

Comparison of Low-Rate Speech Transcoders in Electronic Warfare Situations: Ambe-3000 to G.711, G.726, CVSD

V. Govindu
Department of ECE,
UCEK, JNTUK,
Kakinada, India 533003.

Parthraj Tripathi
Defence Electronics Research
Laboratory, DRDO,
Hyderabad, India 500001.

B. Leela Kumari, PhD
Department of ECE,
UCEK, JNTUK,
Kakinada, India 533003.

ABSTRACT

Continuous efforts are active to reduce the bit rates but maintaining channel noise tolerance, secure transmission and justified MOS(mean opinion score) among various communication networks. These networks at their end-terminals may employ variety of vocoders operating at different individual bit rates. In order to maintain fidelity, transcoders are used to map the information when traffic flows from one channel operating at one bit rate to another channel operating on another bit rate as seen in the case of channels with different capacities. Some networks (like satellite communication and some private networks) employs codecs like AMBE (Advanced Multiband Excitation), CVSD for their very low bit rate, channel noise tolerant attributes and another features. In order to interface networks accompanying the said vocoders with that of public networks containing codec like PCM, ADPCM, we have done feasibility study for justified MOS using AMBE-3000 HDK. Also, we have compared Transcoders against MOS.

Keywords

Low rate, Speech coding, AMBE-3000, G.711, G.726, CVSD, Transcoding

1. INTRODUCTION

Speech coding or compression is a process of obtaining a compact representation for the speech signals, for the purpose of efficient transmission over band limited wired or wireless channels and also for efficient storage. A speech coder is one which converts a digitized speech signal into a coded representation and transmits it in the form of frames. At the receiving end the speech decoder receives the coded frames and performs synthesis to reconstruct the speech signal. The speech coders differ primarily in bit-rate, complexity, delay and perceptual quality of the synthesized speech at which they produce output with reasonable quality. In digital communication system, the quality of the entire communication system has a direct relationship with the performance of speech communication. So we are increasingly tend to compress the voice signal as much as possible in order to conserve the channel resource and promote the communication capacity, of course the compression of the speech signal must be within a certain quality. And it will promote the continuous development of speech compression coding technology and speech coding technology [1] is playing an important role for voice communication.

Compressing the bandwidth of speech signal in a digital speech communication system with speech coding technique, especially with low-rate speech coding technique [2], is still a hotpot which is concerned in the communication field now and in the future. The G.711, G.726, CVSD and AMBE-3000

codec's are the separate representative low-rate speech coding algorithms among the current speech coding research. In this thesis, we design and study the vocoding conversion system between AMBE-3000 and G.711, G.726, CVSD by using the combination of software and hardware.

2. DESIGN OF CONVERSION SYSTEM BETWEEN AMBE-3000 TO AND G.711, G.726, CVSD

There is no direct transcoding algorithm for the AMBE and G.711, G.726, CVSD at present. However, the translation of the AMBE format and the PCM format is achieved by the AMBE-3000 module, and the translation of the G.711, G.726, CVSD format and the PCM format is achieved by the G.711, G.726, and CVSD module. The project intends to adopt a combination of software and hardware, through the conversion.

