

文章编号: 1007-2861(2001)01-0013-05

G 729 CS-ACELP 语音编码算法的 优化及其DSP实现^{*}

许丽红, 阚海鹰, 余小清, 万旺根

(上海大学 通信与信息工程学院, 上海 200072)

摘要: 在研究 G 729 CS-ACELP 语音编码算法基础上, 通过分析其原理及基本特征, 在实际应用中提出了优化算法的两种方案, 大大降低了算法复杂度; 并介绍了利用 DSP 相应的开发环境对简化后的算法进行软件模拟, 实验结果证明其输出语音仍然保持了很高的合成品质; 最后给出了以 DSP 为核心的硬件系统构成框图

关键词: G 729 算法; 语音编码; 算法优化; DSP 芯片

中图分类号: TN 912.32 **文献标识码:** A

A Speech Coding Algorithm Optimization Based on G 729 CS-ACELP and Its Implementation

XU Li-hong, KAN Hai-ying, YU Xiao-qing, WAN Wang-gen

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

Abstract: This paper presents some optimized methods of G 729 (CS-ACELP) standard based on studying its principle and characteristics in applications which can considerably reduce the calculation burden. It also introduces the simulator in C54x CCStudio and the hardware system including DSP. Experiments show that output speech has a high reconstructed quality with the optimized G 729 algorithm.

Key words: G 729 algorithm; speech coding; algorithm optimization; DSP chips

语音编码作为多媒体通信中信息传输的一个重要环节, 越来越受到广泛的重视, 语音压缩编码技术也日趋成熟。早期的波形编码, 语音质量好, 但同时也对传输速率提出了很高的要求; 后来发展起来的参数编码, 逐步降低了编码速率, 但重建语音的质量有相对幅度的下降^[1]。现在更多的编码技术将两种方式结合起来, 兼顾了传输速率与语音质量的要求, 使其发挥各自的长处, 在 4~16 kbit/s 传输速率上

可以得到高质量的合成语音。

G 729 CS-ACELP 语音编码算法就是这样一种以共轭代数码本激励线性预测为基础, 且编码速率为 8 kb/s 的高质量的语音编码标准。由于其在线性预测参数量化上, 使用分裂式矢量量化; 在自适应码本搜索上, 采用开环、闭环相结合, 分数延时码矢量搜索; 固定码本采用具有特殊结构的代数码本; 激励序列的增益, 采用了预测式矢量量化, G 729 获得

* 收稿日期: 2000-07-05 基金项目: 国家自然科学基金(60072031); 上海市教委曙光计划(98SG38)

作者简介: 许丽红(1977~), 女, 河北平山人, 硕士生, 主要从事语音数字信号处理方面的研究

了较高的语音合成质量和较低的语音算法复杂度

在银行系统、公安系统、空中交通管制等领域中,很多语音对话必须能够记录下来以便于核对查询,这就需要高质量的数字语音记录设备。G. 729 语音编码算法由于其具有的算法复杂度低、编码效率高、语音质量好等优点,被数字语音记录设备所广泛采用。算法复杂度越低,实时实现的硬件系统越简单,而 8 kbit/s 的高效率编码使其能够在保证语音质量的前提下,大大减小语音记录所需要的存储空间。

本文在简要阐述 G. 729 语音编码算法的基础上,提出了在实际应用中优化算法的两种方案,并对此优化算法进行的软件模拟,还给出了由 DSP 来实现数字语音记录系统的硬件结构框图,为实现实时处理系统奠定了基础。

1 G. 729 CS-ACELP 算法的特点^[2]

G. 729 编码器将输入信号分为 10 ms 语音帧,对每帧提取 CELP 模型参数(包括线性预测参数、固定码本和自适应码本码矢量以及对应的增益),并将其编码后的比特流(80 bit/frame)送信道传输。图 1 给出了 CS-ACELP 算法编码器原理框图。

