Speech Coding for Bluetooth with CVSD Algorithm

Othman O. Khalifa, S. Khan, MD Rafiqul Islam, Mohd. Bin Muktar and Zahrein Yaacob

International Islamic University Malaysia
Kulliyyah of Engineering
Electrical and Computer department
Jalan Gombak
E-mail: khalifa@iiu.edu.my

Abstract- Speech coding is a procedure used to represent a digital speech signal by as minimum bits as possible, maintaining at the same time a reasonable level of speech quality. Due to increasing demand for speech communications, speech coding technology has received a lot of researcher's interests. Speech coding is performed using numerous steps or operations specified as an algorithm. This paper presents the Continuously Variable Slope Delta (CVSD) speech coding method which is widely used in Bluetooth applications. It is a nonlinear sampled data feedback system which accepts a band limited analog signal and encodes it into binary form for transmission through a digital channel. At the receiver, the binary signal is decoded into a close approximation of original analog signal. This coding technique is implemented in Matlab Simulink and the simulation results shown in coding has good noise immunity. Further, shown in this work is the fact that damaged packets are discarded straight away without having a serious impact on the speech quality making it more suitable for two way audio communications.

1. Introduction

Digital speech Codecs are becoming the dominant technology in today's audio applications. With satellite radio, VoIP, cellular phones, and MP3 players, digital is replacing analog techniques in broadcast entertainment, telephony, and even personal audio recording [1][2]. Data bandwidth is one of the first places developers look when considering audio codecs. Human hearing spans a frequency range of approximately 5 Hz to 20 kHz, with a useful signal range of 120 dB from inaudible to painful [1][2][3]. Taking a brute-force digitization approach, this gives a Nyquist sampling rate of at least 40 kHz, and you need two audio channels for stereo. Compact disc (CD) audio, for instance, uses a 44.1 kHz sample rate with a 16-bit sample (96 dB range) for an audio data rate of 1.4 Mbits/sec to achieve stereo. To the raw data rate systems further overhead information added for encapsulation, channel coding, and error-correction coding where needed [7][12]. The fully-encoded digital data on CDs produces a bit stream at 4.3 Mbits/sec. In many applications, these data rates are much too high and some form of compression is required. Compression comes in two forms: lossless and lossy. Lossless compression exploits the statistical redundancy that occurs in many signals. Unfortunately, lossless compression is not enough to reduce audio data streams to manageable bit rates. There is not enough of a pattern to achieve much more than a factor of two in compression. As a result, audio codecs must turn to lossy compression techniques. A wide variety of digital audio coding standards have arisen over the last decade because of the range of tradeoffs that developers consider. Each application has its own optimal combination of data bandwidth, perceived sound quality, processing delay, and other design considerations, and each has evolved its own audio compression standard. Among these varying standards, however, they can be grouped in three major forms of audio coder/decoder (codec) approaches: direct audio coding, perceptual audio coding, and synthesis coding [2][12][13[17].

In direct audio coding, the main tradeoff is between bit rate and audio bandwidth. By first filtering the sound before encoding, then the Nyquist rate needed can be reduced for proper digitization [3][13]. Also it is possible to limit the dynamic range of the signal. Both limits are standard practice in telephony communications. Human speech puts most of its sound energy in the 200 Hz to 3.2 kHz frequency band. Further, even yelling has a volume level of only about 70 dB above inaudibility, so 12 bits is sufficient to capture conversations. By filtering the signal and sampling at 8 kHz, telephony systems thus reduce their bit rate to 96 kbits/sec. The resulting sound quality is adequate for a listener to recognize the person speaking by the sound of their voice and to catch the nuances of voice inflection, making it an acceptable level of sound quality for conversations. It is considered low quality for audio such as music, however. For instance, AM radio transmissions have a similar audio bandwidth [14][15][16][18].

2. Basic Concepts in Bluetooth

The objective of the Bluetooth standard is to enable seamless communications of data and voice over short-range wireless links between both mobile ad

stationary devices [4][6][7][8]. With its use of spread-spectrum technology, the first generation of the Bluetooth specification permits the secure exchange of data up to a rate of about 1 Mbps even in areas with significant electromagnetic activity and with its use of continuously variable slope delta modulation (CVSD) for voice encoding, the Bluetooth specification allows speech to be carried over short distances with minimal disruption. Bluetooth was designed to allow low bandwidth wireless connections to become so simple to use that they seamlessly integrate into your daily life. The Bluetooth specification is an open specification that is governed by the Bluetooth Special Interest Group (SIG) [9][10]]17]. Bluetooth is a specification to standardize on a low-cost, low-power Radio Frequency solution for wireless transmission between a wide variety of devices such as PCs, keyboards, cordless telephones, cell phones, headsets, printers, monitors, LCD projectors, and Personal Digital Assistants [7][8].

Bluetooth was intended to create a single digital wireless protocol to address end-user problems arising from the proliferation of various mobile devices that need to keep data synchronized (i.e., consistent from one device to another). It operates on both a point-to-point and a point-to-multipoint basis. Although it primarily is promoted as a means of eliminating cables between devices, Bluetooth also has limited LAN potential in the form of a Personal Area Network (PAN). Bluetooth operates in the 2.4 GHz band and makes use of the Gaussian Frequency Shift Keying modulation scheme and Frequency Hopping Spread Spectrum [5][6][7]. The nominal link range is as much as 100 meters, although most devices will be limited to 10 meters, and the theoretical gross data rate is up to 1 Mbps, under optimal conditions [6][11].