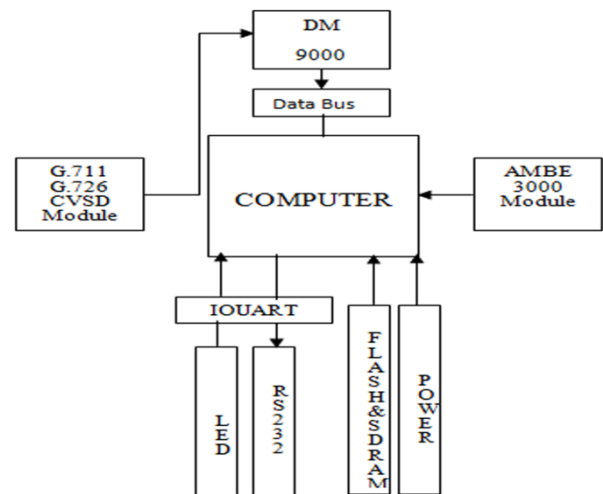


Fig. 1. The System block diagram

The system is segmented into four modules: the AMBE-3000 module, the G.711, G.726, CVSD module, the computer module and the network transmission module. The translation of the AMBE format and PCM format achieved by the AMBE-3000 module, and the translation of the G.711, G.726, CVSD format and the PCM format is achieved by the G.711, G.726, CVSD module, so that the translation of the AMBE format and the G.711, G.726, CVSD format is achieved by PCM format.

3. G.711, G.726 AND CVSD CODECS

As the widely applications of digital multi-media technologies G.711, G.726, CVSD has become most popular audio standards. G.711 PCM codec is used as narrowband audio codec it uses 64kbps PCM A-law and μ -law which gives high

quality speech for VoIP. G.726 is used to convert 64kbps A-law or μ -law PCM to and from a 40, 32, 24 or 16kbps channels for carrying voice overload and data modem signals in DCME. CVSD is a voice coding method. It is a delta modulation with variable step size. It encodes at 1bit per sample, so that audio sampled at 16 kHz is encoded at 16kbits/s.

3.1 Implementation of G.711, G.726 and CVSD Module

After studying the algorithms of G.711, G.726, CVSD, we implemented the codec's of G.711, G.726, CVSD in VC++6.0 platform. To compare the difference between the input speech file and the output speech file, we use MATLAB software to the analysis the changes of the encoded file and the decoded file. The results of the experiments:

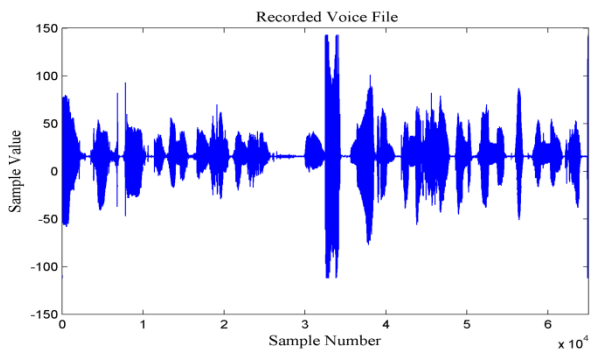


Fig. 2. Time-domain waveform of the input speech file

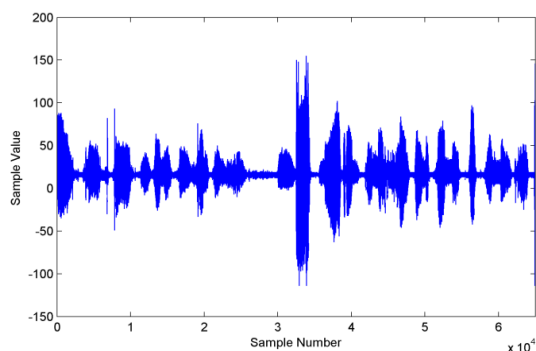


Fig. 3. Time-domain waveform of the output speech file

4. ADVANCED MULTIBAND EXCITATION (AMBE)-3000 SPEECH CODER

The multi-band excitation (MBE) speech compression coding scheme is an ideal coding plan at rate of 2.4~4.8kbit/s, which is proposed by MIT Lincoln Laboratory of the United States in 1988. However, the AMBE algorithm has proven to be a good improvement and complement of the standard MBE algorithm. Its coding rate, algorithm delay and average complexity are minimum in the common parameter coding and hybrid coding algorithm. Additionally, the AMBE algorithm has a strong suppression of noise and a good naturalness.

In the system design, we select the DVSI Company's latest AMBE voice codec chip--AMBE-3000. This chip adopts AMBE algorithm to realize the speech compression coding low bit rate and high voice quality [3]. The designed transcoding speech communication system which is based on

the AMBE-3000 could provide a legible voice at the rate of 2.0kbps. It also has the advantage of low bit rate compression, low power consumption, long distance transmission and so on.