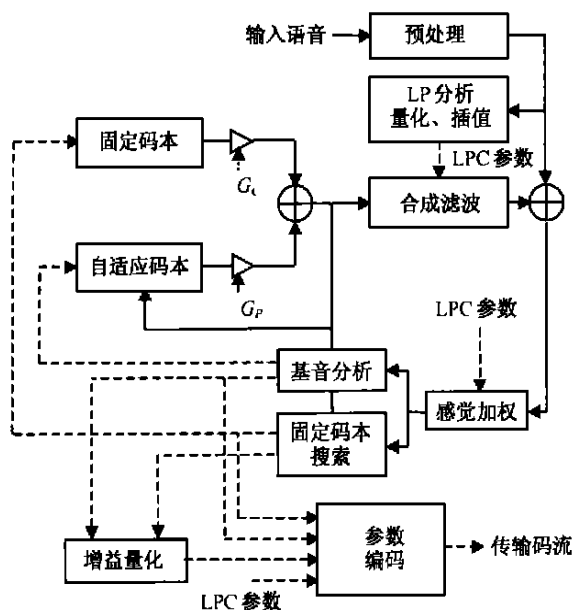


图 1 CS-ACELP 编码器原理图

Fig. 1 The principle of the CS-ACELP encoder

8 kbit/s CS-ACELP 算法的帧长为 10 ms,分两个子帧进行处理。采样输入的语音信号首先经过高通滤波器去除直流分量及低频分量的干扰,然后

按每一帧提取 10 阶 LP 参数并转化为 LSP (线谱对) 参数进行量化、编码;用量化后的 LSP 参数求得的 LP 参数分别构成感觉加权滤波器 $W(z)$ 以及合成滤波器 $1/\hat{A}(z)$;目标信号是从加权语音中除去感觉加权滤波器的零输入响应而求得的,目标信号经过自适应码本搜索及固定码本搜索后产生激励信号,分别乘以各自的增益,并共同激励合成滤波器,产生本地重建语音信号。由此可看出,CS-ACELP 语音编码算法主要包括 LP 参数的提取、量化及编码、自适应码本搜索、固定码本搜索和增益矢量量化四个部分。

1.1 LP 参数的量化、编码

利用 Levinson-Durbin 算法每一帧提取 10 阶 LP 参数,量化时先将其转化为 LSP 参数,以利于系统稳定,然后再按均方误差(MSE)最小原则将其量化为 18 bit 进行传输。由于 LSP 特征矢量的各个分量虽然属于一个整体,但它们对幅度谱的影响只局限于各自的频率区域内,这就使其可以将一维数较高的特征矢量分裂为两个维数较低的特征矢量分别进行矢量量化^[3],即分裂式矢量量化。G. 729 采用两级分裂式矢量量化,首先利用 MA (滑动平均) 滤波器结合前几帧的 LSP 系数预测当前帧的 LSP 码向量^[4],然后根据均方误差最小原则,在码书 1 中搜索出对应的码矢,将其量化为 7 bit;第二级量化将第一级量化得到的最小均方误差矢量分为两个子分量,前五个构成一个子矢量,后五个分量构成一个子矢量,分别量化为 5 bit。采用多级分裂式矢量量化,不但可以大大减少码书的搜索量、存储量,又使整体量化精度得以保持。

1.2 自适应码本搜索

G. 729 自适应码本搜索先采用开环的方法,在每个子帧内利用加权语音的相关性计算出开环延迟后,再用闭环方法在延时为开环延迟的一个小范围内搜索出最佳自适应码矢量。闭环分数延迟的搜索过程分两个子帧进行,第一子帧利用开环延迟求得的参考值搜索,使用 8 bit 编码;第二子帧在第一子帧最佳分数延迟的基础上搜索出本子帧的最佳基音延迟。这是由于语音信号具有短时平稳性,最佳自适应激励码序列前后相差不大,可以采用差分搜索和编码,因此第二子帧可以用差分量化的方法将其量化为 5 bit,从而减少了码本的量化比特,提高了编码效率。

1.3 固定码本的结构及其搜索

传统的 CELP 算法中, 用一个自适应码本中的码字来逼近语音的长时周期性(基音)结构, 而用一个固定的随机码本中的码字来逼近语音的经过短时、长时预测后的信号^[5], 即逼近由自适应预测所得的语音信号与原始语音的误差, 从而为合成语音提供更精确的误差补偿, 提高语音质量

