International Journal of Modern Electronics and Communication Engineering (IJMECE) ISSN: 2321-2152
Volume No.-6, Issue No.-6, November, 2018

Implementation of Speech Companding Technique in Blackfin Digital Signal Processor for Digital Telephone Systems

T. Veeramanikandasamy1st
Department of Electronics and Communication Systems
Sri Krishna Arts and Science College
Coimbatore 641008, India.
E-mail: veeramaniks@gmail.com

M. Venkatachalam2nd
Department of Electronics
Erode Arts and Science College (Autonomous)
Erode 638009, India.
E-mail: eacmvenkat@yahoo.com

Abstract: The companding is a Pulse Code Modulation technique which can be effectively used in different applications, especially in digital telephone systems. A digital telephone system gets an analog speech input signal and it is converted into a digital signal usually it is considered as a linear signal. In order to reduce the transmission bandwidth, the signal is compressed before transmission which would become a non-linear signal. The non-linear digital signal is converted back to a linear digital signal in digital reception. The international standard A-law (G.711) is a speech companding technique, to compress 13-bits linear PCM data down to 8-bits of logarithmic data and expands 8-bits of logarithmic data back to 13-bits linear PCM data that allowing for a bit rate of 64 kbps. This A-law companding technique is implemented by using ADSP BF533 processor which is an enhanced member of the Analog Device Blackfin family that offers significantly higher performance in lower power for new product development.

Keywords: Digital Telephone System, Logarithmic PCM, A-law Companding, Blackfin processor

I. INTRODUCTION

In the field of digital telephony, the digital signal processor has brought tremendous advancements. They include noise-free signal transmission and reception, teleconferencing, and secrecy in transmission through coding. DSP systems are characterized by real-time operation, with emphasis on high throughput rate, and the use of algorithms requiring intensive arithmetic operations, notably multiply-accumulate. The speech coding is concerned with the development of techniques which exploit the redundancy in the speech signal, in order to reduce the number of bits required to present it. The main application areas for speech coding are voice mail systems, a cordless telephone channel, narrow band cellular radio, and military communications [1].

In general speech signal contain frequency with sufficient energies up to about 4 KHz. The telephone speech signal is band limited to 3.3 KHz since the information-bearing formants are concentrated in the frequency region below 3.3 KHz [2]. For this application, the sampling rate F for speech will normally in 8 KHz. The sampled signal is represented by a 16-bit code for high-quality speech processing applications. The ADSP-BF533 is the Blackfin processor offering significantly higher performance in lower power. It has code compatibility benefits and wide on-chip memories. Due to this reason, the DSP hardware and its software are used which provides the simplest and most efficient way to perform companding. This companding technique is the basic method for developing the new ITU standards.

The aim is being focused on to decrease the number of bits by passing through a Pulse Code Modulation process. The G.711 recommendations of the CCITT (Consultative Committee for Telephone and. Telegraph) are specified that a 13-bit information word should be reduced to an 8-bit PCM word correspondingly reducing the bit flow while maintaining an acceptable quality [3]. In order to reconstruct the speech waveform that appears is by the properties of the human voice. Even though it is slightly different from the original voice signal when it reaches to the human ear.

II. SAMPLING THE SPEECH SIGNAL

The speech signal can be divided into three different classes of phonemes. The phoneme is defined as voiced, unvoiced, and plosives. The voiced phonemes are measured deterministic in nature. This sound is produced by forcing air through internal and external muscles of the larynx with the passive longitudinal tension of the vocal folds and they vibrate in a relaxed oscillation. The unvoiced signals are caused by forming a constriction at some point in the vocal cords and forcing air at a high enough velocity to produce turbulence. As a result, unvoiced phonemes have an average power of less than a threshold level and it is considered random in nature [4]. Similar in nature the plosives are consonant sounds that are formed by complete closure of the vocal tract, building up pressure behind the closure, and abruptly releasing it. Each phoneme type gives its own tension to the telephone system. In general, the peak to peak voice amplitude is approximately five to ten times higher than the unvoiced and plosive phonemes [5], as clearly shown in Fig.1.

