Implementation of low bit rate Vocoder for speech compression

Lavanya Krishna[1], Mrs. Rajeswari P [2]

[1]PG student, Dept. of Telecommunication Engineering, DSCE, Karnataka, India
[2]Associate Professor, Dept. of Telecommunication Engineering, DSCE, Karnataka, India

---

Abstract - *Compression of speech signal is an important field in digital signal processing. Because of limited bandwidth in many fields especially in the field of military speech compression has a significant importance in the present world. The other reasons for speech compression are limited transmission and storage capacity. The process of converting human speech signals into encoded representation then converting back into original signal by decoding back to produce a close approximation of the original signal. This paper presents a speech compression by designing a low bit rate Vocoder board. The process includes component selection for doing schematic followed by PCB design. The final board is tested by giving speech input of different languages and then calculating PESQ of both input and compressed speech*

***Key Words*: Vocoder, PESQ, Low bit rate, Codec, UART.**

1. INTRODUCTION

According to information theory, the minimum bitrate at which the conditions of distortionless transmission of any source signal is possible is determined by the entropy of the speech source message. The compression after the maximum level compression results in distortion and loss of signal. Various speech encoding techniques includes LPC, CELP, MELP and TWELP. Compression of speech signal results in low bit rate data which reduces the bandwidth required for transmission. Implementing better efficient compression techniques results in both quality and LBR data. Encoding, decoding and compression of speech signal can be done using VOCODER. VOCODER can be configured either by hardware or by software. In hardware configuration jumpers are used to fix the voltage for configurable pins, where as in software configurations any of the processors or the controllers can be used to configure the pins. The Blackfin device BF548 can be used to configure and also to read and write the speech signals from and to the VOCODER respectively. On top of that the read speech signal can be encrypted. The encryption of compressed speech signal is the main requirement.

A large part of the researches in speech process algorithms is motivated by the need of obtaining secure military communications, to allow effective operation in a

Hostile environment. Since the bandwidth of the communication channel is a sensitive problem in military applications, low bit-rate speech compression methods are used. Several speech processing applications such as Mixed Excitation Linear Prediction are characterized by very strict requirements in power consumption, size, and voltage supply. These requirements are difficult to fulfill, given the complexity and number of functions to be implemented, together with the real time requirement and large dynamic range of the input signals. To meet these constraints, careful optimization should be done at all levels, ranging from algorithmic level, through system and circuit architecture, to layout and design of the cell library. The key points of this optimization are among others, the choice of the algorithms, the modification of the algorithms to reduce computational complexity, the choice of a fixed-point arithmetic unit, the minimization of the number of bits required at every node of the algorithm, and a careful match between algorithms and architecture. This paper concentrates on low bit rate speech coding technology, mainly in TWELP and solved the problem of optimizing the program of TWELP on Digital Signal Processor platform. The algorithm was ported onto a fixed point DSP, Blackfin 537, and stage by stage optimization was performed to meet the real time requirements. The main functions involved were analysis, parameter encoding, parameter decoding and synthesis. The fixed point source code at the TWELP front end was also thoroughly optimized at the C Level. Memory optimization techniques such as data placement and caching were also used to reduce the processing time. The results were obtained show that real-time implementations of a speech Vocoder based on the TWELP standard for low bit rate communications (2400 bps) can be successful on DSP platforms.

1.1 PIN Assignment

The pin diagram of Vocoder chip is as shown in the figure. This is a full-multiplex Vocoder chip. It has built-in FLASH and RAM and can do real-time speech encoding/decoding with single chip, no need for external storage, which decreases the complexity for customer to design their systems. VOCODER supports 600bps, 1200bps and 2400bps coding rate, which is configurable with pins. VOCODER can integrate the codec AD73311 seamlessly and configure it when powering up without user involving. It connects to MCU with UART. User can read and write the speech data using UART and the process is asynchronous and full-duplex. This document describes a demo board, which is used to demo VOCODER chip's external circuits and to show the encoding/decoding effect. The demo board provides a simple reference design. User can follow this board to design the speech encoding/decoding circuit in their specific product.