Fig. 4. AMBE-3000F™ Vocoder Chip

4.1 AMBE-3000 Features And Design Elements

AMBE-3000 is a multi-rate speech codec chip with high performance. It has two operating modes [3]: Codec mode and Packet mode. In Codec mode, the A/D-D/A converter module of speech communication system is connected with the codec module to execute data exchange directly. In Packet mode, the A/D-D/A converter module is connected with the codec module by a MCU to execute data exchange indirectly. AMBE-3000 characterized by low power consumption, low complexity and high performance of voice is marked by Flexible Codec Rate (from 2000bps to 9600bps), Forward Error Correction (FEC), and Voice Activity Detection (VAD), Echo Canceller, Comfortable Noise and DTMF signal detection. So it is very suitable for digital speech communication, encrypted speech communication and other needs of digital speech processing applications.

Some of the AMBE-3000 hardware design considerations are following:

- Speech and FEC rates. These rates are preferred by using a configuration control packet, or through hardware configuration pins.
- Mode of process (Codec mode or Packet mode).
- Choice of Codec interface. (McBSP, SPI)-for Codec mode only.
- Choice of Packet interface. (UART, McBSP, PTT).

There are four physical interfaces (UART ,SPI, Parallel and McBSP) used to allocation the data to/from the AMBE-3000 Vocoder Chip. For codec mode, the user must choose two physical interfaces: one for the packet data and one for the codec data. The choices for the codec interface are McBSP or SPI. The choices for the packet interface are UART or McBSP or Parallel Port.. The McBSP cannot be used for both the packet interface and the codec interface. For packet mode, the user must choose one physical interface to be used for packet data. The packet interface is used to transfer both the compressed channel data and the speech data samples. The choices for the packet interface are UART or McBSP or Parallel Port. And the obtainable interface arrangements are shown in Table I Physical Interface Selection.

Choice of A/D-D/A chip: The AMBE-3000 Vocoder Chip can be formed to transmit and receive digitized speech to and from maximum a-law, μ -law or linear A/D-D/A codec's. The format of the outgoing and incoming speech data streams are

fixed, that is to say they must be the same format (8-bit a-law, 8-bit μ -law or 16-bit linear).

Table 1. physical interface selection

| Interface Configurations | | | | | |
|--------------------------|--------------------------------|------|------|--------------------|---------------------|
| MODE | IF_SELECT Configuration Pin | | | Codec Interface | Packet Interface |
| | 4/B1 | 3/C3 | 2/C2 | | |
| Codec Mode | 0 | 0 | 0 | SPI | UART |
| Codec Mode | 0 | 0 | 1 | SPI | PPT |
| Codec Mode | 0 | 1 | 0 | SPI | McBSP* |
| Codec Mode | 0 | 1 | 1 | McBSP* | UART |
| Codec Mode | 1 | 0 | 0 | McBSP* | PPT |
| Packet Mode | 1 | 0 | 1 | Not used | UART |
| Packet Mode | 1 | 1 | 0 | Not used | PPT |
| Packet Mode | 1 | 1 | 1 | Not used | McBSP* |

The AMBE-3000 Vocoder Chip proposals several interfaces for flexibility in integration into a selection of design configurations. The special functions of the AMBE-3000 Vocoder Chip, such as echo cancellation, voice activity detection, power mode control, data FEC rate selection, and etc. can be controlled through either hardware control pins or the packet interface. AMBE-3000 will be initialized according to its hardware configuration after system power on. There are as many as 30 kinds of software configuration packet command of AMBE-3000. These commands can be used to execute many common function or special function configurations to cover the hardware setting.

4.2 The Interface Design Between AMBE-3000 And TLV320AIC14

In order to ensure proper performance from the voice coder, it is necessary to select a proper A/D-D/A chip between the microphone and AMBE-3000. The selected A/D-D/A chip is crucial to the quality of speech communication system. A 16 bit linear A/D-D/A chip is suggested taking the quality of voice into account. However, the SNR and filtering performance of A/D-D/A chip is considered at the same time. It is worth noting that the A/D-D/A chip must operate at an 8 kHz sampling rate in order to ensure the voice rate [4].