As a result, a large range of signal amplitudes is observed from the telephone system. The lower amplitude in speech signal represents unvoiced and plosive phonemes that contain more information and thus, higher entropy then voiced phonemes. Thus, the lower amplitude signals are more responsible for the higher resolution in the telephone system. In addition, the bandwidth is restricted with respect to the human speech and auditory ranges in the telephone network. The bandwidth of the speech signal for adults is approximately 10 kHz. In contrast, the minimum and maximum auditory range of humans is 20 Hz - 20 kHz and typically hearing bandwidth of adults is 15 kHz [6]. This maximum auditory range is normally limited to the children.

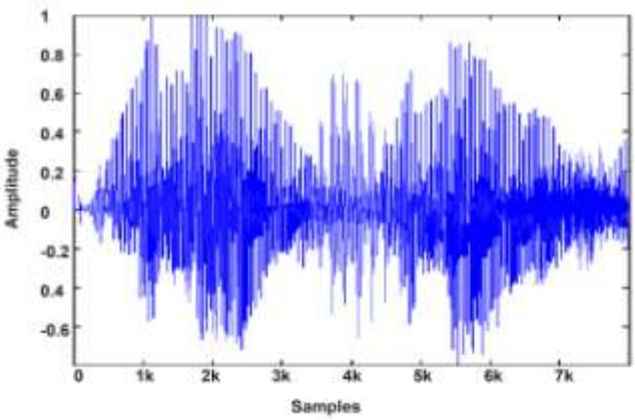

Figure 1. Sample of Speech Signal

The telephone network limits transmission bandwidth of voice to a 3 kHz especially from 0.3 to 3.3 kHz. This frequency range is supposed to coincide with the region of entire speech signal including all three formant frequencies of the sampled speech signal. The unused spaces in the reduced bandwidth are 0 to 0.3 kHz and from 3.3 to 4 kHz. A guard band is allocated due to unused space which provides a buffer against conversation interference. A total bandwidth for the telephone network is obtained by summing the transmission band and guard band, normally it is about 4 kHz. Due to the occasional occurrence of high energy voiced phonemes, the wide range of signal amplitude is required for the transmission in the telephone system [7]. Within a limited bandwidth, the accomplishment of these concurrent tasks can be achieved by Pulse Code Modulation (PCM) and companding.

In the first step, the higher frequency components in the speech signal are removed by the digital filter. The energy of speech information lies somewhere between 300 Hz and 2800 Hz. The typical bandwidth for the speech and voice communication is 3000 Hz. The input speech signal is band-limited with Nyquist criterion by the digital filter to prevent aliasing. If the sampling frequency is less than the highest frequency of the input analog signal causes the undersampling. Normally the low-pass filter is used to reconstruct the original input signal and it creates a new signal [8].

In the second step, the filtered input signal is continuously sampled at a constant sampling frequency by pulse amplitude modulation process. The sampling frequency is two to three times higher than the maximum input analog speech signal per second. The sampled data are quantized and digitized for the transmission over a telephony network. This process can be done by the Pulse Code Modulation. PCM decodes each analog input speech sample into an 8-bit PCM binary word. In PCM, an analog-to-digital converter samples the analog signal which exist in the source side, a digital-to-analog converter is residing on the destination side and it encodes the sample by quantization with 1 to 256 levels [9].

III. SPEECH CODING SYSTEM

The speech coding is important in digital telephone systems so that the speech coding methods have progressive consideration in recent years. The CODEC is the combination of coder and decoder. The requirements for a good speech CODEC can be:

- Quality of speech should not be affected
- A speech signal is compressed in a lesser bit size
- A small delay may be introduced in the coding and decoding part
- CODEC is not sensitive to errors during the transmission
- The computation for coding and decoding should be faster, and the technique used for this purpose should not damage the quality too much.

