送开销.

3.5 死锁避免

本文算法采用阻塞方式发送和接收数据, 由于所有处理器都采用先发送后接收的方式传递数据, 会产生死锁现象. 要避免死锁必须指定其中 1 个处理器以先接收后发送的方式传递数据(图 4), 以破坏死锁产生的环路条件.

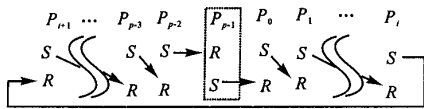


图 4 接收发送方式

P 表示处理器; S 表示发送; R 表示接收.

3.6 矩阵数据的获取

本文提出的算法处理的矩阵数据为事先输入, 并且以文件方式存储, 矩阵 A 以行顺序存储, 而矩阵 B 以列顺序存储, 其目的是使得列的划分问题转

化为行的划分, 从而使行、列的划分采用同样的代码实现. 为便于划分, 在向文件写入矩阵数据之前将矩阵 A 的行数、列数及矩阵 B 的列数先行写入文件.

4 结束语

现实世界中许多现象都表现出并行性, 众多问题的求解过程都有并行的可能性, 但由于人们习惯用 SISD 计算模型上的思维, 使得编写并行机执行程序变得不合常规, 其实, 底意识的并行才更接近问题. MPI 程序的 SPMD 编程模式给人们进行并行思维以很好的训练, MPI 的通信机制为人们在连网工作站上编写并实现并行程序提供了舞台, 使得问题求解变得更加自然.

参考文献:

- 1 孙家昶, 张林波, 迟学斌, 等编著. 网络并行计算与分布式编程环境. 北京: 科学出版社, 1996.
- 2 都志辉编著. 高性能计算并行编程技术——MPI 并行程序设计. 北京: 清华大学出版社, 2001.
- 3 Barry Wilkinson, Michael Allen 著. 并行程序设计. 陆鑫达, 等译. 北京: 机械工业出版社, 2002.

(责任编辑: 黎贞崇)

(上接第 216 页)

各个国家的研究人员对现有的语音编码算法进行深入研究, 也取得一定的成绩(如 0.8~1.5 比特/采样的窄带语音编码), 但各种输入信号的编码的统一模型还没有形成. 在不久的将来, 需要更为灵活的高质量数字宽带语音编码, 这将对宽带语音算法提出更高的要求.

参考文献:

- 1 ITU-T. Recommendation G. 723. Extensions of Recommendation G. 721 adaptive differential pulse code modulation to 24 and 40 kbit/s for digital circuit multiplication equipment application, 1988.
- 2 ITU-T. Recommendation G. 729-Coding of Speech at 8kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP) 1996. Annex A: Reduced complexity 8kbit/s CS-ACELP speech codec, 1996.
- 3 CCITT. 7KHZ Audio Coding within 64kbit/s. In Recommendation G. 722, vol. III. 4 of Blue Book, Melbourne, 1998. 269~341.
- 4 Schnitzler Ju Krgen, Vary Peter. Trends and perspectives in wideband speech coding. Elsevier Science Signal Processing, 2000, (80): 2267~2281.
- 5 ITU-T. Recommendation G. 711, Pulse Code Modulation (PCM) of Voice Frequencies, 1988.
- 6 ITU-T. Recommendation G. 721, 32 kbit/s adaptive differential pulse code modulation (ADPCM). 1988.

- 7 Guyader A Le, Lamblin C, Boursicaut E. Embedded algebraic CELP/VSELP coders for wideband speech coding. Speech Communication, 1995, (16): 319~328.
- 8 Schnitzler Jurgan. A 13.0 kbit/s inband speech codec based on sb-acelp. Proc of ICASSP, 1998.
- 9 Combes Pierre, Schnitzler J. 16, 24, 32 kbit/s Wideband Speech Coder Based On ATCELP ICASSP, Phoenix, AZ, USA, IEEE, 1999.
- 10 Ubale Anil, Gersho Allen. A Low-delay Wideband Speech Coder At 24 kbps. In Proc. IEEE Int Conf Acoustics, Speech, Signal Processing, 1998. 165~168.
- 11 Kataoka A, Kurihara S, Sasaki S, et al. A 16-kbit/s wideband speech codec scalable with G. 729 Proc EURO-SPEECH'97, 1997. 1491~1494.
- 12 Asakawa Yoshiaki, Okamoto Toru, Susa Akiko, et al. A 24-kbit/s Wideband Speech Coder for Videoconferencing Systems. 1990. 1671~1674.
- 13 McCree A V. A 2.4Kb/s MELP coder candidate for the new U. S. Federal standard. Proc. ICASSP, 1996, 200~203.
- 14 McCree A V, Barnwell T P. A mixed excitation LPC vocoder moder for low bit rate speech coding. IEEE Trans. Speech & Audio Processing, 1995, 3 (4): 242~250.
- 15 覃团发, 刘运毅, 张淑仪. G. 729A 声码器的 VC++ 面向对象封装实现. 电声技术, 2003, (12): 13~16.

(责任编辑: 黎贞崇)