

G.729 语音编码算法的分析 及其基于 ARM9 的优化实现

张辉¹, 陈小云²

(1.大同市电子科技研究所, 大同 037008; 2.中国科学院 长春光学精密机械与物理研究所, 长春 130033)

摘要: 本文对 G.729 编解码原理进行了研究, 提出了在 ARM9 平台上实现 G.729 语音编码器的方案, 并介绍了从算法、C 程序等方面对程序进行优化, 发现比优化前的算法大大减小了编码时间, 完全达到实时的要求, 并且语音质量没有明显下降。

关键词: 语音编码; G.729; ARM9; 优化

中图分类号: TN912

文献标识码: A

文章编号: 1672-9870(2011)03-0164-03

The Analysis of G.729 Audio Signal Coding Algorithm and Its Implement Based on ARM9

ZHANG Hui¹, CHEN Xiaoyun²

(1. Datong Institute of Electronic and Technology, Datong 037008;

2. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Science, Changchun 130033)

Abstract: This paper introduces the algorithm structure of G.729. It is proposed to implement a scheme of G.729 speech codex based on ARM9 platform. It is considered how to optimize the program from algorithm, C language. After optimization, the performance of speech compression process is improved to match the real time requirement, and the sound quality not obviously decreased.

Key words: speech coding; G.729; ARM9; optimization

G.729 是国际电信联盟标准部门在语音编码方面取得的最大标准化成果。G.729 语音编码标准为语音信号的高效传输提供了一种有效方法, 其本质就是在较高的语音质量指标下, 降低数字化语音的码速率。本文细分了 ITU-T G.729 算法, 然后提出了在 ARM9 平台上用三星公司的 S3C2440 处理器实现 G.729 语音编码器的方案, 并对程序进行了优化。

1 G.729 编解码原理

G.729 音频编码算法结构图如图 1 所示。G.729 协议^[2]是基于码激励线性预测(CELP)声码器模型, 对于 8K 的语音信号其数据流为 80 个采样值而言, 该模型所运行的帧长为 10ms。对于一帧语音信号,

经过 LP 分析, 得到 CELP 模型的编码参数(LP 滤波器系数)。这些参数属于时域参数, 将其转换成线谱对(LSP), 同时使用预测型两级矢量量化(VQ)器进行量化为 18bit。本方法中, 利用理解加权的失真测度激励将原始信号和重构信号之间的误差进行最小化。这可以通过理解加权滤波器实现, 理解加权滤波器一般采取自适应算法, 以便改善输入信号的性能。

对于 8K 语音信号, 需要将一个帧分为两个子帧, 即 5ms, 40 个采样点。这样激励参量需要一个帧确定。在第一个子帧, 使用内插的 LP 滤波器的系数(量化的和未量化的)。每 10ms(一帧)估计一次, 开环音调延迟, 接下来每一子帧重复一次。

加权合成滤波器 $W(z)/A(z)$ 是 LP 通过加权残差

收稿日期: 2011-06-19

作者简介: 张辉 (1967-), 男, 工程师, 主要从事电子技术与计算机应用方面的研究。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

获得了,本项目目标信号。也就是滤波器的初始状态是由LP激励和残差的误差来获得的。也是实现了加权语音信号中去除滤波器的零状态响应。滤波器的响应时间需要通过计算来获得。

以上为开环语音信号分析,下面进行作闭环音调分析,即开环音调延迟搜索,仍然使用目标信号 $x(n)$ 和脉冲响应 $h(n)$ 。音调延迟编码原则是在第一个子帧是用8bit编码,而在第二个子帧是用5bit差分编码,并使用1/3分辨度的分数进行延迟。目标信号 $x(n)$ 除去了自适应编码后,得到新的目标信号 $x'(n)$,也是经过修正后的目标信号。将形成的新的目标信号,用在固定码本的搜索中,容易得到最佳激励。固定码本激励是使用一个17bit的代数码本。自适应码本和固定码本的增益用7bit矢量化。最后,滤波器存储器使用已经确定的激励信号来修正。

