

The Optimization of G.729 Speech codec and Implementation on the TMS320VC5402

Geng wang^{1, a}, Wei cheng^{2, b}

¹The School of Electronic and Information Engineering, Hubei University of Science and Technology, Xianning, 437100, China

²The School of Electronic and Information Engineering, Hubei University of Science and Technology, Xianning, 437100, China

^a wanggeng20001982@126.com, ^b799040609@qq.com

Keywords: G.729; TMS320VC5402; TLV320AIC23; Speech Codec

Abstract. This paper designs a realtime speech codec system based on the TMS320C5402 platform, from two aspects of data transmission interface and control interface for the TMS320C5402 and TLV320AIC23 interfacedesign, and through the g. 729 speech decoding algorithm principle research and program optimization on the platform to realize the real-time algorithm.

Introduction

The G.729 algorithm is very complicated, and the real-time effect is dissatisfactory, but with rapid development of the DSP which has a strong operation ability and research on all kinds of optimization methods, this algorithm has become the mainstream in the moderate and low bit rate speech coding algorithm. The input signal of the algorithm process is a 16-bit linear PCM encoded speech signal by 8kHz sampling. The output rate is 8kbps of binary bit stream with a compression ratio of 16: 1 or so after the encoding processing. When algorithm DSP processed speech signal is used, the communication module need peripheral speech signal acquisition and necessary processing, such as signal amplification, A / D. Bitstream coded after storage and transmission can recover high-quality voice signal through the corresponding decoding process. Currently, G.729 codec has been widely used in various fields of data communications, such as H.323 and IP Phone.

Codec Program Design And Optimization

Codec Program Design

The object processed by the encoder is each 10ms frame of speech, which is 80 sampling points, and the average signal for each frame is divided into two sub-frames. Each frame signal is analyzed to extract the relevant model and excitation parameters, which are to be encoded. To reduce the computational overflow phenomenon, the amplitude of the input signal is processed by half with a cutoff frequency of 140Hz high-pass filter to filter out low frequency noise during the preprocessing phase. In linear prediction analysis stage, the autocorrelation coefficients should be calculated first, and then the 60Hz bandwidth expansion autocorrelation coefficient correction, and last linear prediction coefficients can be obtained using the resulting coefficients after Levinson - Durbin algorithm processing[1]. For quantization and interpolation, line spectrum pair should be transferred into prediction coefficients. The line spectrum pair coefficients are quantized and interpolated, and finally reduced to line spectrum pair coefficients from the linear prediction coefficients. Perceptual weighting is used to linear prediction coefficients without quantization. Analysis of the adaptive codebook search range is obtained by ring opening of genetic analysis to reduce the codebook search complexity[2]. Adaptive codebook and fixed codebook search are performed for each subframe. Then, the next subframe update the parameters of the synthesis filter and weighting filter should be used, and finally the resulting parameters could be encoded in a certain order. The decoding process is relatively simple. The parameter should be extracted firstly, the interpolated line spectral pair coefficients should be converted into linear prediction coefficients for each subframe;

and the adaptive codebook and fixed codebook should be multiplied by the excitation signal after it has been gained[3]. A linear predictive synthesis filter reconstructed speech could be made by the excitation signal, together with the completion of the speech after adaptive post-reconstruction filtering and high-pass filtering processing.

Optimization of Codec Program

In order to achieve real-time systems, series of optimization needed to be done for the source code G.729 from ITU, whose optimization degree can be divided into the algorithms level, C language level and assembler-level optimization[4]. The main algorithm-level optimization has the following work: First, cancel 5ms preview: In the LP analysis, the original algorithm which has the forward-looking 5ms data, all of which can be replaced with zeros in operation, bringing you a large reduction of computation; Second, the open-loop gene search using roughened search mode. In the original algorithm, when the correlation coefficient is calculated, the search is increased in steps of 2, saving half of the time, due to smaller changes of the continuous addition of speech data frames in pitch delay value. When changes decreased within a certain range, search is unnecessary, which could be directly replaced by the value of the previous frame. Third, in all the multiplication operation, the operation result of zero could be abandoned. Fourth, the fixed codebook search algorithm should be changed into the reset pulse sequence method: The first 40 possible pulse positions may be turned into equation (1) to calculate the contribution to the value of a single pulse, and then it should be reordered according to the size of the contribution to the value in the same track. The first track of the location of each of the four pulse Reset post could be selected to search for:

$$\frac{C_k^2}{E^2} = \frac{(\sum_{n=0}^{39} d(n)C_k(n))^2}{C_k^t \phi C_k} \quad (1)$$

The methods of C language level and assembler-level optimization are mainly as follows: Omitting unnecessary overflow judgment; the arrangement of the same nature with the same function, which will help the compiler to compile it into a code having parallel computing structure; when you call the cycle, the loop should be shortened as much as possible, and the transfer of the judge sentences should be avoided; fewer merge command functions, such as the autocorrelation function, windowing functions to save on the stack of the operating time[5]; The number of calls to the number of instructions are fewer functions, whose former name with a keyword inline, the compiler when comparing the province, which is a space for time optimization tools. In a PC, the algorithm is implemented in software platform CCS2.0, which could be used to some of this optimization method based on proven platforms, such as the use Intrinsic functions, the options open of C / C ++ compiler of CCS, Release-mode compilation, Debug information exclusion, which have a greater impact on system performance.