G 729 的代数码本结构采用插入单脉冲排列(ISPP)设计方案, 这种码本的每个码矢都有四个位于不同样点子集的单位脉冲构成, 同时又采用解析码本方案, 解码器利用传输码流本身的信息就可以重构码矢和码本增益, 而不需要象通常的矢量量化方法那样, 去搜索一个具体的码本, 从而提高了效率

1.4 增益的矢量量化

增益量化分为自适应码本增益和固定码本增益两种, 由于自适应码本增益大小与语音信号准周期性强弱密切相关, 它在各子帧之间变化较大, 可直接进行量化; 而固定码本的增益则由于语音信号的短时平稳性而具有较强的帧间相关性, 可以通过增益预测来提高增益的量化精度和量化效率

2 算法的优化

在实际应用中可以根据需要, 在保证合成语音质量的前提下对 G 729 算法进行优化。优化过程从两个方面来考虑, 一是在满足精度要求下, 进一步将算法简化; 二是对高级语言程序代码的优化

2.1 在满足精度的要求下简化算法, 降低计算量

在对 G 729 算法分析中发现, 它大量利用了多级查询或多级量化的方法, 即首先根据某种原则将查询和量化范围逐步缩到某一小区域内, 再进一步提高搜索精度, 缩小范围, 直到找到最终结果, 这样可以减小码本的搜索过程和计算量。G 729 中 L P 系数向 L S P 系数的转化、L S P 系数的量化、基音延迟的搜索等都是这一方法的典型应用

在算法优化过程中, 为了降低复杂度, 可以考虑在满足精度要求的前提下, 针对多级查询或多级量化, 进一步粗化第一步的搜索过程, 从而减少总的搜索时间。例如基音延迟的确定利用先开环搜索, 进而再闭环搜索, 开环搜索将基音延迟的可能范围(20~143)缩小在 t_{\min} 与 t_{\max} 之间, ($t_{\max} - t_{\min} = 9$, 分辨率为 1), 随后的闭环搜索在小范围 t_{\min} 与 t_{\max} 之间找到分辨率为 1/3 的基音延迟值, 算法的优化可以通过粗

化开环搜索后的范围来降低计算量。每一子帧利用判断加权语音的相关性来计算开环基音延迟, 其中相关函数表示为

$$R(k) = \sum_{n=0}^{79} sw(n)sw(n-k),$$

其中, $sw(n)$ 为加权语音; $R(k)$ 为相关值; k 的取值范围为 20~143。简化后相应的表达式为

$$R(k) = \sum_{n=0}^{39} sw(2n)sw(2n-k),$$

其他相关的计算公式也作相应的修正, 即简化后由 80 个样本点中的 40 个偶数点的相关性来代替原来的计算 80 个样本点的相关性, 使计算量降低一半。其他运算过程, 如 L P 系数向 L S P 系数的转化、L S P 系数的量化等也可以用类似的方法简化算法。图 2 给出了原算法与简化后的算法的合成语音波形比较, 结果证明粗化第一步的搜索过程不会影响最终结果, 却可以大大减少计算量



图 2(a) 原 G 729 算法生成的合成语音波形

Fig. 2(a) Reconstructed speech by original algorithm



图 2(b) 简化后的算法生成的合成语音波形

Fig. 2(b) Reconstructed speech by simplified algorithm

2.2 充分利用 DSP 软件开发环境, 优化算法

用 C 语言编写的算法代码, 若直接加载到 DSP 软件开发平台中进行编译, 生成的可执行文件会很庞大, 一般很难与实际的 DSP 硬件环境相匹配。即使利用 DSP 自带的优化编译器, 虽然可将代码的容量及速度优化到原来的 70% 左右, 但效果仍不够理想, 这就需要充分开发 DSP 内部资源, 优化算法。在实际应用中, 可以将定点运算中原来用 C 语言编写的加、减、乘、除、移位、乘加等基本运算的函数操作直接转化为宏汇编指令操作^[6,7], 这样不仅可以节省寄存器的占用, 而且可以提高运算效率。另外一些常调用的函数, 如卷积、相关、FIR 滤波等可结合汇编编程的特点, 直接用汇编编写, 开发出高效率的程序代码