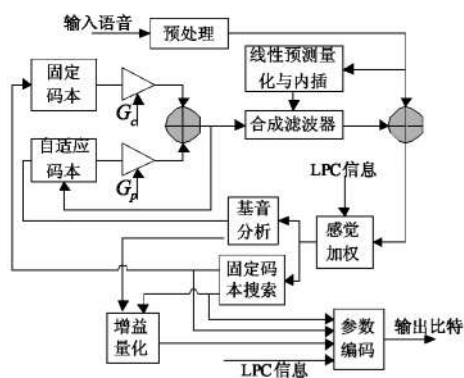


图1 G.729音频编码算法结构图

Fig.1 Structure diagram of G729 audio coding algorithm

G.729语音解码器的原理如图2所示。首先,在接收到的语音流信号中,提取编码是所提供的参量指针,也就可以读取响应的一帧即10ms,80个采样的编码参量也就是LSP系数。其中可以演算获得音调延时参数、固定码本增益等参数。其次,每一个子帧的LSP系数要进行内插,并转换为L滤波器系数。最后,对美5ms的子帧完成下面的运算:

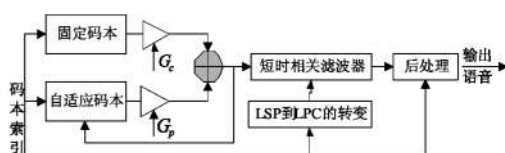


图2 G.729音频解码算法结构图

Fig.2 Structure diagram of G729 audio decoding algorithm

(1)将自适应码本矢量与固定码本矢量进行增益定标;

(2)定标之后的自适应码本矢量与固定码本矢

量进行相加;

(3)将目标信号 $x(n)$ 信号通过滤波器滤波获得合成语音 i 信号。

(4)将合成语音信号通过自适应滤波器和高通滤波器进行滤波和定标获得,最后的语音信号的输出,即进行了解码。

2 编码器的实现

2.1 编码器硬件平台

本系统硬件平台采用的CPU为Samsung S3C2440 ARM920T内核控制器,实现了MMU、AMBA BUS和Harvard告诉缓冲体系结构。主频400MHZ,最高可达533 MHZ。系统配置了一片nand flash(64M),采用两片半字(half-word)SDRAM器件共同组成一个32位数据宽度的SDRAM系统,提高了其与CPU的通信效率。通过提供一整套完整的通用系统外设,S3C2440A减少整体系统成本和配置额外的组件^{[3][4]}。

音频处理器选用的UDA1341是Philips公司推出的音频芯片,UDA1341支持IIS总线数据格式,采用位元流转换技术进行信号处理,完成声音信号的模数转换,具有可编程增益放大器和数字自动增益控制器,具有低电压、低功耗的特点。UDA1341对外提供2组音频信号输入接口,每组包括左右2个声道。支持IIS总线数据格式,IIS总线具有正常传输、DMA、传输/接收三种工作模式。

本文针对UDA1341在ARMLinux下的驱动程序,重点分析了缓冲区的设计和音频数据的编码和解码,并实现了基于该芯片的录制和播放。最后通过实现其功能函数,完成了UDA1341驱动程序的开发。

微型处理器与音频处理器的原理如图3所示。UDA1341TS由2组音频信号输入线、1组音频信号输出线、1组IIS总线接口信号线和1组L3总线组成。音频信号输入线、音频信号输出线主要用于实现音频信号A/D及D/A的变换;通过IIS总线接口,实现与S3C2440之间音频信号的通信。UDA1341的I音频输入输出(VIN,VOUT)分别和麦克风,扬声器连接;IS引脚分别接到S3C2440对应的IIS引脚;UDA1341的L3接口相当于一个Mixer控制器接口,可以用来控制输入/输出音频信号的音量大小,低音等。L3接口的引脚L3MODE,L3DATA和L3CLOCK分别连接到2440的IICSCL,IICSDA和GF6引脚。

