Optimization and Realization of G.729 Protocol Based on ARM 9 Platform

Chunming Wu
Department of Information Engineering
Northeast Dianli University
Jilin, China
e-mail:wuhi1966@126.com

Bowen Zhang
Department of Information Engineering
Northeast Dianli University
Jilin, China
e-mail:814404286@qq.com

Yanxia Li
Department of hydraulic engineering
Hunan Technical College of Water Resourcesa and Hydro Power
Hunan, China
e-mail:lyx_spring@163.com

***Abstract*—In order to simplify the complexity of the algorithm structure of G.729 protocol and make the speech transmitted real-time on the ARM 9 platform, the present paper analyzes the G.729 protocol and designs the hard and software of the ARM 9 platform. Optimization and upgrading of Language C and algorithm based on the ARM 9 platform are conducted. The result of the research shows that the optimized complexity of the algorithm structure of G.729 protocol is obviously reduced and the requirements of real-time transmissions are met.**

Keywords- ARM9; G.729; speech coding; optimization

I. INTRODUCTION

Proposed by the ITU-T group, G.729 protocol enjoys a high level of speech quality and a low level of delayed speech coding. Its transmission rate stands at 8kb/s and the algorithm used is conjugate structure arithmetic code linear prediction (CS-ACELP). G.729 protocol is widely applied in various fields of data communication now[1]. However, since G.729 protocol enjoys a high degree of complexity in terms of algorithm, in most cases, the algorithm not optimized can not meet the requirements of speech real-time transmission. It is based on this consideration that the present paper conducts an in-dept analysis of the G.729 protocol and optimizes its algorithm, whose purpose is to make it possible for speech to be transmitted in a real-time manner on the ARM 9 platform.

II. EN-AND DECODING PRINCIPLES OF THE G.729

As can be seen from figure 1, the input signal is pre-processed by high-pass filtering. A linear prediction (LP) of a speech frame whose running time is 10ms is analyzed. Factors like the LP filter co-efficient calculated are transferred into line spectrum pairs (LSP). Quantization of 18bit is conducted by prediction quantity and vector quantization (VQ). This encoding machine is mainly based on the analysis synthesis, which can reduce the energy of differential signal between the combined voice and the actual voice when hearing is weighted. Adaptive algorithm is used when considering the weighting filter in order to improve the input signals[2].Every one frame (5ms, 40 samples) of incentive parameters (adaptive and fixed coding parameters) is conducted.The unquantified and quantified LP filters co-efficient is applied into the second frame, whereas the inserted LP co-efficient is applied into the first. An estimation is made on the delay of open loop genes every 10ms based on the perceived speech signals. Practice follows has to be repeated on every frame. The target signal $x(n)$ is got after reducing the weighting filter $W(z)/A(z)$.The initial stage of the filter is modified by LP residual and error between incentives, which is equivalent to cancel the zero input response of weighting filter from the weighting speech signals. The impulse response of weighting filter $h(n)$ is calculated and then with the $h(n)$ and $x(n)$ as the method, explores the set numerical nearby the open loop genes and analyzes closed loop genes, which is to search for gains and delays of adaptive coding book. Resolution ratio of factors genes delays is at a sample interval of 1/3. Encoding of genes delay is conducted at 8bit in the first frame, whereas it is at 5bit in the second frame. Thus, the contribution of adaptive coding book is decreased from the target signals. The new target signal $x'(n)$ is searched and employed by the fixed coding book, the encoding of which is conducted by the book at 17bit. Combined quantization at 7bit is employed in gains of fixed coding book and adaptive coding book. Finally, fixed incentive signal is adopted when modifying filter memory[3].

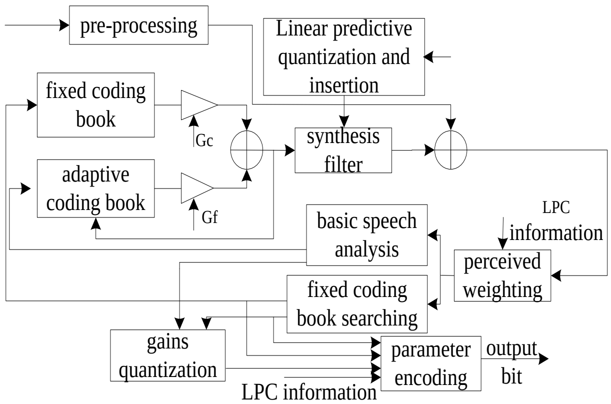


Figure 1 Algorithm Structure of G.729 Protocol

Figure 2 shows the principles of G.729 speech encoding machine. First, the serial numbers of parameters are collected from the received coding. The corresponding encoding parameters of 10ms speech frames are got by decoding these serial numbers. These encoding parameters are LSP parameters, consisting of two fixed coding book vectors and two factors genes delays and two groups of adaptive coding. Every frame of LSP parameter is inserted and converted into LP filters co-efficient, and then combines into the speech at 5ms per unit. Steps adopted are as follows:

Fixed coding book and adaptive coding book multiply their respective gains, and then are added to make incentives.