No perfect CODEC algorithms should satisfy all the requirements because part of the requirements is contradictory. However, by making different compromises, a large number of coding standards for different applications have been developed [10]. All the above requirements are essential for recording a speech signal in the database whereas the computational load and error resiliency is inessential, only the quality and good compression ratio are essential.

There are plenty of coding methods available to the speech compression, but they can be generally divided into two main classes: waveform coding and source coding. In waveform coding, an effort is made to maintain the waveform of the original input signal and the coding is based on quantization and removal of redundancies within the waveforms [11]. In source coding, the speech parameters are coded which enabling the reconstruction in the decoder. The bandwidth of speech is the same as in telephone network which is 300-3400 Hz and the sampling frequency is 8 kHz. This is so-called narrowband speech. In some applications, the narrowband network is not being used and also higher sampling frequency may be used such as video conferences. In these applications, the bandwidth is usually 50-7000 Hz and the sampling frequency is 16 kHz. This is so-called wideband speech.

A-law companding is used in European telephony. North American telephony employs µ-law companding. The µ-law and A-law standards are specified in the International Telecommunication Union (ITU) recommendation G.711. For the variation in input levels, this quantizer technique maintains an approximate 35-dB signal to the quantization noise ratio of an 8-bit/sample. This noise level is almost low for speech communication over telephone systems. The speech sampling rate is 8 kHz which yields an overall bit rate for coded speech is 64-kbps [12].

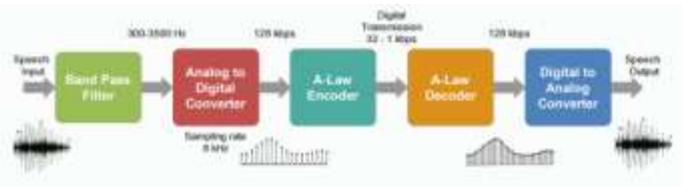

Figure 2. The Speech Coding System

The block diagram of the speech coding technique is shown in Figure 2. The first block involved in digital speech is sampling and amplitude quantization which is done by CODEC. The sampling frequency for the speech is always 8 kHz. The input analog signal is band-limited by the band pass filter with the band of frequencies 300-3400 Hz before sampling. The amplitude has to be quantized in 256 levels into a binary representation. The representation of the amplitude can be either linear or logarithmic. In a linear representation, the step size is always the same between two consecutive samples. The total bit rate for raw speech with 8 kHz sampling rate and 16-bits quantization is 128 kbps. It can be reduced to 64 kbps by A-law/µ-law compression. The A-law technique is implemented by using the ADSP-BF533 Blackfin Processor. It is a 600 MHz high performance and low power embedded processor from an Analog Device which provides the greatest flexibility for signal processing applications. At receiving end, the ADSP is used to expand or decode this 64 KB data into 128 KB data. The 16-bit digital speech data is converted into analog signal by CODEC. The output of CODEC is almost equal to the amplitude of original speech signal.

A sample value is characterized by the segment number and its position or level within the segment. The A-law PCM companding technique has 8-bit word format which consists of three parts [13, 14]. The most significant bit is the sign bit, the next three bits representing the segment number, and the last four bits representing the position or level within the segment.

IV. A-LAW COMPANDING ALGORITHM

The block diagram of encoding and decoding process with the help of A-Law companding algorithm in digital telecommunication systems are shown in Figure 3 and Figure 4.

A. *Encoding Process*

The 16-bit uniform input from CODEC is pickup by data receive register. This 16-bit data is retained to 13-bit by mask upper 3 bits. The 13th bit is a sign bit which cannot be used in the encoder and it is directly given to the output. The remaining 12-bit magnitude is encoded by A-Law companding algorithm into 7-bit magnitude. As a result, the 8-bit word is obtained.

