# Implementation of G.729E Speech Coding Algorithm based on TMS320VC5416

YANG Xiaojin[1, a], PAN Jinjin[2,b]

[1]Department of Information Engineering, Guilin University of Aerospace Technology, Guilin, China

[2]Department of Electronic Engineering, Guilin University of Aerospace Technology, Guilin, China

[a] yxj@guat.edu.cn, [b] jphan@guat.edu.cn

**Keywords:** Speech Coding G.729E DSP

**Abstract.** Implementation of G.729E Speech Coding Algorithm based on TMS320VC5416 G.729E algorithm is an excellent speech coding algorithm. G.729E designed to speech with background noise and even music, it will be widely used in multimedia communication. This paper introduce the base algorithm, and it is emphasized the hardware and software implementation in TMS320VC5416DSP. It is offered the testing result on hardware.

## Introduction

In recent years, the Internet got rapid development, with the development of IP technology convergence network research framework and VOIP technology, data communication network has been integrated into the traditional voice business. But uncompressed speech signal real time data volume is very large, e.g. A voice signals,   at 8 kHz samples, via the A/D conversion, and with 16 bit PCM coding, will produce 128 kbps amount of data . Considering the shortage of channel resources, with the development of IP telephony and video phone, it is hoped that the speech coding algorithm have as low as possible rate yet smaller delay. So appears the G. 729 Algorithm, which has those features. Meanwhile with the development of multimedia technology, the speech coding in multimedia application is becoming increasingly important, and, is put forward higher requirements as well. In November 1998, the ITU - T set the G. 729E standard-- a high rate   base on the enhanced version of G.729 , it has   higher voice quality than G.729 . And with the particular voice signal processing algorithms to the music signal and background noise, G729E is more suitable for conference calls, VOIP and other multimedia communication,   and its coding rate is 11.8 kbps which is only slightly higher than the G.729.

## G.729E Speech Coding and Decoding Algorithm

### G.729E Speech Coding Algorithm

G.729E uses the conjugate structure-Algebraic Code Excited Linear Prediction (CS - ACELP), the selected frame length is 10 ms (sampling rate is 8000 Hz/s, which means 80 sample points per frame). For every 10 ms frame data, the encoder gives a 10 order to LP forward analysis and a 30 order to LP backward analysis to calculate the LP filter coefficient. For the LP forward analysis, it is needed to convert coefficient of LP to imitate the LSP parameters, and to carries on the vector quantization. By comparing the prediction gain of forward and backward insertion and parameters of the previous frame data , we can determine to adapt which prediction, forward or backward. and then use the Analysis Synthesis Method to extract the excitation parameters according to the perceptual distortion minimum weighted criteria between the synthesis signal and the original signal .

The calculation of the Excitation parameters (fixed code and adaptive code) is conducted according to the sub frame, namely every 5 ms calculation, and the calculated LP coefficient used for the second sub frame, and the first sub frame uses the interpolated coefficient of LP. For backward LP filter coefficients, if the number of frames is less than 10 frames   after using the LP analysis , then the calculated LP coefficient needs to be smoothed, namely press type calculation:

(Abwd means the backward LP prediction coefficients after direct calculation, Afwd the forward LP prediction coefficients.)

$$A(n+i) = 0.1*i*A_{bwd}(n+i) + (1-0.1*i)*A(n+i-1), 0 \le i \le 9 \quad (1)$$

$$A(n+i) = A_{bwd}(n+i), \quad i \ge 10 \quad (2)$$

$$A(n-1) = A_{fwd}(n-1) \quad (3)$$

After perceptual weighting, each sub-frame data should be conducted open and close pitch analysis to determine the delay and gain of the adaptive code . Different LP analysis will have different perceptual weighting, the perceptual weighting coefficient is based on the smoothness of signal .  As for the forward LP model, when the signal of a smooth degree is low, inverse filter constituted of  LP coefficient which un-quantification; when it's high, use quantification . For the backward LP model, when the signal of the smooth degree is low, un-quantitative LP coefficient is adopted to form inverse filter, the weighted factorγ1＝0.9 and γ2=0.4, or directly use the backward filter, the perceptual weighting factor  γ1＝0.98 andγ2＝0.4. G. 729E has large capacity of codebook and the number of pulses in a code word number is also very large(the number of pulses forward mode for 10, after the number of pulses to the pattern of 12). And the pulse of search algorithm is used for its suboptimal search algorithm.

**The Decoding Algorithm**

Figure 1 is a codec principle block diagram. G. 729E is decoded by frame, after decoding the standard code stream   which meets the standard of G. 729E, the corresponding parameters are get, and then generates the synthesized speech according to the voice .
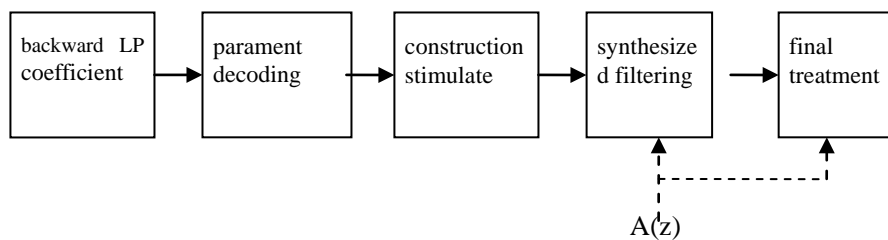


Fig.1 G.729E codec principle block diagram

For decoding part, no matter the decoding frame is forward LP model or backward LP model, always backward LP analysis goes first, and then decode parameters are carried respectively , based on different LP models, then goes the computing reconstruct speech, finally the post-processing, which includes adaptive filtering, high-pass filter and signal amplification. .When the signal to frame after the high for 20 consecutive frames is smooth, do not need to undertake after filtering.