In this paper, we select TLV320AIC14. It features one 16-bit A/D channel and one 16-bit D/A channel. It provides high-resolution signal conversion from analog-to-digital (A/D) and from digital-to-analog (D/A) using oversampling sigma-delta technology with programmable sampling rate. It is an accurate high-performance, low-power, low-cost, mono voice codec with ant aliasing filter (AAF), and selectable low-pass FIR/IIR filters integrated. However, the TLV320AIC14 codec presents a low cost solution for use with AMBE-3000 vocoder chip. The control registers in the TLV320AIC14 codec must be initialized for accurate operation. The proposed process is to prepare the TLV320AIC14 by writing data to 5 control registers via packet from the AMBE-3000 vocoder chip after the system power on. The main register configuration data as follow:

Table 2. Control register value for the TLV320AIC14

| Control Register | Configuration Data | Note: |
|------------------|--------------------|---|
| 1 | 0x41 | Set 16 bit DAC mode, set configuration data transfer mode |
| 2 | 0xA0 | Set TURBO=1(SCLK=MCLK/P), keep I2C addr=4 |
| 4 | 0x83 | Set M=3 |
| 5C | 0xB8 | Side tone=MUTE |
| 6 | 0x02 | set input MICIN selfbiased at 1.35 V |

AMBE-3000 could send the configuration data above to make TLV320AIC14 run into the following working state:

MCLK=18.432MHz;SCLK=MCLK/P=2.304MHz;FS=MCLK / (16*M*N*P) =8 kHz.While MCLK is from the external clock input and the default value of P is 8. Register 4 is configured with M=3. Register 6 is configured with the speech output port (OUTP2, OUTP3) off. Using OUTP1 as the speech input port. As there is an integrated operational amplifier inside TLV320AIC14, the output speech signal could drive the headphones directly, which results in a simple and flexible circuit design. The interface circuit between AMBE-3000 and TLV320AIC14 is shown in Fig 5.

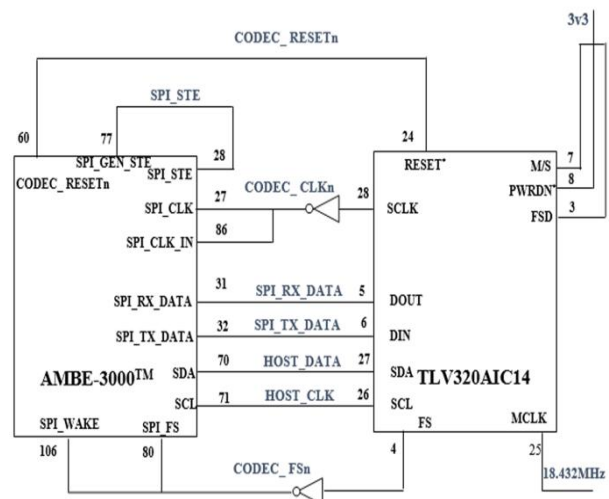


Fig. 5. AMBE-3000 Vocoder chip and TLV320AIC14 interface block diagram

4.3 AMBE-3000 Vocoder Chip And MCU Interface Design

The speech data sampled by A/D will be processed and cached in MCU before AMBE-3000 Companding. And the receiving data from the channel will be done in the same way before AMBE-3000 decoding. There are three connection

ways between MCU and AMBE-3000: McBSP, UART and Parallel Interface. AMBE-3000 provides UART serial interface and parallel interface which is different from AMBE-2000[5]. This makes it more convenient and simple to connect with MCU. In the design, we adopt asynchronous serial communication mode, MCU will receiving speech data automatically through open receive interrupt by judging the TX_RDY level of AMBE-3000. In the same way, MCU would send a frame data to AMBE-3000 every 20ms through open transmit interrupt.