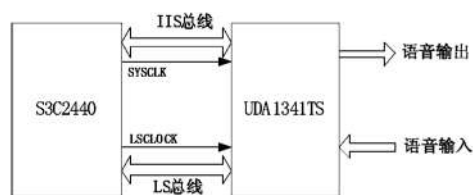


图3 语音采集芯片与2440连接电路连接图

Fig.3 Circuit connecting diagram of audio acquisition chip and 2440

2.2 编码器软件实现

2.2.1 G.792编码器的C程序实现

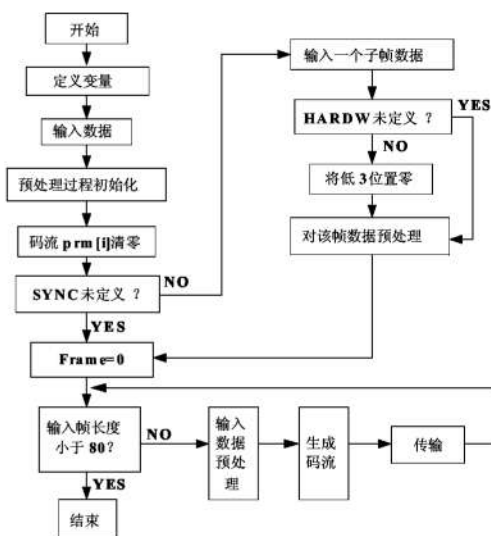


图4 主程序流程图

Fig. 4 Flowchart of main program

图4中,我们以主程序为例来介绍其流程及调用关系。在驱动程序开发过程中完成了信道编码函数(prm2bits-1d8k()),在主程序中直接调用即可。主程序首先对数据区及代码区进行初始化操作,将输入信号存在缓冲区的数据进行提取,利用bits.c程序中prm2biits-1d8k函数进行信号编码,获得比特流。

这些文件及函数分别在程序运行的不同阶段被调用。其中main.c作为这个语音处理子系统的主程序文件。

算法实现

1. 算法优化

固定码本矢量 $c(n)$, 由构造公式得^[5]:

$$c(n) = c_0\theta(n-m_0) + c_1\theta(n-m_1) + c_2\theta(n-m_2) + c_3\theta(n-m_3) \quad n=0, \dots, 39 \quad (1)$$

由文献[5]可知, $c(n)$ 在采样时 n 只有在前4个点上取值,而在其他点上都为零。因此,可以利用这一特点对 $c(n)$ 的计算进行优化。在进行增益量化的过程中,需要计算 $z(n)$:

$$z(n) = \sum_{i=0}^n c(i)h(n-i) = c(n)*h(n) \quad n=0, \dots, 39 \quad (2)$$

其中 $h(n)$ 为合成滤波器系统冲击响应。由式(2)可知, $n=39$ 时,一共需要计算780次乘累加。

利用上面阐述的文献[5]的 $c(n)$ 的特性,进行增益量化优化后,得到:

$$z(n) = c(n)*h(n) = \sum_{j=0}^n h(j) \cdot \sum_{k=0}^3 c_k \theta(n-j-m_k) = \sum_{k=0}^3 c_k h(n-m_k) \quad n=0, \dots, 39 \quad (3)$$

而 c_k 为 ± 1 , 所以上边的乘法计算实际上是加减运算。通过比较(2)(3)式可知,运算量骤减,减少了近5倍,计算量大大减小。

2.C程序的优化与实现

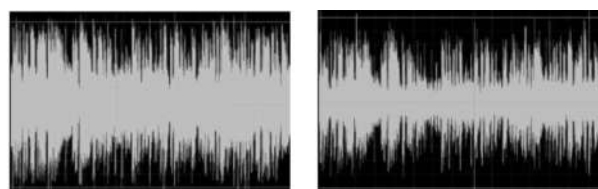
对上述的算法,采用C语言进行了实现,具体方法如下:

首先,用预读取的40个采样数据来取代,原先预读5ms的数据。通过仿真结果看,取消预读5ms数据后的语音效果比取消前的MOS值没有明显下降,同时优化后,能节省5ms的延迟时间。其次,由于对目标信号已经进行了除2操作,大大降低了结果的溢出,因此取消部分对结果的溢出校验,提高了运算的速度。

2.2.3 仿真结果

通过以上的优化方法,在进行一帧语音信号的编码中,延迟时间能够控制在10ms以内,程序的执行效率明显提高。

最后,我们在 S3C2440 对上述实现的 G.729 算法进行测试,我们对事先存放在 FLASH 里的一段原始 WAV 音频文件进行编码。将编码的参数以文件方式存储,并通过网口传回基于 windows 平台的 PC 机。最后在 PC 机上用常用的 G.729 解码软件进行解码,图5展示了编码前的原始文件和编码后的解码文件的波形。



(a) 原始语音文件

(b) 解码后语音文件

图5 软件仿真后的文件波形

Fig.5 File waveform of software simulation

(下转第169页)

III)若 $\dim F = n = \dim E$, 则 $\dim Y = 0$, Y 是一个点. 此时 $n+1 \geq l(F) \geq l(X) = l(R) = n-1$, 所以 $l(F) = n+1$ 或者 n 或者 $n-1$.

