Design and Implementation of AMBE based Voice Codec Module over custom FPGA platform

Kapil Mahawar Vishal Kumar H.O. Gupta
Department of Electrical Engineering, Indian Institute of Technology, Roorkee, Roorkee-247 667, India
mahawar.kapil@ieee.org,vksaxfee@iitr.ernet.in,harifee@iitr.ernet.in

Abstract— **This paper presents the design and implementation of a Voice Codec Module (VCM) over a Field Programmable Gate Array (FPGA) platform for low bit rate voice communication using Vocoder AMBE (Advanced Multi Band Excitation)-2000 and Codec AD-73311. Low bit rate voice communication has become essential to many applications like digital radio, satellite communication, underwater acoustic communication etc. where bandwidth is at a premium and voice intelligibility is imperative. The hardware circuit of VCM is designed using FPGA instead of following the traditional DSPs/Microcontrollers approach so as to fulfill the objective of designing all the major subsystems of voice/data communication baseband unit inside a single FPGA chip. The Codec AD73311 used is a high performance A/D & D/A chip with programmable input and output gain control, and AMBE-2000 is the vocoder which is based on the AMBE compression algorithm. Using this VCM, the call voice is clear and legible at the rate of 2.0~2.4 Kbps which is very low compression bit rate.**

Keywords-Voice Communication; FPGA; AMBE-2000; Codec AD73311

I. INTRODUCTION

The trend in voice communication is to compress the voice signal as much as possible in order to conserve the channel resource and promote the communication capacity [1]. Thus there is continuous development in speech compression coding technology and thereby number of researches have been undertaken to develop Voice Codec Modules (VCM) based on different compression coding algorithms using various implementation approaches [2]-[3]. A VCM is one which is capable of accepting the analog speech either from Telephone/Microphone and delivering the analog speech signal back either to Telephone/Speaker.

While the DSPs/Microcontrollers is one of the implementation approach capable of realising the voice communication system as shown in Fig 1, the FPGA chip is attracting increasing attention in these areas as the FPGA is getting faster and more versatile than before [4]- [5] . The trend in hardware design is towards implementing a complete system intended for various applications on a single chip i.e. System on a Programmable Chip (SoPC) which primarily uses Field Programmable Gate Array (FPGA) for SoPC design. In pursuit of this, voice communication systems are also employing FPGAs instead of traditional DSPs/Microcontrollers in order to exchange data between vocoder and modem as depicted in Fig 2.Apart from realizing the voice communication module in FPGA, it is also possible to incorporate various other functions of digital communication system like modem, crypto engine etc. inside the same FPGA. Furthermore, one can customize it with 32-bit RISC soft processors and also use standard peripherals such as Ethernet in a single FPGA device thereby enabling the designer to implement a complete FPGA based SoPC design.

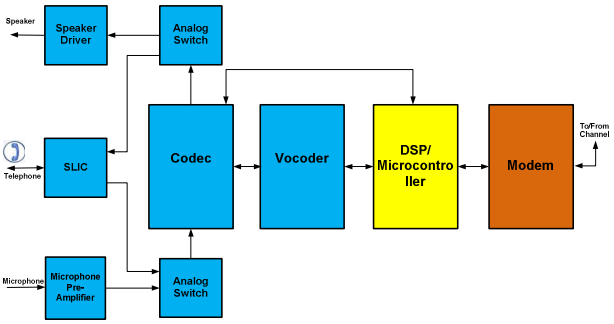

Figure.1. General Block Diagram of any Traditional Voice communication system

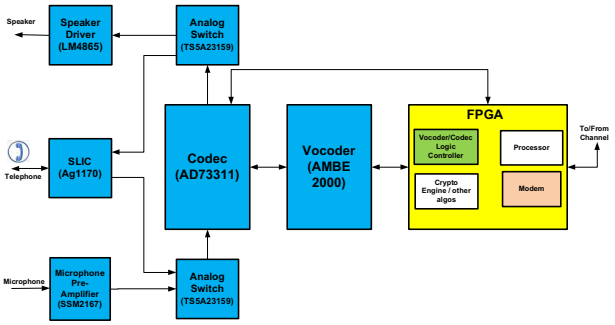

Figure.2. General Block Diagram of any SoPC based Voice communication system

In this paper, our VCM design refers to the Advanced Multi Band Excitation (AMBE), which has proven to be a good improvement of the standard MBE algorithm speech compression coding scheme [6]-[7], an ideal coding plan at rate of 2.4~4.8kbit/s.