3. Continuously Variable Slope Delta modulation

Delta modulation (DM) and Continuously Variable Slope Delta modulation (CVSD) are differential waveform quantization techniques [3][6][17]. Both employ two level quantizers (one bit). CVSD is basically DM with an adaptive quantizer. Applying adaptive techniques to a DM quantizer allows for continuous step size adjustment. By adjusting the quantization step size, the coder is able to represent low amplitude signals with greater accuracy (where it is needed) without sacrificing performance on large amplitude signals.

CVSD is used in tactical communications where communication quality is required yet the option for security must be available. MIL-STD-188-113 (16 Kb/s and 32 Kb/s), and Federal Standard 1023 (12 Kb/s CVSD) are examples of a tactical communication systems using CVSD. With the tremendous worldwide growth in wireless technology, secure communication is becoming important to everyone. In addition to point

to point communication, CVSD is commonly used in digital voice recording/messaging and audio delay line [14][17].

4. Continuous Variable Slope Delta Modulation Speech Coding

The speech coders are specified in the standard, either a 64kb/s log PCM format A-law or m-law, or a 64 kb/s Continuous Variable Slope Delta Modulation is used. The latter Continuous Variable Slope Delta Modulation is used. It is implemented in this model. A Wav-read block first brings in an 8 kHz signal from the wave file. It then interpolates the signal up to 64 kHz ready for speech coding. The CVSD speech coder encodes each sample with a single bit using a differential coding scheme, which transmits a 1 bit if the speech sample increases from the previous value and a 0 bit if it decreases [5][17]. This scheme is very robust to bit errors. If a bit is corrupted in transmission, then the decoded speech at the receiver will only be in error by a small fraction of the total amplitude range. Here, the CVSD block is constructed under the Bluetooth specification standard. The output of the coder is a 64 kbps stream, which is buffered into frames of 240 bits that are equivalent to 3.75ms of speech (HV3 packet type).

5. Results and Conclusions

Three types of packets are implemented and simulated in AWGN and Rayleigh fading channels. In AWGN, the three packet types show practically no differences but, however, in Rayleigh channel there is some dBs lag. It can be seen that HV1 packet type transmission generates more errors for the same BER as compared to an HV3 type. CVSD has several attributes that make it well suited for digital coding of speech. One bit words eliminate the need for complex framing schemes. Robust performance in the presence of bit errors makes error detection and correction hardware unnecessary. Other speech coding schemes may require a digital signal processing engine and external analog to digital/digital to analog converters to convert the analog signal in to a form that can be processed digitally the entire CVSD codec algorithm, including input and output filters, can be integrated on a single silicon substrate. Despite this simplicity, CVSD has enough flexibility to allow digital encryption for secure applications. Finally, CVSD can operate over a wide range of data rates—it has been successfully used from 9.6 kB/s to 64Kb/s. At 9.6Kb/s audio quality is not particularly good, however, it is intelligible. At data rates of 24 Kb/s to 48 Kb/s it is judged as quite acceptable. And above 48 Kb/s it is comparable to toll quality. All of these attributes make CVSD attractive to wireless communication systems. The defiance industry has been using CVSD for decades in wire line and wireless. Bluetooth is a short-

distance wireless technology that is optimized for low cost and low power consumption. In many ways you can think of Bluetooth as a cable replacement technology that eliminates the need for numerous and inconvenient cable attachments for connecting mobile devices. Because of its simplicity and low power usage, Bluetooth is ideal for devices like cell phones and digital cameras.

6. References

[1] J. C. Bellamy, *Digital Telephony*, Wiley and Sons, New York, 1982.

[2] J. A. Greefkes and K. Riemens, "Code Modulation with Digitally Controlled Companding for Speech Transmission," *Philips Tech. Rev.*, pp. 335-353, 1970.

[3] A. Gersho, "Principals of Quantization," *IEEE Transactions on Circuits and Systems*, pp. 427-436, July 1978.

[4] Rene Nusser, Bluetooth-based Wireless Connectivity in an Automotive Environment, *IEEE 2000*.

[5] K. Sairam, Bluetooth in Wireless Communication, *IEEE Communications Magazine* June 2002.

[6] C. Bisdikian, An overview of the Bluetooth Wireless Technology, *IEEE Communications Magazine* Dec 2001.

[7] Robert E. Van Dyck, *Wireless Personal Area Networks*, whitepaper on Bluetooth July 2001.

[8] R.Shepherd, Bluetooth wireless technology in the home, *E. & Communication Engineering Journal*, October 2001.

[9] Ivan Howitt, *IEEE 802.11 and Bluetooth Coexistence analysis methodology*, EECS 2002.

[10] A. Soltanian, Performance of the Bluetooth System in Fading Dispersive Channels and Interference, *NIST* Aug 99.

[11] S. Zurbes, " *Radio Network Performance of Bluetooth*," *IEEE* 2000.

[12] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice-Hall, Englewood Cliffs, N. J., 1984.

[13] A. B. Jerri, "The Shannon Sampling Theorem— Its Various Extensions and applications: A Tutorial Review," *Proceedings of the IEEE*, pp. 1565-1596, November 1977.

[14] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, N. J., 1978.

[15] M. Schwartz, *Information, Transmission, Modulation, and Noise*, McGraw Hill, New York, 1980.

[16] R. Steele, *Delta Modulation Systems*, Pentech Press, London, England, 1975.

[17]http://www.mathworks.co.uk/company/newsletters /digest/nov01/bluetooth_pt2.html