4.4 System Software Design

The software design of speech communication system is mainly programming for MSP430F5438. First step is to initialize MCU, including the clock serial communication port and other I/O ports. AMBE-3000 will be initialized according its hardware configuration after power on. MCU would configure some features of AMBE-3000 through sends some control commands at the same time. Some related control packets are sent to AMBE-3000 by MCU to realize TLV320AIC14 initialization.

Some of the software design considerations are following:

- Codec configuration packet sequence. There are three Codec control command packets: PKT_CODECCFG ! PKT_CODECSTART ! PKT_CODECSTOP. Data configuration packet order is first to exit low-power mode by PKT_LOWPOWER, second to configure register by PKT_CODECCFG and the last is to start Codec mode through PKT_CODECSTART.

Data exchange between MCU and AMBE-3000. MCU reads data from AMBE-3000 every 20ms through interrupt receiving. There must be the same data rate when MCU sends data, which could be judged by the signal from TX_RDY.

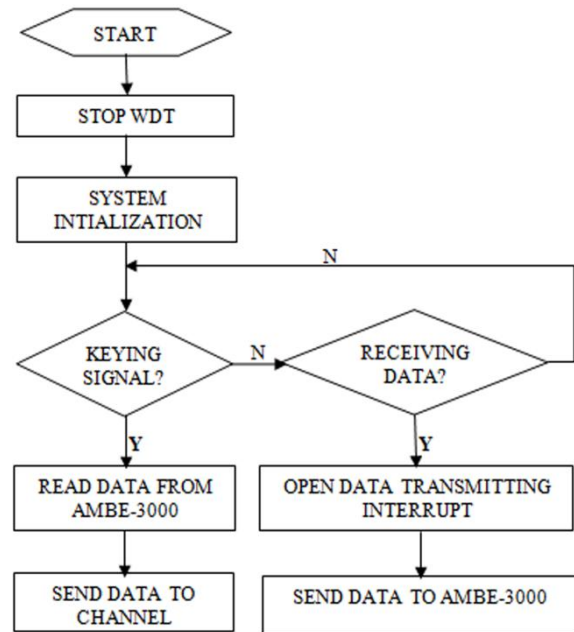


Fig. 6. System main program flow as follow

5 THE PROPOSED TRANSCODING ALGORITHM

The architectures of the tandem system and the proposed transcoder are depicted in Figures 7(a) and 7(b), respectively. For the tandem process, the G.711, G.726, CVSD encoded speech is first decoded by the G.711, G.726, CVSD decoder to obtain decoded speech, which is then compressed by the AMBE-3000 encoder to obtain AMBE-3000 coded speech. Similarly, the AMBE-3000 encoded speech is transformed into G.711, G.726, CVSD coded speech. However, the tandem approach wastes many useful coded speech parameters that exist in the compressed speech in the other format. In this paper, we proposed a transcoding method (depicted in Figure 7(b)) too directly and effectively convert the LSP and open-pitch parameters from G.711, G.726, and CVSD (AMBE-3000) to AMBE-3000 (G.711, G.726, CVSD) coded speech. With the proposed transcoding method, we can dramatically reduce the computations required for the encoder in retrieving the LPC and the open-loop pitch parameters.