In the system design, the analog speech can be inputted through Microphone/Telephone. Subscriber Line Interface Card (SLIC) is being used between the telephone and the Codec in order to provide the necessary line conditions to the conventional handset [8]. Ag1170 [9] was the component we used and it had several other capabilities such as detecting on/off hook conditions, ringing the phone etc. The Codec A/D-D/A chip AD73311 [10] samples and quantizes the input voice data, then the AMBE-2000 [11] runs the speech

compression coding. The encoded voice data is sent to the FPGA where it is processed and sent to the channel. Process of receiving voice data is just opposite of transmitting, where the analog speech can be fed back to either Speaker/Telephone.

The organization of the paper is as follows. Hardware platform and software tool used is discussed in Section II, followed by system architecture in Section III. The system hardware design and implementation issues are discussed in Section IV. Finally implementation results and conclusion are given in Section V and VI.

A. *Hardware Platform and Software Tool*

In the present work, a FPGA based custom development board is used whose block schematic is depicted in Figure.3. The board consists of single PCB containing digital and analog components. Although this board is designed to serve as a full-fledged terminal controller for baseband applications in military based communication, so not all the components are in use for the present work. Here primarily for the implementation of VCM, we are using Altera's Cyclone II FPGA (EP2C70F672I8) [12], vocoder interface from DVSI i.e. AMBE-2000,Audio Codec AD73311,SLIC Ag1170,Speaker driver LM4865, Microphone Pre-Amplifier SSM2167 and analog switch TS5A23159 as depicted in Figure2.

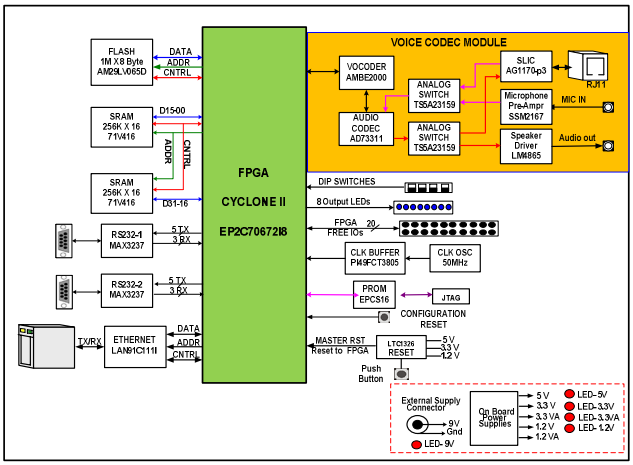

Figure.3. Block Schematic of FPGA based custom development board

The CAD tool used for the development is Altera's Quartus II which is an integrated development environment for facilitating all the tasks in the FPGA design flow namely:

1. Design entry
2. Design Simulation
3. Design synthesis
4. Synthesis simulation
5. Implementation
   a. Technology mapping
   b. Placement & Routing
   c. Back annotation
6. Timing Verification
7. Downloading configuration into FPGA

The entire design has been coded in VHDL [12] and verified using Quartus II [13].

II. System Architecture

The system architecture and the interconnection between various sub-modules of Voice Codec Module are depicted in Figure. 4.The interface between the codec and the vocoder chip is a serial port. In the present design, FPGA performs tasks such as:

1. Loading of configuration data into the control registers of the codec upon initialization.
2. Extraction of data into FPGA from vocoder chip and frame processing.
3. Rudimentary formatting of the data to/from the vocoder chip in 16-bit format.
4. Interface support for SLIC
5. DTMF (Dual Tone Multi Frequency) command decoding.