## Algorithm Implementation

### The hardware design

System block diagram is shown in figure 2. Working principle of the system is speech signal is sent through the microphone into PCM3002 stereo AD/DA converter, then digital signal is concerted and sent to VC5416DSP storage. There should be an interrupt after storing every frame signal, during which, speech coding algorithm is used to give compressing processing and data storage, and then send the data. On the other hand it can be compressed data decompression process, that is, when receiving a frame of voice compression data, an interrupt appears, call voice decoding algorithm to extract the data and then send the data to the PCM3002 stereo AD/DA converter through the DSP serial port and then ,after getting the reduction of speech signal, the reconstruction of voice can be shown through headphones and other equipment. Interface design of the circuit is shown in figure 3,it shows the interface between VC5416DSP and the other chips.
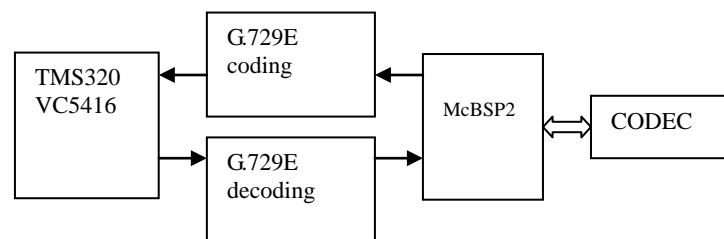
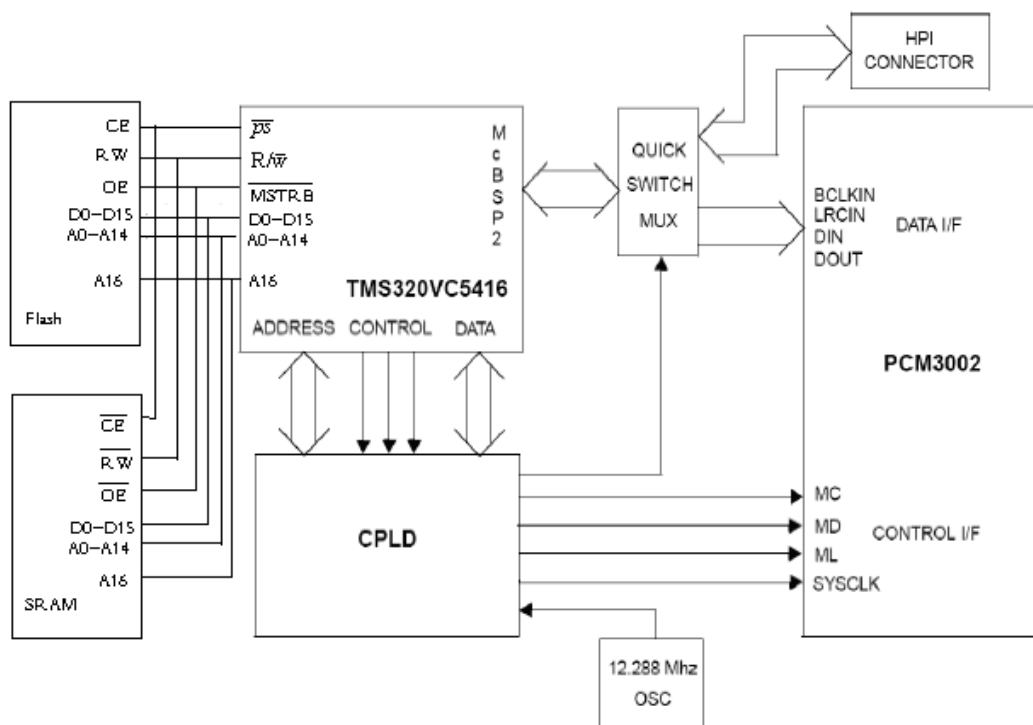Fig. 2 system chart of G.729E algorithm implemen tation

Fig. 3 interface between VC5416DSP and the other chips

### The algorithm implementation and optimization

Press with CCS (Code Composer Studio) built-in DSP/BIOS, to establish a DSP/BIOS configuration file accordingly,, and add it to the project file. Software implementation scheme is to establish real-time voice and data channels by firstly using the PIP function and then by using interrupt function to call voice coding or decoding procedures to decode arithmetic. The main program is really just completed some initialization Settings, and then wait for the interrupt.