例如 FIR 滤波器中有大量的乘加操作, C 语言编写的程序代码经编译后, 生成的汇编代码大多遵

循两个数先相乘(利用MPY指令),再利用ADD指令将结果累加到某一个寄存器中,这样既需要众多的寄存器暂存数据,又会因指令的繁多(乘、加、装载、存储等)而增加运算时间。优化代码过程中可以充分利用DSP擅长的乘加一次性操作,即MAC指令,使得系统可以在一个指令周期内完成乘加操作,还可以结合寄存器的合理安排,实现连续乘加,而不需暂存中间数据,从而大大提高了运算效率。另外,一些大于一个指令周期才能完成的指令,若在下一条指令中要用到它的结果时,这两条指令之间就必须错开相应的指令周期。对于这种情况的处理,由C语言直接编译生成的汇编代码,往往插入空指令(即NOP指令,除延时一个周期外,不做任何操作);在直接用汇编编写的程序中就可以根据实际需要,运行其它与前面的等待结果不相关的指令,提高系统的并行处理能力,从而提高程序运算速度。

3 优化后 G 729 算法的软件模拟实现

为了能够在基于DSP的数字语音记录系统中实现实时语音压缩,在对G 729算法进行了上述优化处理并将此算法C代码编译通过后,采用相应的DSP集成开发工具C54x CCSstudio (Code Composer Studio (Simulator))来模拟实现整个编解码过程。

软件模拟实际上是借助PC机本身的CPU及内存资源仿真DSP的处理过程,因此在软件模拟的操作环境下,模拟器只是用来检测算法在目标DSP系统上实现的可行性与准确性。为了能够真实地反映目标系统,在模拟之前要进行模拟器的环境配置,其中包括为了使CCStudio能够作为软件模拟仿真器(Simulator),利用CCS(Code Composer Setup)进行的系统配置;另外,还包括在DSP可执行代码装载前,设置内存映像,来实现PC机内存的DSP存储器模拟,其中包括程序空间的映射、数据空间的映射及I/O空间映射。程序在编译连接时,将目标模块的各段映射到连接命令文件所指定的地址区域,这些地址信息最终都包含在编译输出的out文件中,仿真时装载的out文件根据所包含的映射信息把程序的各段放在软件模拟的内存环境中,同时进行内存环境的校验。在配置时要注意,若使用中存在数据/程序空间重叠映像,如数据空间同时映射为程序空间,那么在软件模拟仿真中需要在程序加载前

设置相关的DSP寄存器标志位进行区分。

优化后的G 729算法在模拟实现中,原始语音采用16位、抽样率为8000 Hz的PCM信号,帧长10 ms,针对每帧提取CELP模型参数,并将其编码后的比特流(80 bit/frame)送信道传输或存储,编码速率为8 kbit/s,解码端经处理后输出与输入一致的PCM语音信号。根据G 729算法的复杂度要求,采用C5402 DSP的系统配置,C5402的16 K在片DARAM为程序/数据空间,可以用来加载执行代码;外部的48 K空间全部映射为数据空间,用于初始化数据的加载或程序运行时创建和保存中间数据^[8],内存映像也根据C5402在片资源来设定。连接文件(后缀为cmd的文件)应该与根据C5402在片资源所设置的内存映像相一致。

CCStudio为软件开发人员提供了高效的调试工具,其中包括单步调试、断点设置、变量值的实时读取、信号波形的实时观测等。软件仿真时原始语音信号采集、编码后的码流的存储及合成语音的回放等接口的模拟,都通过在相应的位置设置探测点(Probe Point),利用控制文件的读取、写入方式,设置数据读取、存储的起始地址、标定数据每帧的读取长度等来完成。调试过程中要查看变量值和图形显示,就需在相应位置处不仅设置探测点,而且要设置断点(Breakpoint),其中探测点控制数据的读取,断点控制数据、图形界面的手动刷新。另外利用系统提供的描述点(Profile Point)设置功能,还可以对算法复杂度进行实时测试,在需测试的一段子程序两端设置描述点,系统会统计出此段程序的指令周期数、执行时间等,时间测定所遇到的时钟设置要依据具体的DSP硬件环境,这里用到的C5402芯片,处理速度为100 MIPS,其指令周期即为10 ns。