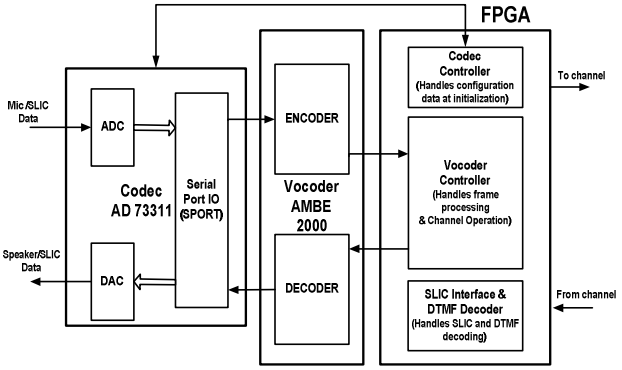

Figure.4. System Architectural Diagram of Voice Codec Module

III. System Hardware Design and Implementation

A. *AMBE-2000 features and operation*

Digital Voice System Inc's AMBE-2000 is a multi-rate speech codec chip which adopts AMBE algorithm to provide superior voice quality at low data rates. AMBE-2000 is characterised by low power consumption, low complexity and it offers a high degree of flexibility in selecting the speech and FEC (Forward Error Correction) data rates. It supports data rates of 2.0 kbps to 9.6 kbps in 50 bps increment. Additionally it offers Voice Activity Detection (VAD), Echo Cancellation, Comfort Noise Insertion (CNI) and DTMF signal detection and generation. This chip is capable of compressing a 128 kbps data stream, into a 2 kbps data stream at an acceptable quality as shown in Figure. 5.

Since our aim was to produce a voice communication system with conventional telephone sets as well, therefore the functionalities listed above were all ideal and hence we used AMBE-2000 chip in our system.

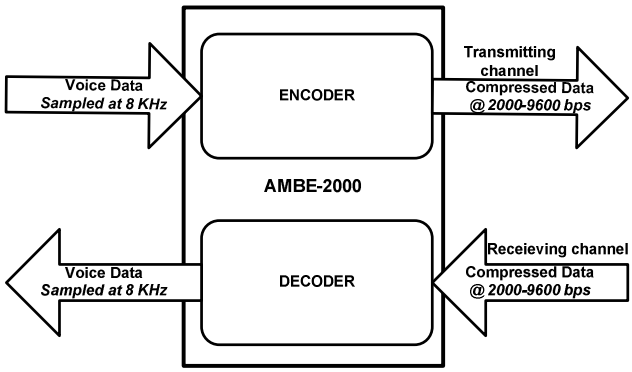

Figure.5. Basic Operation of AMBE-2000 chip

Operation: The basic hardware unit of the AMBE-2000's channel interface is a serial port which can run in passive or active modes and the data has two formats, formatted and unformatted. In our design we have chosen active formatted mode where data strobes for pulling the data is provided by the chip and a known structure is preceded before the encoder data. Under normal operation, every 20 ms the encoder outputs a frame of coded bits, and the decoder needs to be delivered a frame of coded bits. The signals in Table I make up the serial channel interface [1]. The serial channel mode transfers data in and out of the AMBE-2000 using 16 bit words in two data lines CHAN_RX_DATA and CHAN_TX_DATA.

TABLE I : AMBE-2000: Channel Serial Interface Description

Pin Symbol	Description
EPR	Encoder Packet Ready: Output signal after every 20 ms to indicate that the encoder has a frame of data to output.
CHAN_RX_DATA	Serial Data Input: 16 bits of channel data are input synchronous to chan_rx_clk with each chan_rx_strb pulse.
CHAN_RX_CLK	Serial Input Clock: Data is latched on the falling edges of this clock.
CHAN_RX_STRB	Input (Write) Data Strobe: Signal indication to the chip when the data on chan_rx_data will be latched by chan_rx_clk.
CHAN_TX_DATA	Serial Data Output: 16 bits of channel data are output synchronous to chan_tx_clk with each chan_tx_strb pulse.
CHAN_TX_CLK	Serial Output: Data is output on the rising edges of this clock.
CHAN_TX_STRB	Output (Read) Data Strobe: Signal indication to the chip when to bring the data to the chan_tx_data pin.