Speech is re-organized by the incentive LP combination.

The re-organized speech signal is post positioned, including short-term combination filter, long-term post positioned filter and high-pass filtering[4].

The analysis and processing of CS-ACELP encoding machine is carried out by speech signal at 10ms per unit.4 Apart from that, since a future frame is needed during encoding, the total delayed time is 15ms. In actual practice, the total delayed time should include: (1) the transmission time in connecting communications. (2) the processing time of en-and decoding. (3) the combined time delayed of combined speech data and other data.

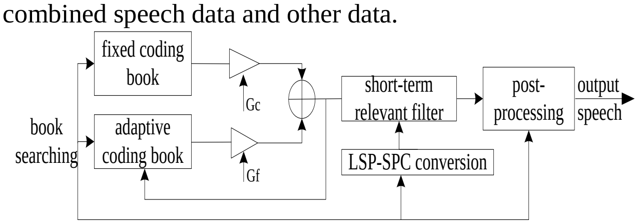


Figure 2 Decoding Algorithm Structure of G.729 Protocol

III. Realization of the Encoding Machine

A. *Software Realization of the Encoding Machine*

As is illustrated by figure 3, procedures and call relations are introduced by the main program. Channel coding function { prm2bits-ld8k ( )}is completed in the process of developing of drivers, which can be directly adjusted and called in the main program. Areas of coding and data are initialized in the first place by the main program. Stored data of the input signals in the buffer are collected. The function of prm 2bits-ld8k in the bits.c program is used to encode the signals, receiving bit stream.

These functions and documents are called in different phases in the running of the program respectively. This branch of speech processing system uses main.c as the main program document.

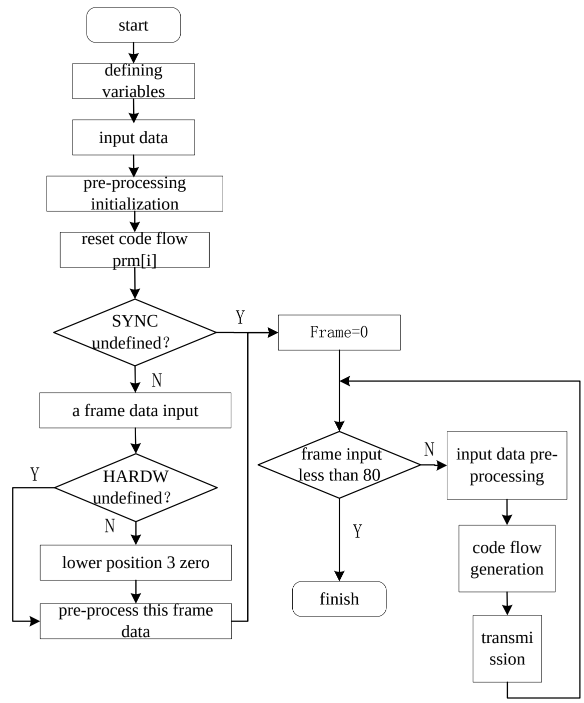


Figure 3 Flowchart of the Main Program

B. *Hardware Realization of the Encoding Machine*

Hardware is shown by figure 4, through which MMU,AMBA BUS and Havard high-speed buffer systems can be realized. A CPU of Samsung S3C2440 ARM920T is the core controller, with a basic frequency of 400MHZ, and 533MHZ being the highest. The system is equipped with a nand flash(64M) and two half-word SDRAM system, thus ensuring an improved communicating efficiency of CUP. Extra consumption of S3C2440A is reduced by employing a whole set of peripheral devices of communicating system[5]. UDA1341, developed by Philips Corporation is adopted as the audio processing machine.