System Design

System block diagram is shown in Figure 1. Voice signal input has been made through the line or microphone, which generates 16-bit linear PCM by TLV320AIC23 chip pretreatment[6]. TLV320AIC23 chip is controlled by DSP through the on-chip MCBSP0, and exchanged data through the on-chip MCBSP1.

PC through the JTAG port of the device is programmed into the DSP, 16-bit PCM voice signal through MCBSP1 enter DSP, then processed by compression encoding algorithm to generate bit stream. When bit stream through MCBSP1 to an external communication system or module, voice could be restored at the decoding side. If it is to verify the correctness of the algorithm and reconstruction of voice quality distortion on a single DSP, after the bit stream decoding, the data could be sent back to TLV320AIC23 chip, and be reconstructed speech with D / A and its amplification.

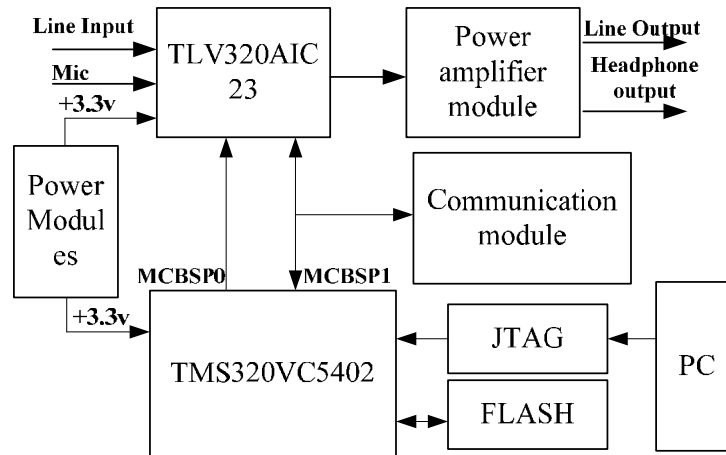


Fig.1. System block diagram

Hardware Design

Hardware design is mainly that of the interface between AIC23 and DSP. TLV320AIC23 ADC and DAC components of highly integrated in the chip, can provide 16bit, 20bit, 24bit and 32bit sampling in the frequency range of 8K ~ 96K. Voice signal can produce 16-bit linear PCM signal sampling rate of 8KHz after TLV320AIC23 through acquisition, providing an input signal processing algorithms in line. TLV320AIC23 pins can be divided into signal input and output pins, control pins, pin data transfer and power supply pins, etc., in which the signal input, output pins and the power pin connection are relatively simple, which could be completed in reference to typical circuit chip t materials. There are four control pins, respectively, SCLK, CS, SDIN, MODE, which are used to coordinate the host DSP initialization TLV320AIC23, there are five data transmission pins, namely BCLK, LRCIN, LRCOUT, DIN, DOUT, which are used with the host DSP voice data exchange. DSP's over six-channel buffered serial port pins can be divided into control pins and data pins. The control pin is the clock sending and receiving pins BCLKX, BCLKR, frame sending and receiving pins BFSX, BFSR, and the data pins is BDX, BDR. Pin connections between TLV320AIC23 and DSP design shown in Figure 2 and 3.

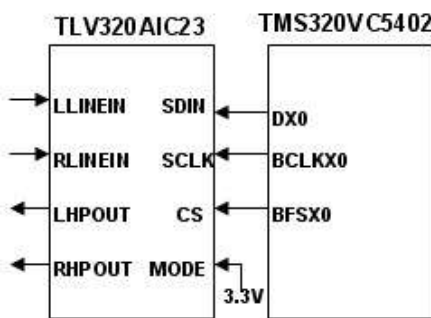


Fig.2. Control Interface

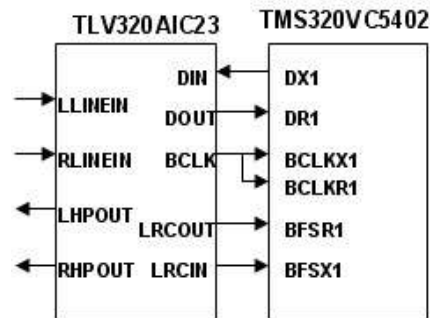


Fig.3. Data interface

Software Design

TLV320AIC23 chip initialization and work processes are implemented through software control, so are the DSP receiving, processing and transmitting data. The software design uses the modular, structured programming thought tradition. The program can be divided into five modules: DSP initialization, MCBSP0 communication with TLV320AIC23, MCBSP1 exchange data with TLV320AIC23, G.729 coding procedures, G.729 decoder. The overall software design process is shown in Figure 4:

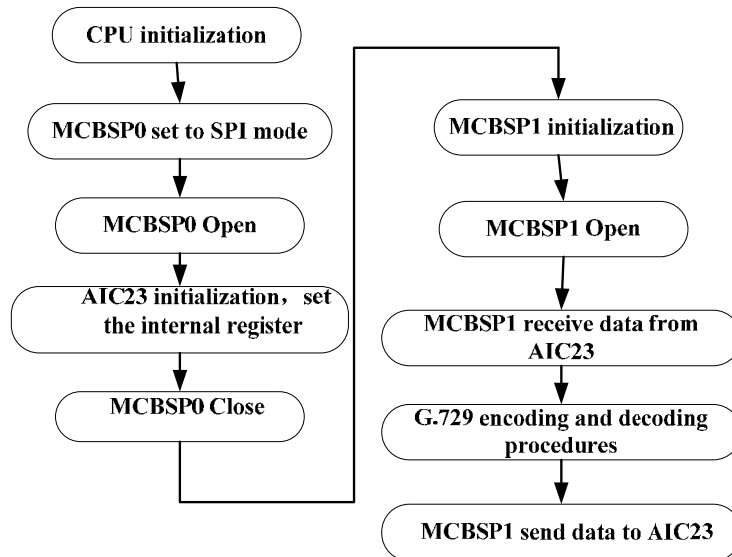


Figure4. Software design flow

Experimental Results

The optimized program is downloaded to the DSP through the emulator and JTAG, and then one's own voice recording could be input to the input interface of TLV320AIC23 through the PC line. The speech codec can be restored and reconstructed by the system. Compared with the original speech, only a very small distortion occurs. The original speech and the reconstructed speech waveform are shown in Figure 5 and 6. If a series of in-depth optimization could be implemented to the coding and decoding procedure and system, a better effect would be obtained.

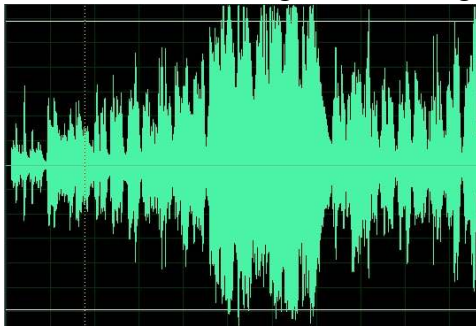


Fig.5. Original speech waveform

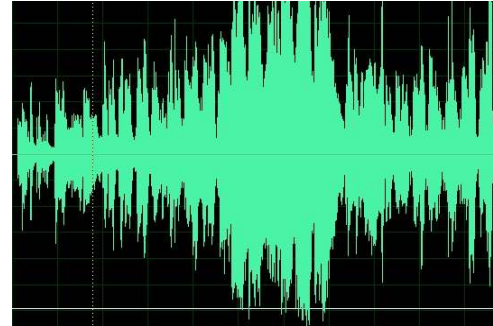


Fig.6. Reconstruction speech waveform

Acknowledgement

Fund Project: Hubei University of Science and Technology University's scientific research project(KY12042).

References

- [1] ITU-T Rec.G.729, "Coding of Speech at 8kbit/s using Conjugate Structure Algebraic Code Exited Linear Prediction(CS-ACELP) ,"March 1996.
- [2] M. Banerjee, B.A.Vani, and G.R.Krishna, "Optimalreal time DSP implementation of ITU G.729 speech codec," in Proc.IEEE 60th Vehicular Technology Conf. (VTC '04) vol. 6, pp.3908 3912, Sep. 2004.
- [3] Zhengwen Zhang,Geng Wang,Yongjie Zhang."Implementation of G.729 algorithm based on ARM9,"The International Conference on Consumer Electronics, Communications and Networks (CECNet 2012), China, 2012.

- [4] Tsai S-M, Yang J-F. Efficient algebraic code-excited liner predictive codebook search[J]. IEEE Proc Vis Image Signal Process, 2006, 153(6): 761. 768.
- [5] Hochong Park, Yunchang Choi, Doyoon Lee. Efficient codebook search method for ACELP speech codecs[C]. Speech Coding, IEEE Workshop Proceedings, 2002: 17-19.
- [6] TMS320VC5402 Data Sheet(Rev.F)Texas Instruments,February 2005.