The AMBE-2000 expects an encoder packet to be read approximately every 20 msec. After initial reset, when AMBE-2000 has a valid initial packet it will generate EPR. FPGA waits for EPR to go low and read the initial packet. For all the following packets, FPGA waits for EPR and follows the procedure below:

1. Wait for EPR to go low.
2. Assert CHAN_TX_STRB and read word on CHAN_TX_DATA.
3. If transmitted word not 0x13EC, discard it and repeat step 2.
4. If transmitted word 0x13EC, read 23 more words (rest of packet).

TABLE II: AMBE-2000: A basic Formatted Input Format

		Word No	Description
20 ms frame	24 sixteen-bit words = 48 bytes = 384 bits	1	Header always set to 0x13EC
		2	Power Control ID (8bits) Control Word 1 (8 bits)
		3	Rate info 0
		4	Rate info 1
		5	Rate info 2
		6	Rate info 3
		7	Rate info 4
		8	Unused in Input
		9	Unused in Input
		10	Unused in Input
		11	DTMF Control
		12	Control Word 2
		13	Channel Data
		14	Channel Data
		15	Channel Data
		16	Channel Data
		17	Channel Data
		18	Channel Data
		19	Channel Data
		20	Channel Data
		21	Channel Data
		22	Channel Data
		23	Channel Data

The Framed format is a 24 by sixteen-bit word format for a total of 48 bytes or 384 bits. Every 20 milliseconds the encoder outputs 24 sixteen-bit words, and likewise the decoder expect to receive 24 words. The format of the input frame is detailed in Table II. The first 12 sixteen bit words are made up of header, ID and status or control information. The remaining 12 sixteen bit words make up the encoded data bit field. These 12 words, or 192 bits, will be fully populated with relevant voice data only when the AMBE-2000™ is operating in a 9600bps mode (9600 bits/sec, 50 frames/sec = 192 bits/frame).Otherwise, when the data rate is less than 9600bps, the coded voice bits are filled starting from the MSB of the first word in the field, leaving any unused bits as zeros.

The special functions of the AMBE-2000 Vocoder chip, such as echo cancellation, voice activity detection, comfort noise insertion, data FEC rate selection etc. can be controlled either through hardware control pins or through the setting of corresponding bits in the input frame format.

B. Interface design between AMBE-2000 , AD73311 and FPGA

In this paper, we selected Analog devices AD73311 codec chip for use with AMBE-2000 vocoder chip. The AD73311 is a complete front-end processor suitable for applications in speech and telephony area. It features a 16 bit A/D and D/A

conversion channel having programmable input and output gain control and a serial port (SPORT) for interfacing of codec with DSP/Microcontroller/FPGA logic via a six-wire interface. SPORT is used to transmit and receive digital data and control information. It has a very flexible architecture that can be configured by programming internal control registers having five read/write registers each 8 bits wide. The first two control registers CRA and CRB are reserved for controlling the SPORT while the other three registers CRC, CRD and CRE are used to hold control setting for the ADC, DAC, reference and power control sections of the device

The interface block diagram b/w the FPGA, codec AD73311 and AMBE-2000 is depicted in Fig 6 which shows that the configuration data is loaded into the codec chip upon power-on reset before the codec chip exchanges data with the vocoder chip. The objective of this FPGA logic is also to tri-state the output of the AMBE-2000 CODEC_TX_DATA which allows the FPGA to communicate with the AD73311 to send the desired configuration data as per Table III [3].