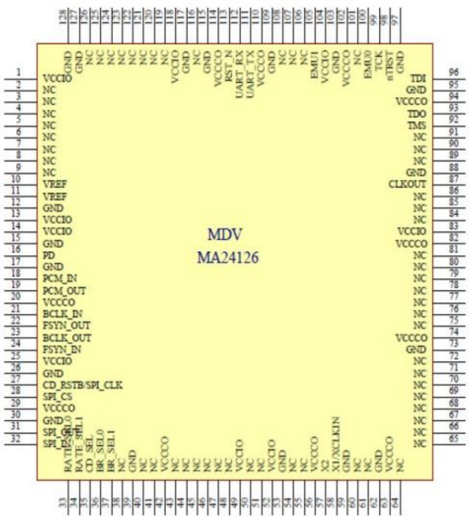

Figure 1: PIN Asignment

1.2 Block Diagram

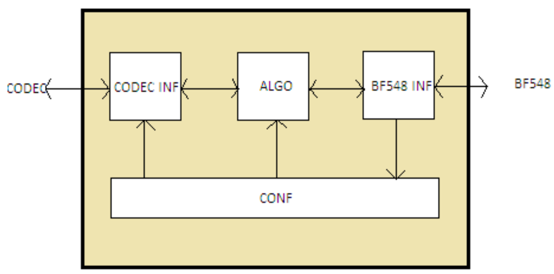

Figure 2: Block Diagram

The above block diagram represents the functional overview of the Vocoder. Each block functions are described below.

1.2.1 Algorithm block:

The function of algorithm block is to implement the functions related to encoding/decoding algorithm. This block is the core module of Vocoder chip. During encoding the algorithm block receive speech data from the codec interface block then compress and encode the data and then send to the BF548 interface clock to transmit. During decoding the algorithm block receive data from the BF548 block, decodes the data and then send to the codec interface module to playback. It consists of codec interface block, ADSP BF548 interface block, algorithm block and configuration block.

1.2.2 CODEC interface block:

The codec interface block connects to the external codec to which the speech data to be compressed has to be send. During encoding the codec interface block receives the speech data from external codec and then sends it to an algorithm block to do compression encoding. During decoding the codec interface block receives the decoded speech data from algorithm block and then send it back to the external codec to play

1.2.3 BF548 interface block:

The BF548 block connects to the external BF548 and is used to transport encoded/decoded data and also the configuration data. During encoding the BF548 block receives data from algorithm block frame them and send to external BF548 unit. During decoding, the BF548 interface block receives speech data frame from external BF548 unit decode the frames and then send to algorithm block for speech data encoding. During configuring BF548 interface block receives configuring data frame from external BF548, decode the frames and then send to the configuration block for parsing and configuring. The communication between BF548 and configuration block is full duplex therefore the configuration data, encoded data and decoded data can be send simultaneously

1.2.4 Configuration block:

This block configure the chip function according to configure pin status or external configure data. When powered up, configure block samples the configure pins' status to configure the chip accordingly. When in operating, the configure block accepts the data from BF548 interface block BF548 and configure related blocks after parsing the data.

2. Configuration

Vocoder chip can be configured in hardware or software method. One hand, the chip samples the voltage of the configure pins to finish the configuration. On the other hand, user can configure it via software protocol when running.

Rate selection

Vocoder chip works at three rates: 2400bps, 1200bps and 600bps. The pins are illustrated in the table below:

Table 1: Rate selection

RATE_SEL1	RATE_SEL0	Coding Rate
0	0	2400bps
0	1	1200bps
1	0	600bps
1	1	Codec loop

Codec Loop Mode means skipping the processing of coding and decoding with playing back the speech data directly. It is used to test or debug. Rate selection can also be configured with software protocol.

Codec Selection

With external pin selection, Vocoder chip can connect to AD73311 and AIC23 seamless.

The pin CD_SEL is given as below:

Table 2: CODEC selection pins

CD_SEL	Function
0 Select	AD73311
1 Select	AIC23

MCU Interface Rate Selection (Baud rate):

The chip connect to external MCU via asynchronous UART and the Baud rate can be selected using the pins as following

Table 3: Baud Rate selection pins definition

BR_SEL1	BR_SEL0 Baud	Rate
0	1	15200bps
0	1	9600bps
1	0	4800bps
1	1	2400bps

Baud rate of the serial port can be configured with software protocol.