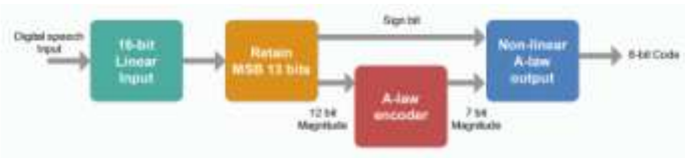

Figure 3. Encoding Process with A-Law Companding Algorithm

The step-by-step encoding algorithm for A-Law companding is described as follows,

1. The A-law is a 13-bit segment code and hence first, the 16-bit number is reduced into a 13-bit number by shifting it by 3-bits to the left.
2. The 13th bit is a sign bit. If it is zero, a number is a positive number and if it is one, the number is a negative number.
3. If it is a negative number, get 2's complement of that number and go to step 4.
4. The sign bit is extracted and hence the 13-bit number becomes a 12-bit number (Magnitude).
5. Count the number of leading zeros and hence to find the segment S, using the formula,
   S = 7 - Number of leading zeros.
6. If leading number of zeros is greater than 7, S=0.
7. Extract the index bits L depending on the value of segment S, using the formula [15,16],
   If S=0, drop the seven leading zeros and the following four bits are index bits.
   If S>=1, drop (7-S) + 1 MSB bits and the following four bits are index bits.
8. The compressed word having sign bit (7th bit), 3-Segment bits (6th-4th bits), and 4-Index/Level bits (3rd-0th bits).

B. *Decoding Process*

In decoding, the 8-bit word is divided into 1 sign bit (MSB) and 7-bit magnitude bits. This sign bit is directly taken at the output. The 7-bit magnitude is given to A-Law decoding algorithm which converts 7-bit into 12-bit magnitude. Finally, a 13-bit data is obtained including the sign bit.

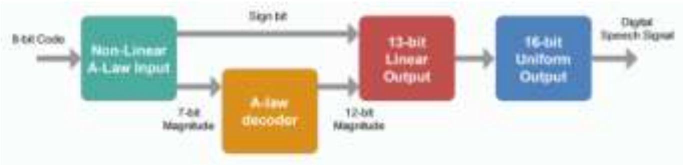

Figure 4. Decoding Process with A-Law Companding Algorithm

1. The sign bit is extracted and hence, the 8-bit number becomes a 7-bit number (Magnitude).
2. Get the segment S by shifting the input by 4-bits right.
3. Check if S=0 and go to step 4 otherwise go to step 5.
4. Find the magnitude from the decoded output using the formula [17, 18],

   $$(S,L)=2L+1,\ for\cdot S=0$$

   where S is the segment bit and L is the index bit.
5. Find the magnitude from the decoded output using the formula [17, 18],

   $$(S,L)=2S^{-1}\left(2L+33\right),\ for\cdot S\geq 0$$

   where S is the segment bit and L is the index bit.

V. CONCLUSION

A-Law companding algorithm has been implemented using the ADSP-BF533 Blackfin embedded processor on the speech signal. The companding is a compression technique which includes an encoding and decoding algorithm to reduce the data rate of audio signals. The cost of speech signal transmission depends on the data rate used in the digital telecommunication system. This algorithm requires minimum memory size and reduced overhead with register initializations that are executed at maximum speed. The signal amplitude and bandwidth are limited by A-Law companding. The source code of Blackfin processor for A-Law companding algorithm is made to minimize the tradeoff between required memory and cycle time, that is, companding by the direct implementation is achieved using 16 to 21 words requiring only 10 to 13 cycles. The DSP software provides the simplest and most efficient way to perform the companding. The execution speed of an encoding and decoding algorithms is 61 ns and 49 ns, respectively.

ACKNOWLEDGMENT

Authors are grateful to the head and faculty members of the Department of Electronics and Communication Systems, Sri Krishna Arts and Science College, Coimbatore 641008 for their valuable comments in the preparation of this manuscript. Authors are sincerely grateful to the authors of cited references.

REFERENCES

[1] Mark Hasegawa-Johnson, Abeer Alwan, Speech Coding: Fundamentals and Applications, Wiley Encyclopedia of Telecommunications, 2003.