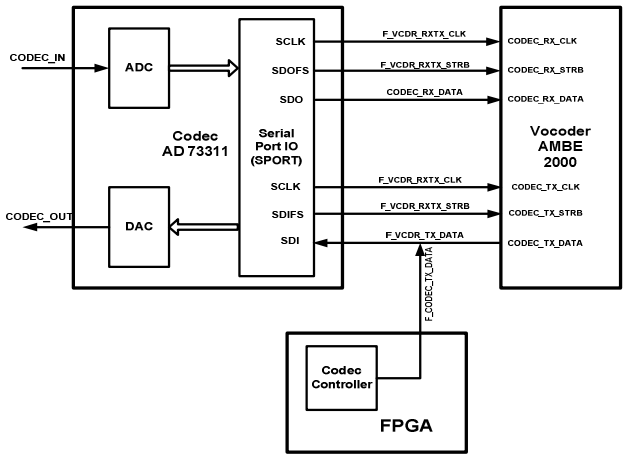

Figure.6. Interface Block Diagram between Codec, Vocoder and FPGA

Control Register	Configuration Data	Notes
CRB	0x13	MCD = 1 Sets DMCLK=MCLK/2 SCD = 0 Sets SCLK=DMCLK/8
CRC	0x79	D6 = 1 Reference Out Enabled D5 = 1 Reference Power On D4 = 1 DAC Power On D3 = 1 ADC Power On D0 = 1 Power On
CRD	0x00	Gain set to 0
CRA	0x01	D0 = 1 Puts CODEC in Data Mode

TABLE III: AMBE-2000: Channel Serial Interface Description

C. Interconnection for achieving desired voice compression

The functionality of the entire voice codec module is described in Fig. 7 which shows how different modules are interconnected to achieve the desired voice compression (say 2.4 kbps) from the input 128 kbps digitised voice . The encoder of the compressor produces a word at every 20ms, which had payload and overhead. This frame had to be pulled into the channel processor by the use of two clocks for data and strobe respectively. We used FPGA logic to do this job of pulling the data and the processing required as mentioned in earlier sections. Here to achieve 2.4 kbps data transmission at channel ,only 48 bits out of the 384 bits are transmitted to the receiving end where they are placed in the same positions of another 384 bit frame and fed into the decoder at receiver.

D. Other Modules implemented in FPGA

DTMF algorithm is based on the DTMF field of the voice codec frame (Word no 11 out of 24 words) which is triggered upon the occurrence of the edd (encoder dtmf detect) signal. When a digit is dialled, its DTMF is transmitted by the voice codec in every 20ms. This is identified by the DTMF command decoder logic implemented in FPGA.

SLIC interface is also implemented in FPGA for decoding logic of the Ag1170 chip for functions like generating ringing, dial tone and ring tone etc.

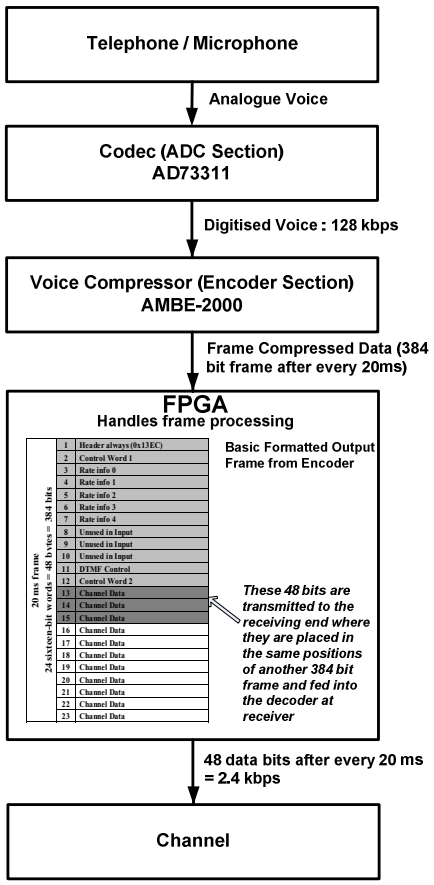

Figure.7. Achieving 2.4 kbps data transmission: figure describes how different modules are interconnected to achieve the desired voice compression

IV. IMPLEMENTATION RESULTS

We have used Quartus II 8.1 to create a VHDL model of the Voice Codec Module in Cyclone II EP2C70F672I8 FPGA. As it is difficult to simulate the complete design in totality, so we have shown the implementation results directly. Out of

many sub-modules implemented, one of them is vocoder controller which has two FSMs, one for transmitter and other for receiver having register architectures. Each such FSM coding is verified part-wise with the help of Signal Tap Logic Analyser tool in the Quartus II development environment. The logic analyser tool helps in real time verification while capturing even the internal signals in the design. Finally for the real time testing of the complete design:

- The line in jack of the custom development board is connected to the laptop/Microphone to input the analog speech signal.
- The line out jack of the development board is connected to the Speaker/Headphone to output the analog speech signal.