3. INTERFACES

The three interfaces in this set are

MCU Interface

VOCODER connects with MCU using UART port. The speed can be selected by hardware or software. The interface pin includes UART_TX and UART_RX. The port time sequence adopts the standard UART time sequence.

Codec Interface

VOCODER support kinds of Codec, which is selectable by hardware or software. The interface pins include: BCLK_IN, FSYN_IN, PCM_IN, BCLK_OUT, FSYN_OUT, and PCM_OUT. The port time sequence can be configured by software protocol.

Configure Interface

When VOCODER is powered up, it sample the configure pins to configure the working mode. It can also be configured with software protocol when running. The configure pins include: coding rate select pins (RATE_SEL0, RATE_SEL1), Codec select pins (CD_SEL), MCU port speed select pins (BR_SEL0, BR_SEL1).

The communication protocol between BF548 unit and internal MCU with configuration block is UART protocol.

A UART (Universal Asynchronous Receiver and Transmitter) is a device allowing the reception and transmission of information, in a serial and asynchronous way.

- A UART allows the communication between a computer and several kinds of devices (printer, modem, etc.), interconnected via an RS-232 cable. Data transmission is made by the UART in a serial way, by 11-bit blocks:

- A 0 bit marks the starting point of the block
- Eight bits for data
- One parity bit
- A 1 bit marking the end of the block
- The transmission and reception lines should hold a 1 when no data is transmitted.
- The first transmitted bit is start bit data parity bit stop bit
- The first transmitted bit is the LSB (least significant bit)
- The parity bit is set to 1 or 0, depending on the number of 1's transmitted: if even parity is used, this number should be even; if odd parity is used, this number should be odd. If the chosen parity is not respected in the block, a transmission error should be detected
- The transmission speed is fixed, measured in bauds.

4. Frame structure

The frame length is fixed as 16 Byte and the frame structure is as below: following are the fields in the frame

HEADER	CMD_TYPE	LEN	PAYLOAD	CHECKSUM

(1) HEADER

Frame head, 2 bytes length. The content is fixed as 0x4C4E

(2) CMD_TYPE

The command type, 1 byte length.

(3) LEN

The payload length, 1 byte length.

(4) PAYLOAD

The payload data, 11 bytes length.

(5) CHECKSUM

Checksum is a 1 byte length. Add the first 15 bytes in a frame (it means the total frame excluding checksum itself) and get the low 8 bits of the sum as the checksum.

5. APPLICATIONS

- Under water acoustic communication
- Mobile communication
- Satellite communication
- Secret communication
- HF communication
- Embedded speech data storage
- Digital mobile radio station

6. RESULTS

Work done till now

The TWELP Vocoder board supporting the speech compression at three coding rates 600bps, 1200bps and 2400bps is developed with a special feature of interfacing it with ADSP BF548 and compression of speech in English is done. Figure 8.1 and 8.2 shows the raw speech signal and the compressed one.