[2] Brian B. Monson, Eric J. Hunter, Andrew J. Lotto, Brad H. Story, The perceptual significance of high-frequency energy in the human voice, Front Psychol 5 (2014) 587.

[3] Yusuke Hiwasaki, Hitoshi Ohmuro, ITU-T G.711.1: extending G.711 to higher-quality wideband speech, Communications Magazine IEEE 47 (2009) 110-116.

[4] Peter Birkholz ; Dietmar Jackel ; Bernd J. Kroger, Simulation of Losses Due to Turbulence in the Time-Varying Vocal System, IEEE Transactions on Audio, Speech, and Language Processing 15 (2007) 1218-1226.

[5] Ilse Lehiste, Gordon E. Peterson, Vowel Amplitude and Phonemic Stress in American English, Journal of the Acoustical Society of America 31 (2005) 428.

[6] Christopher A. Shera, John J. Guinan Jr., and Andrew J. Oxenham, Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements, PNAS 99 (2002) 3318-3323.

[7] G. Faria, J.A. Henriksson, E. Stare, P. Talmola, DVB-H: Digital Broadcast Services to Handheld Devices, Proceedings of the IEEE 94 (2006) 194-209.

[8] Jarek Krajewski, Anton Batliner, Martin Golz, Acoustic sleepiness detection: Framework and validation of a speech-adapted pattern recognition approach, Behavior Research Methods 14 (2009) 795-804.

[9] I. Kalet, J.E. Mazo, B.R. Saltzberg, The capacity of PCM voiceband channels, Proceedings of ICC '93 - IEEE International Conference on Communications, 2002.

[10] A.W. Rix, J.G. Beerends, M.P. Hollier, A.P. Hekstra, Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs, International Conference on Acoustics, Speech, and Signal Processing. IEEE Proceeding, 2002.

[11] A. M. Kondoz, Digital Speech: Coding for Low Bit Rate Communication Systems, John Wiley & Sons, 2005.

[12] Noboru Harada, Yutaka Kamamoto and et all, Emerging ITU-T standard G.711.0 — lossless compression of G.711 pulse code modulation, International Conference on Acoustics, Speech and Signal Processing, 2010.

[13] A.S. Spanias, Speech coding: a tutorial review, Proceedings of the IEEE 82 (1994) 1541-1582.

[14] Fraidoon Mazda, Telecommunications Engineer's Reference Book, Butterworth-Heinemann, 2014.

[15] Douglas A. Lyon, The μ -law CODEC", in Journal of Object Technology, 7 (2008) 17-31.

[16] Tomasi, Introduction to Data Communication and Networking, Pearson Education India, 2007.

[17] H. Kaneko, A unified formulation of segment companding laws and synthesis of codecs and digital compandors, The Bell System Technical Journal, 49 (1970) 1555-1588.

[18] Mohamed Konate, Peter Kabal, QUANTIZATION NOISE ESTIMATION FOR LOG-PCM, IEEE Canadian Conference on Electrical and Computer Engineering 2011.

AUTHOR'S BIOGRAPHIES

Dr. T. Veeramanikandasamy is an Assistant Professor of Electronics and Communication Systems at Sri Krishna Arts and Science College, Coimbatore 641 008, took his Doctorate degree at Bharathiar University, Coimbatore in 2016. His current research interests are in Nanomaterials Characterization, Embedded Systems, and Digital Signal Processing. He has published 8 research papers in international journals. He has presented more than 16 research papers in national/international conferences. He has delivered more than 10 technical lectures in various institutions.

Dr. M. Venkatachalam, presently working as Head and Associate Professor in Department of Electronics, Erode Arts and Science College (Autonomous), Erode – 638 009, India, has 30 years of teaching experience. He has published more than 90 research papers in National / International peer-reviewed journals and conferences. His research interests are in Thin Film Technology and Embedded Systems. He has undertaken minor and major research projects of U.G.C. He has also successfully organized various national and internal seminars and conferences. He has produced a number of M.Phil. and Ph.D. holders.