As in design, provision is made to input/output the compressed bit stream through uart controller. Thus when the TxD and RxD pins are shorted the same audio signal is loop backed and thus possible to check the results with headphones or speakers even with the help of one development board. And in presence of two such boards, both were connected with the help of a cross cable to verify the setup. It results in legible voice at the low rate ranging from 2400 bps to 9600 bps. Table IV presents the resources to implement the design in the Cyclone II EP2C70F672I8 FPGA .The columns falling under LE,LR,MB,EM, and PLL labels describe logic elements, logic registers, embedded multiplier ,memory bits and phase locked loops

	LE	Pins	LR	MB	EM	PLLs
Occupation	4019	49	3228	0	0	0
Total	68,416	422	68,416	1,152,000	300	4

TABLE IV: Implementation Results of Compiling in FPGA

V. CONCLUSION

In this paper, it is demonstrated that how FPGA can replace the traditional DSP/Microcontrollers to realize a Voice Codec Module for low bit rate voice communication using AMBE-2000 Vocoder and Codec AD73311.With the high performance of AMBE-2000 chip, it results in a legible voice even at the low bit rate of 2000 bps. This particular design is able to realize effective speech communication subsystem in a FPGA which can be used in conjunction with other important subsystems of communication like modem, crypto engine, processing engine etc in the same FPGA to implement a complete system on a single chip. As it is quite evident from literature that the combination of programmable logic and IP cores including micro-processors can speed up the system design, thereby this FPGA based solution for low bit rate speech communication is of prime importance.

ACKNOWLEDGMENT

Kapil Mahawar would like to thank Shri R.C. Agarwal, Director, DEAL, Dehradun for granting permission to undergo M.Tech programme at IIT,Roorkee where this work was carried out. Authors would also like to thank Sh. R.K. Agrawal, Associate Director, and Sh. Manoj Dhaka for providing hardware resources to carry out the work and also Head, EED, IITR for valuable guidance and support for the completion of the work.

REFERENCES

[1] T. S. Rappaport, Wireless Communications: Principles and Practice, Prentice-Hall, second edition, 2002.
[2] A. S. Spanias, "Speech coding: A tutorial review," Proceedings of the IEEE, vol. 82, pp. 1541-1582, Oct. 1994.
[3] A. Gersho, "Advances in speech and audio compression," Proceedings of the IEEE, vol. 82, pp. 900-918, June 1994.
[4] Luo Jian, Zhou Hongwei, Li Yi and Lu Minglei, "The design of underwater speech communication codec system based on AMBE-3000,",IEEE Conference on Mechatronic Science, Electric Engineering and Computer, pp 2471-2475, Aug 2011.
[5] M. Samarawickrama , K. Pathirage ,H. Samwarsekara A. Wickramasinghe and K. Samarasinghe, "Power Line Intercom – Narrowband Voice Communication over Power Line," IEEE Conference on Industrial and Information System, pp 37-42, Dec 2009.
[6] D. W. Griffin and J. S. Lim, "Multiband Excitation Vocoder," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol 36, No 8,August 1988, pp. 1223-1235.
[7] B. Atal and M. Schroeder, "Stochastic Coding of Speech at Very Low Rates", Proceedings of ICC, 1984, pp. 1610-1613.
[8] J. C. Bellamy, Digital Telephony, John Wiley & Sons, 2000.
[9] "Ag1170 – Full featured low power SLIC," Silver Telecom data sheet V1.10, February 2007.
[10] "AD73311 – Codec," Analog Devices Inc , May 2000.
[11] "AMBE-2000 vocoder – High performance voice compression solution," Digital Voice Systems, Inc. data sheet, May 2010.
[12] J. R. Armstrong and F. G. Gray, "VHDL Design, Representation and Synthesis", 2nd ed., Prentice Hall, 2000.
[13] (2011) The Altera Corp website. [Online]. Available: http://www.altera.com/