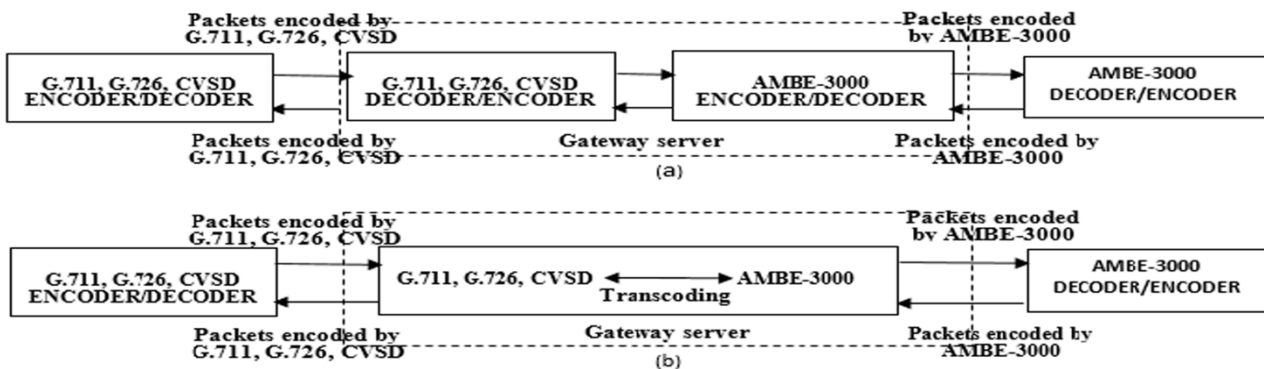


Fig. 7. (a) System block of the tandem approach, (b) system block of the transcoding method

6. EXPERIMENTAL AND ANALYSIS RESULTS

This section analyses, measures and compares different types of codec's performance in campus environment. The MOS will be shown as a single number in the range 1 to 5, where 1 represents the lowest quality, and 5 is the highest quality. The MOS is generated by averaging the results of a set of standard, subjective tests where a number of listeners rate the heard audio quality of test sentences read separately by both male and female speakers over the communications medium being tested. MOS of anything above a 4.0 is considered toll grade (see Fig 8 and Table 3).

Table 3. MOS Rating Measurement

| MOS | QUALITY | IMPAIRMENT |
|-----|-----------|------------------------------|
| 5 | Excellent | Imperceptible |
| 4 | Good | Perceptible but not annoying |
| 3 | Fair | Slightly annoying |
| 2 | Poor | Annoying |
| 1 | Bad | Very Annoying |