(1) 当 $l(F) = n+1 = \dim F + 1$, $(F, L_F) = (P^n, \mathcal{O}_{P^n}(1))$, 由 $l(X)$ 的定义, 存在极小极端有理曲线 $C \subset X$, 有 $l(X) = -K_X \cdot C = n-1$, $L_F \cdot C = L \cdot C = 1$ 故 $K_X \approx -(n-1)L$, X 是 Del Pezzo 流形.

(2) 当 $l(F) = n = \dim F$, $(F, L_F) = (Q^n, \mathcal{O}_{Q^n}(1))$, 同(1)得, X 是 Del Pezzo 流形.

(3) 当 $l(F) = n-1$, $K_F + (\dim F - 1)L \approx \mathcal{O}_F$, (F, L_F) 是 Del Pezzo 流形.

②若 $l(R) = n$, 代入※式, 则 $\dim E = n$, $\dim F = n-1$ 或者 n .

I)若 $\dim F = n-1$, 则 $\dim Y = 1$, Y 是一条光滑曲线. 而 $n = \dim F + 1 \geq l(F) \geq l(R) = n$, 所以 $l(F) = n = \dim F + 1$, 于是我们有 $F \cong P^{n-1}$, X 是光滑曲线 Y 上的 P^{n-1} 丛.

II)若 $\dim F = n$, 则 $\dim Y = 0$, Y 是一个点. 此时 $n-1 = l(X) = l(R) = n$, 矛盾.

③若 $l(R) = n+1$, 代入※式, 则 $\dim E = \dim F = n$, 同上理, 矛盾. 证毕.

(上接第 166 页)

通过对算法和原代码的分析可以看出 G.729 标准对语音信号的压缩有较好的压缩比, 延迟时间也较小, 同时可以满足系统要求质量. 系统中选用此压缩算法是可行的.

3 结束语

本文深入研究了 G.729 语音编码算法, 以及在 ARM9 硬件平台上的实现, 结合 linux 下的 vi 编辑器和 ARM9 指令的特点对 G.729 算法分别进行了算法级、C 语言级和 make 编译工具级的优化, 并在 2440 处理器上测试了该算法, 测试结果表明 G.729 压缩编码算法经过优化后, 性能明显提高, 语音质量符合要求, 并且运算时间大大降低, 能够达到实时传输的目的.

参考文献

- [1] Cho K, Miyaoka Y, Shepherd N I. Characterizations of projective space and applications to complex symplectic manifolds [J]. Advan Stud Pure Math, 2002, 35: 1-88.
- [2] Mori S. Threefolds whose canonical bundles are not numerically effective [J]. Ann Math, 1982, 116: 133-176.
- [3] Fulton W. Intersection Theory [M]. New York: Springer-Verlag, 1998.
- [4] Lazarsfeld R. Positivity in Algebraic Geometry I, II [M]. Berlin: Springer-Verlag, 2004.
- [5] Kawamata Y, Matsuda K, Matsuki K. Introduction to the minimal model problem. Advan Stud Pure Math, 1987(10): 283-360.
- [6] Ionescu P. Generalized adjunction and applications. Math Proc Camb Phil Soc, 1986, 99: 452-572.
- [7] Wisniewski J A. On contractions of extremal rays of Fano manifolds. J Reine Angew Math, 1991, 417: 141-157.
- [8] Hartshorne R. Algebraic Geometry. New York: Springer-Verlag, 1978.
- [9] Fujita T. Remarks on quasi-polarized varieties. Nagoya Math J, 115: 105-123 (1989).

参考文献

- [1] 胡航. 语音信号处理 [M]. 哈尔滨: 哈尔滨工业大学出版社, 1987.
- [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.
- [3] 于明, 范书瑞, 曾祥烨. ARM9 嵌入式系统设计与开发教程 [M]. 北京: 电子工业出版社, 2006: 80-127.
- [4] 杜春雷. ARM 体系结构与编程 [M]. 北京: 清华大学出版社, 2003.
- [5] 王佳. 激光通信中传输音/视频信号的技术研究 [D]. 长春理工大学硕士学位论文, 2010.