单步调试过程中可以对照原C语言程序,经过CCStudio编译形成的汇编代码、汇编指令详细说明来监测程序的运行及寄存器、标志位的变化。对于遇到的不可单步执行指令,利用光标、断点控制等可以先跳出,再继续。一般在运行DSP可执行文件前,要先将DSP对应的模拟寄存器复位(初始化),避免由于各种标志位的不合理设置而影响整个调试过程。

经过对优化后的G 729算法在上述系统中进行的软件模拟测试,一些子程序(如高通滤波、开环基因搜索等)的运算时间可减少到优化前相应算法的4%左右,而整个算法的运算时间也减少到优化前G 729算法所需时间的8%,从而大大降低了算

法复杂度, 输出语音的质量却没有明显的变化。下图3(a)所示为输入的原始语音波形, 运用优化后的G. 729算法对其分析合成后, 其输出语音波形如图3(b)所示, 对比发现两个波形很接近, 听力实验结果表明输出的合成语音质量达到了预定的要求。



图3(a) 输入语音波形

Fig. 3(a) The original speech



图3(b) 输出的合成语音波形

Fig. 3(b) The reconstructed speech

4 基于优化后 G. 729 算法的数字语音记录系统构成

作为语音处理平台^[9,10], 不仅要能够完成对语音信号的实时采集、解码语音回放及数字语音的编解码运算等, 而且要为语音处理提供必要的外围电路及硬件设备, 这包括DSP系统所需的电源管理, DSP芯片与PCI总线的接口电路, 编码后的数据存储设备(这里用高速硬盘)及系统操作界面等。图4给出了语音处理系统的硬件构成框图。

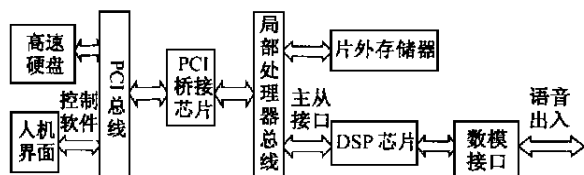


图4 数字语音记录系统结构图

Fig. 4 The structure of speech recorder system

在充分权衡优化后G. 729算法复杂度的要求以及DSP市场价格后, 选择TI公司的TM S320VC5402作为语音编码处理芯片。C5402 DSP的处理速度为100MIPS, 能处理至少4路基于G. 729算法的语音编解码, 它的高性价比使其成为系统设计的首选。

5 结论

本文针对G. 729算法的基本原理, 提出了实际应用中优化算法的两种方法, 并在此基础上利用C54x CCStudio对优化后的算法进行了软件模拟, 同时给出了一个实时的DSP语音处理系统硬件设备构成。经过软件模拟测试, 简化后的算法运算时间可缩减到原来的8%, 算法复杂度大大降低, 而输出的合成语音质量也达到了预定的要求, 为数字语音处理系统在实际应用中的实时实现奠定了基础。

参考文献:

- [1] Bao Changchun, Fan Changxin. A review of speech coding [J]. Journal of the China Institute of Communications, 1998, 19(5): 45- 55.
- [2] ITU-T Recommendation G. 729-Coding of Speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP) [R], 1996.
- [3] 唐 昆, 等. 高质量4~8 kbit/s变速率有限状态ACELP语音编码算法研究 [J]. 电子学报, 2000, 28(1): 21- 25.
- [4] Redwan Salami, Claude Laflamme. Design and description of CS-ACELP: a toll quality 8 kb/s speech coder [J]. IEEE Transactions On Speech And Audio Processing, 1998, 6(2): 116- 130.
- [5] 杨行峻, 迟惠生, 等. 语音信号数字处理 [M]. 北京: 电子工业出版社, 1995. 251- 264.
- [6] TM S320C54x C Source Debugger User's Guide [R]. Texas Instruments, 1998.
- [7] TM S320C54x DSP Mnemonic Instruction Set, Reference Set [R]. Texas Instruments, 1996, 2.
- [8] TM S320C54x DSP CPU and Peripherals, Reference Set [R]. Texas Instruments, 1999, 1.
- [9] 李勇华, 万旺根, 等. 一种新型语音处理系统 [J]. 上海大学学报(自然科学版), 1999, 5(3): 265- 270.
- [10] 郑伯诺, 万旺根, 等. 基于TM S320C25 DSP的一种实时数字语音处理系统 [J]. 上海大学学报(自然科学版), 1998, 4(2): 212- 216.