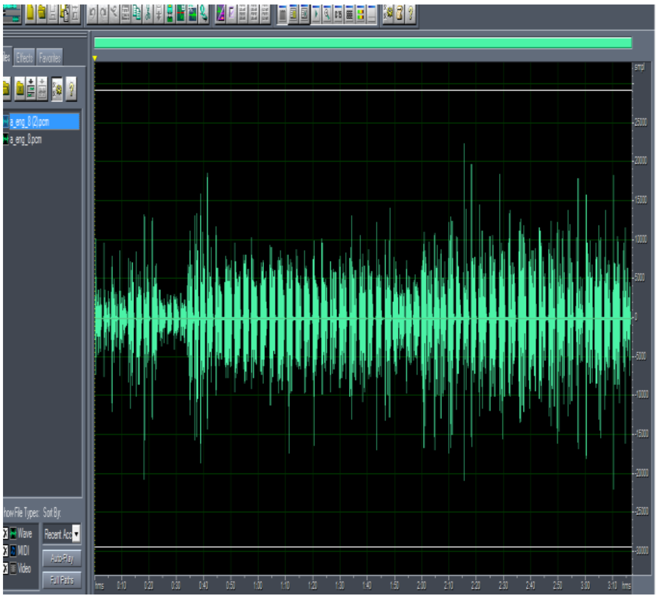

Figure 3: Input (raw) speech signal – English

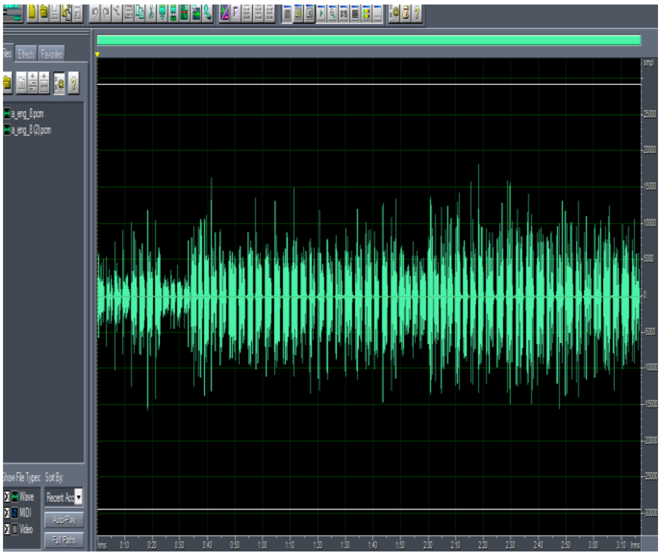

Figure 4: Output (compressed-1200bps) speech signal – English

PESQ result

PESQ result for the English speech signal is obtained VEC\INP\P50\b_eng_8.pcm VEC\OUT_1\b_eng_8.pcm 2.826 2.569

7. Future work

The compression of speech in different other languages has to perform. The PESQ results for input (raw) and output (compressed) has to find out to compare the quality of compressed speech with original speech. The same thing has to be performed at the rate of 2400 bps.

REFERENCE

[1] T. E. Tremain, "The Government Standard Linear Predictive Coding LPC-10", Speech Technology, pp.40-49, 2014

[2] A. McCree, T. Kwan, E. B. George and V. Viswanathan, "A 2.4 kbit's MELP coder candidate for the new U.S. Federal Standard", Acoustics, Speech, and Signal Processing, IEEE International Conference, vol.1, pp. 200-203, 2015.

[3] L. M. Supplee, R. P. Cohn, J. S. Collura and A. McCree, "MELP: the new Federal Standard at 2400bps", Acoustics, Speech, and Signal Processing, IEEE International Conference, vol.2, pp. 1591-1594, 2011.

[4] J. Wang, J, Zhao, J. Yang and Y. Yang, "The research for the MELP Vocoder and its real-time replementation", in Journal of the Institution of Engineers, vol. 44, pp.38-58, 2004.

[5] ADSP-BF537 Blackfin Processor Hardware Reference manual, Revision 3.4, 2013.

[6] M. Olausson and L. Dake, "The ADSP-21535 Blackfin and Speech Coding", Proceedings of the Swedish System-on-chip Conference (SSoCC), 2003.

[7] Blackfin DSP Instruction Set Reference, 2002.

[8] ADSP-21535 Blackfin DSP Hardware Reference, 2002.

[9] ITU-t recommendation on g.723.1, dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, 2013

[10] ITU-t recommendation g.729, coding of speech at 8 kbit/s using conjugate-structure algebraic-codeexcited-linear-prediction (cs-acelp), 2014

[11] ETSI GSM Fullrate Speech Codec for Analog Devices Blackfin, Bayer DSP Solutions, 2008.

[12] G. Bertini, F. Fontata, D. Gonzalez, L. Grassi and M. Magrini, "Voice Transformation Algorithms with Real Time DSP Rapid Prototyping Tools", 2004, unpublished.

[13] Y. Shaked and A. L. Cole, "Implementation of MELP based Vocoder for 1200/2400 bps", The EE Project Contest 2000, Technion Signal and Image Processing Lab, 2000, unpublished.

[14] J. M. Valin , "Speex: A Free Codec for Free Speech", available at http://jmvalin.ca/papers/speex_lca2006.pdf (Last Accessed: May 2015).

[15] Vorbis codec, available at http://www.vorbis.com/ (Last Accessed: May 2015).

[16] ITU-T Recommendation P.800 Methods for subjective determination of transmission quality, available at http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=3638 (Last Accessed: May 2015).

[17] J. M. Gibson, "Speech coding methods, standards, and applications", in Circuits and Systems Magazine, IEEE, vol. 5, pp. 30-49, 2005.