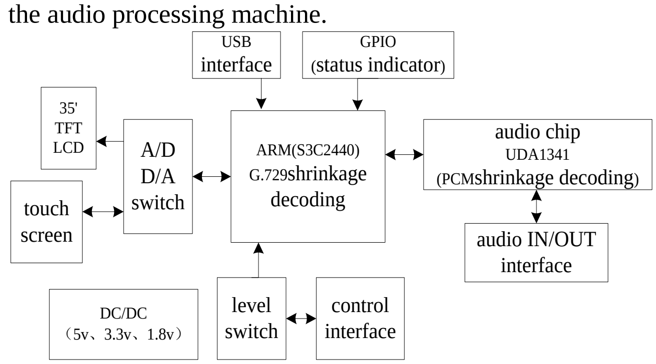


Figure 4 Structure Chart of Hardware

UDA1341 supports the IIS bus data format and adopts bit stream conversion technology to conduct signal processing and completes the analog-to-digital conversion of speech signals. It is equipped with programmable gain amplifier and digital automatic gain control unit and enjoys features of low voltage and low energy consumption. Two groups of audio signal input sockets with left and right sound tracks each are provided by UDA1341. The IIS bus data format is supported. There are three working modes of IIS bus, namely, DMA, normal transmission and transmission and reception[6].

C. Overall Design of the System

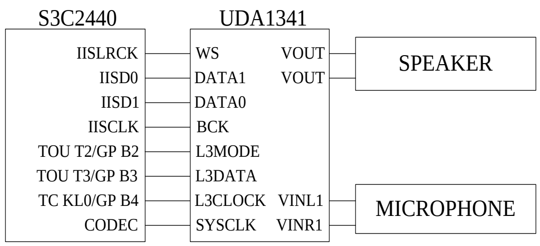


Figure 5 Hardware Connection Chart

As can be seen from figure 5, IISLRCK signal is connected to the pins of WS in the S3C2440 chip. BCK pin of UDA1341 is connected with IISCLK pin of the clock signal. The audio chip of UDA1341 has two audio ways. One is used to output, the other is input. The corresponding connections are as follows. IISD1, the audio input pin of S3C2440 is connected with DATA0, the audio output pin of UDA1341. IISD0, the audio input pin of S3C2440 is connected with DATA1, the audio input pin of UDA1341. The corresponding pins of IIS of UDA1341 are connected with the corresponding pins of IIS of UDA1341. VOUT, pin of audio output is connected with speaker, and VIN, pin of audio input is connected with microphone. L3 can be taken as a Mixer control interface, which can be used to adjust volume and bass[7].

IV. ALGORITHM OPTIMIZATION

A. Optimization at the Language C Level

Standard G.729 protocol defines a complete 16 bit and 32 bit arithmetic and relevant advanced algorithm in documents such as basop32.c. In the actual practice, however, the entire plus and minus are calculated to prevent data from being exposed. Besides, this program also defines an overall variable overflow which is used to mark whether saturation occurs or not. If saturation occurs, the variable will be posted. Efficiency of en-and decoding is severely affected by these factors. Indicators of jumping are increased as a result of the large quantities of if sentences in the system. As such, fluency of streamlines is severely affected and running time of program is increased, leading to a fact that jumping indicator is the most time consuming indicator in the arm system. Therefore, not only are unnecessary amount of coding is brought in but also the efficiency of en-and decoding is reduced to a large extent because of the saturation calculation. So, it is of vital importance to optimize these frequently-called C programs and to delete the saturation of these calculations[8].

B. Optimization at the Algorithm Level

In most cases, the optimization of the C language can not meet the requirements of real-time transmissions. As is known to all, the high degree of complexity that G.729 algorithm enjoys is the major reason for the inability of speech real-time transmissions. The present paper puts forward an optimized method targeted at the highly-complicated G.729 speech algorithm, which simplifies the quantified gains structure. The multiplied figure zero in the simplified structure is simplified into plus, decreasing the algorithm complexity[9].

In the gains quantization process, every single frame should be calculated by $h(n)$ perceived weighting combined impulse response of filter $W(z)/\hat{A}(z)$ and $c(n)$ fixed coding book vector. $z(n)$ is got from multiplying them. The formula of $z(n)$ is as follows[10]:

$$z(n)=c(n)*h(n)=\sum_{i=0}^{n}c(i)h(n-i), n=0,\cdots\cdots,39 \quad (1)$$

Calculate (1) directly. 780 times of multiplying and added are required.

By analyzing G.729 algorithm, the formula of $c(n)$ can be got as:

$$c(n)=\sum_{k=0}^{3}b_k\delta(n-m_k), n=0,\cdots\cdots,39 \quad (2)$$

There are only four non-zero impulse in $c(n)$ of formula (2), with each range of impulse +1 or -1.