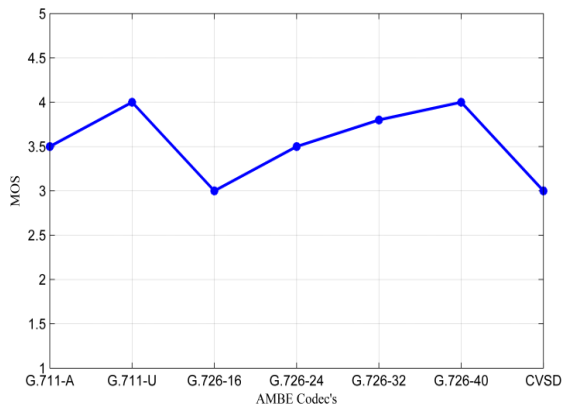


Fig. 8. MOS of AMBE-3000 Codec's

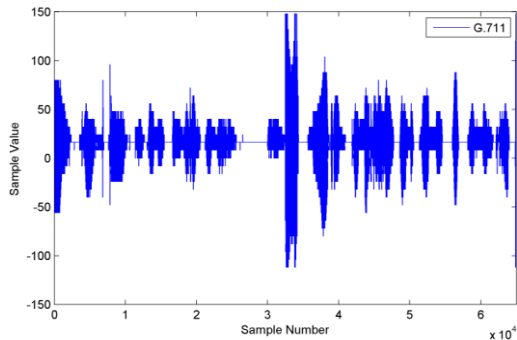


Fig. 9. Time- domain waveform of G.711 input speech

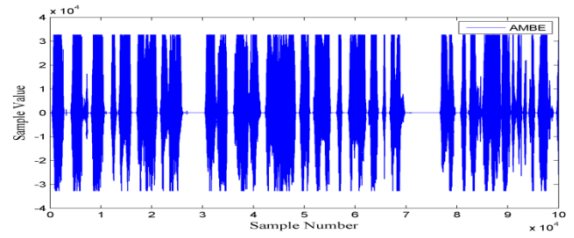


Fig.10. Time- domain waveform of AMBE output speech

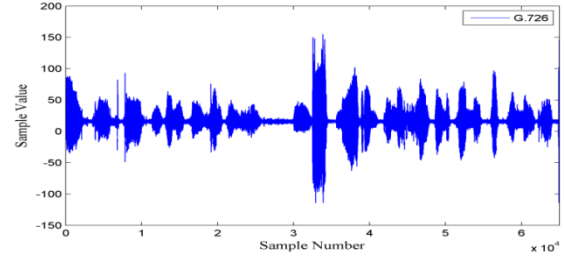


Fig.11. Time- domain waveform of G.726 input speech

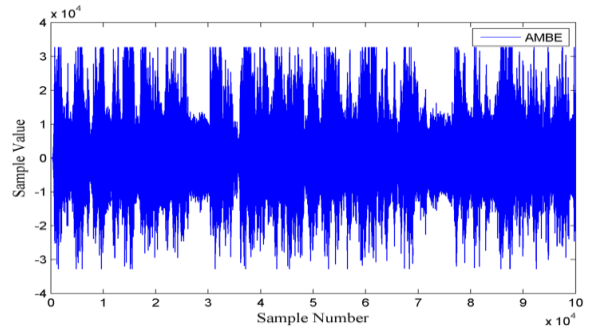


Fig.12. Time- domain waveform of AMBE output speech

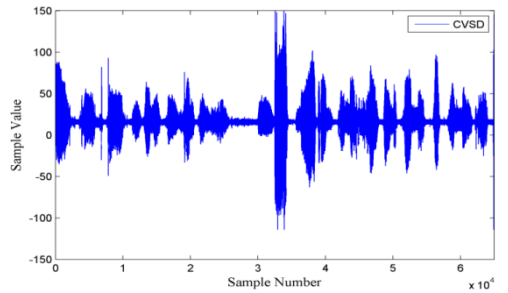


Fig.13. Time- domain waveform of CVSD input speech

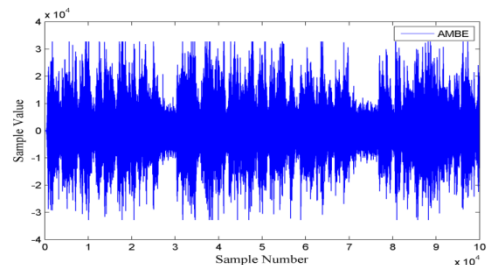


Fig.14. Time- domain waveform of AMBE output speech

7. CONCLUSION

In this paper, a low-rate vocoding conversion system between AMBE and G.711, G.726, CVSD is designed and the outputs of the AMBE-3000 HDK were observed. For these AMBE decoded outputs Mean Opinion Scores were justified. Among these systems, we observe that MOS of PCM is high compared to other systems. The conversion between G.711, G.726, CVSD and PCM is achieved by software, and the conversion between AMBE and PCM is implemented based on Computer.

8. FUTURE SCOPE

The low-rate vocoding conversion is becoming more and more popular in nowadays, and it will have boarder application prospect with the further improvement of this system.

9. REFERENCES

- [1] Jayant, N., "Digital coding of speech waveforms: PCM, DPCM, and DM quantizers," in *Proceedings of the IEEE*, vol.62, no.5, pp.611-632, May 1974 doi: 10.1109/PROC.1974.9484.
- [2] Tremain T E. The Government Standard Linear Predictive Coding Algorithm: LPC-10[J] *Speech Technology*.1982, 4:40-49.
- [3] AMBE-3000(TM) Vocoder Chip Users Manual Version 2.2.Digital Voice Systems Inc., May, 2010.
- [4] QIANG Wei, XU Yu-bin, SHA Xue-jun, GU Shuo. Application of AD73311 A/D converter in digital speech coding [J]. *Journal of Harbin Institute of Technology*, vol. 38, No. 5, 2006, pp. 780-782.
- [5] DU Jun, GAO Jun, LI Nan. Design of the Multi-mode Digital Speech Communication System Based on AMBE-2000 and DSP [J]. *Ship Electronic Engineering*, vol. 28, No. 2, 2008, pp. 80-82.