Interrupt program is the core of the whole decoding arithmetic. The DSP implementation of G. 729 E speech coding algorithm can be divided into real-time voice data acquisition, speech coding, speech decoding and speech output several modules, among which, the speech coding is the key part. DSP will interrupt once at every collected 80 sample points, calling code program for processing. Coding procedure for data preprocessing goes first, and then come the forward and backward LPC arithmetic respectively. After calculating the corresponding LSF coefficient of forward LPC coefficients,, and then determine which LPC should be used to process the frame data, if it is the forward LPC, then the LSF coefficient should be undertaken quantitative interpolation processing, or directly just interpolation to the backward LPC. In order to improve the quality of voice, In order to improve the quality of voice, obtain the LPC coefficients of perceptual weighting processing after reconstructing speech signal, and get the vector by using the adaptive code search, and finally encode the transmission parameters after storage.

Based on the the standard C code of G. 729 E provided by ITU, the test shows that using a 10 ms VC5416DSP coding frame needed more time than 10 ms and it is unable to realize real-time decoding arithmetic. Therefore in order to improve the code operation efficiency , and the small footprint assembler code is chosen, which means using C and assembly mixed encoding to optimize the program. Program optimization is to repeatedly call the basic computing unit and to optimize the nested loop.

Because the implementation of the standard C code in the process of programming is rigorous fixed-point number operation, thus low running rate will be caused. For example, in a process , repeatedly called such as the rank of mattrix L_mac, L_mult, L_shr can use an assembly instruction to complete in only one operating period, while  the C program needs dozens of operation cycles ;On the other hand, G. 729E algorithm using a large number of filter and vector quantization functions, it needs great amount of computation. The main operation of these functions are nested loop, in the study by optimizing nested loop part of the program, or in assembly language implementation, its computational speed is one hundred times higher than C program.


**Hybrid coding test and results**

After completing the optimization rewritten of the G.729E algorithm , the test results of the code is as follows: in CCS2.0 continuously static voice file decoding 5000 frames , each frame coding takes an average of 201,000 clock cycles, and decoding 1 frame takes up 51000 clock cycles. C5416, working in a 160 MHZ, can finish 10 ms speech codec in 1.6 ms. And can realize real -time voice encoding and decoding arithmetic on C5416 system . Figure 4 is a frame of data with the original voice and data after decoding the spectrogram. The left-hand is a broadband spectrum diagram of female voice with background music, and the right-hand is the data after decoding.
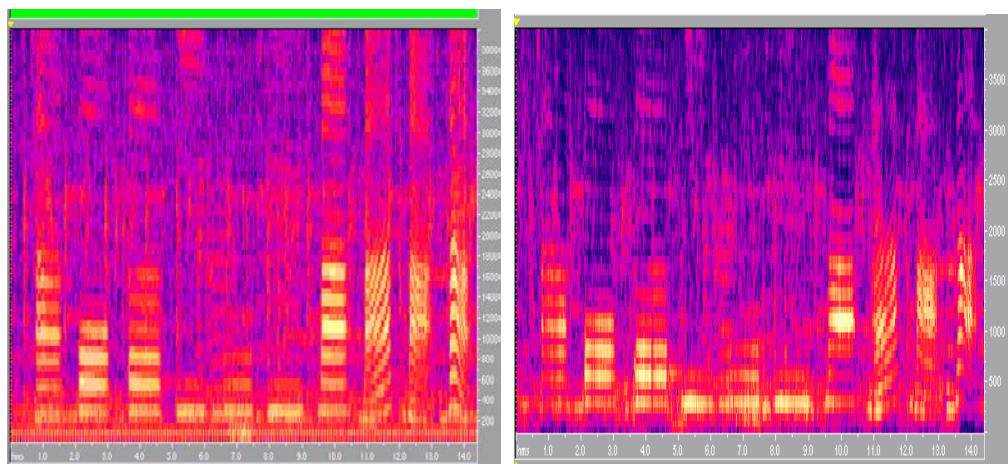


Fig. 4 a frame of data with the original voice and data after decoding the spectrogram

From broadband spectrum diagram, it can be seen that the decoding speech and the original speech have almost the same resonance peak. By subjective listening, whether with background noise, voice or music signal, the decoding voice has better quality .

## Conclusion

G. 729E speech coding algorithm particularly by adding voice processing with background noise or music signal, improves its voice quality, and it will be widely used in multimedia communication. In this paper, the study of its real-time implementation in TMS320VC5416DSP has practical significance; it can be used for speech processing in VOIP, conference calls and etc.

## References

[1] ITU－T　Recommendation G.729 Annex E: 11.8kbit/s CS-ACELP speech coding algorithm[S]. 1998

[2] J.H.James, Bing Chen.Voice over IP and Quality of Service [J]. IEEE Communications Magazine July 2004

[3] Texas Instrument.TMS320VC5416 Fixed-Point Digital Signal Processor Data Manual[S]. 2005

[4] Texas Instrument.TMS320VC54X DSPReference Ser.Volume2: Mnemonic Instruction Set [S].1998

[5] Yang xiaojin，Liu jianhua. The Implementation of IKEV2 for IPSec [J]. Micro computer information ，2006，8-3：55-57