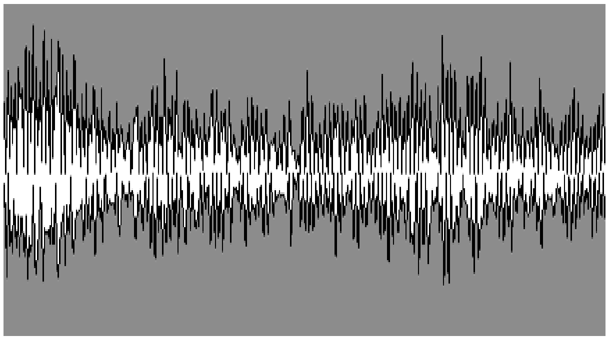
Thus, the formula of $z(n)$ can be turned into

$$\begin{aligned} z(n)=c(n)*h(n)&=\sum_{i=0}^{n}h(n-i)\sum_{k=0}^{3}b_k\delta(i-m_k) \\ &=\sum_{k=0}^{3}b_k h(n-m_k), n=0,\cdots\cdots,39 \end{aligned} \quad (3)$$

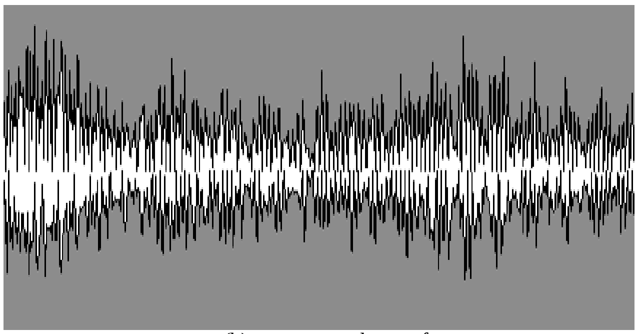
Again, since b_k takes +1 or -1, only 160 times of plus and minus are needed when calculating $z(n)$. Algorithm is simplified in this way, decreasing the complexity of G.729 algorithm.

V. SIMULATED RESULTS

Figure 6(a) and figure 6(b) are speech waveform figures of input and output respectively. From the two charts, it can be seen that by en-and decoding in the above-mentioned manners, there is no obvious distortion of the output speech waveform, completely meeting the requirements of hearing. Through calculating running time of the whole system, it can be found that processing time of one frame on average is decreased from 17.45ms to 8.36ms, which is less than one frame speech length of G.729, whose length is 10ms.



(a) input speech waveform



(b) output speech waveform
Figure 6 Simulated Speech Waveforms

VI. CONCLUSION

The present paper conducts an in-depth analysis of G.729 protocol and its realization based on ARM 9 platform. By combining the high degree of complexity of linux and G.729 speech encoding algorithm, G.729 algorithm is optimized and upgraded at the language C level and algorithm level respectively. This method of algorithm is tested in the S3C2440 processing machine. Simulated results demonstrate that after optimizing G.729 condensed encoding algorithm, speech quality satisfies the requirements and operating efficiency greatly improves, meeting the requirements of real-time transmissions.

ACKNOWLEDGMENT

This work was supported by Science and Technology Development Plan of Jilin Province (20130206050GX).

REFERENCES

[1] Xu Lihong, Kan Haiying, Yu Xiaoqing. G.729 CS-ACELP Optimization and Realization of Speech Encoding Algorithm. Journal of Shanghai University, vol.7(1),pp.13-17,Feb.2001.

[2] ITU-T Recommendation G.729 Coding of speech at 8kbit/s using conjugate structure algebraic-code-excited linear-prediction(CS-ACELP),Nov.1996.

[3] Li Zhiyuan, Li Fu. Prospects of Speech Technology Applied in IT Industry. Modern Intelligence, 2003,4, pp.60-61.

[4] Zhang Jidong, Yang Zhen, Li Xiaofei. ITU-T G.729 CS-ACELP Speech Encoding System Functional Analysis. Journal of Nanjing University of Posts and Telecommunications. Nature and Science, vol.20(4),pp.91-94,2000.

[5] Yu Ming, Fan Shurui, Zeng Xiangye. ARM9 Inserted System Design and Development Course. Beijing: Publishing House of Electronics Industry, pp.80-127,2006.

[6] PHILIPS Corporation. UDA1341TS_datasheet.pdf,2002.

[7] Tian Ze. ARM9 Inserted Development :Experiment and Practice. Beijing: Beijing University of Aeronautics and Astronautics Press, 2006.

[8] Bao Changchun,Dou Gengxin.Fast Codebook Search Algorithm Based on ACEL.Journal of Beijing University for Industry, vol.33(8),pp.824-827,2007.

[9] Wang Wei, Li Yonglu. Classification, Application and Development of Speech Encoding Technology. Communication Net Technology,vol.29(4),pp.38-40,2003.

[10] Zou Huihua,Li Shuangtian. Variable Rate Speech Encoding Algorithm Analysis: Based on Frequency Slope Constraints.Signal Processing, Vol.17,Supplementary issue, pp.193-198,2001.