

Fixed Point Implementation of 16 Kbps Modified CVSD algorithm

Nitin Tandon
DRDO Dehradun
nitin.tandon@deal.drdo.in

Vikas Bhatia
DRDO Dehradun
renurani@deal.drdo.in

LC Mangal
DRDO Dehradun
lcmangal@deal.drdo.in

Abstract—Continuously Variable Slope Delta (CVSD) modulation has been extensively used in wireless environments and it has been adopted by Bluetooth also. In particular, CVSD is especially suitable for Internet and mobile environments due to its robustness against transmission errors, absence of a need for synchronization and simplicity of implementation. A modified CVSD (MCVSD) scheme proposed in [5], provides better performance compared to CVSD for low speech levels and during silence period besides providing better background noise rejection. The performance of modified CVSD algorithm is comparable to CVSD for higher signal level thus providing improvement in overall dynamic range.

A fixed point implementation of 16 Kbps MCVSD speech-coding algorithm on C64x+ digital signal processor of TMS320DM6446 System-on-Chip is described in this paper.

Keywords— CVSD, slope overload, granular, quantization, syllabic, MCVSD, reconstruction integrator, SNR, EDMA, McBSP.

I. INTRODUCTION

Nowadays, digital communication has become indispensable and is being used worldwide. Digital representation of information has certain advantages over analog such as better noise immunity, provision of encryption, storage, multiplexing and many more. Speech signal extends from 300 Hz to 3300 Hz and if sampled just above Nyquist rate at 8 KHz with 8 bit/sample, results in 64 Kbps data rate. It incurs bandwidth requirement nearly 20 times the original analog signal and this motivates the use of speech compression or coding [1]. It reduces data rate and hence resulting transmitted bandwidth.

As human ears are more susceptible to lower amplitude speech signals, one requires a quantization strategy which uses smaller quantization levels for lower amplitude signals and fewer, coarser quantization levels for larger amplitude signals. This is known as non-uniform quantization [2]. Another approach leverages advantage of correlation between adjacent speech samples, and thus quantizing the amplitude difference between adjacent samples as opposed to the entire sample amplitude. It requires fewer quantization levels for the same signal quality which eventually leads to bandwidth reduction. Algorithms employing this technique are classified under the broad category of differential quantization or differential PCM (DPCM). Even more bandwidth reduction is made feasible by combining adaptive quantization with DPCM, generally referred as adaptive DPCM (ADPCM). Both the Delta modulation (DM) and Continuously Variable Slope Delta modulation (CVSD) are differential waveform quantization techniques employing two level quantizers (one bit) [1]. CVSD is basically DM with an adaptive quantizer.

Continuous quantization step size adjustment can be performed by applying adaptive techniques to a DM quantizer. 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 and thus manages slope overload and granular noise issues commonly encountered in DM.

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 of CVSD in the presence of bit errors reduces the role of error detection and correction hardware. Also, CVSD can operate over a wide range of data rates. CVSD is widely employed in tactical communications systems where “communication quality speech” is required along with the option for security [1]. It has been used in several systems such as MIL-STD-188-113 (16 Kb/s and 32 Kb/s), and Federal Standard 1023 (12 Kb/s). Software implementation of CVSD can be employed in Software Defined Radio (SDR) based communications where the basic requirement is to implement all sub module of radio like source coding, channel coding, modem, etc. in software to make it reconfigurable, upgradable, reusable and portable.

In this paper, we discuss the DSP implementation of modified CVSD (MCVSD) algorithm which provides better performance compared to CVSD for low speech levels and during silence period and provides better background noise rejection at the expense of slightly degraded performance at higher signal levels [5].

The paper is organized as follows. In Section II, the basics of CVSD and MCVSD schemes are described. Simulation of MCVSD scheme and its comparison with CVSD is explained in Section III. DSP hardware details are given in Section IV. Implementation details of MCVSD scheme and results are given in Section V. Finally concluding remarks are given in Section VI.

II. MCVSD BASICS

Block Diagram of MCVSD scheme is shown in Fig.1. Common blocks employed in MCVSD and CVSD schemes include comparator, syllabic integrator and reconstruction integrator. MCVSD scheme contains an additional block of energy detector. The comparator compares the input speech signal with feedback approximation signal generated at the reconstruction integrator output. This comparison produces the digital error signal which is provided to the 3-bit shift register. The slope overload detection algorithm operates on the output

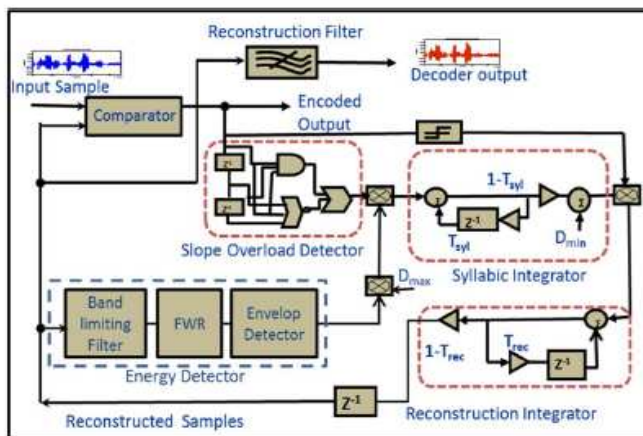


Fig.1. Block Diagram of MCVSD Codec

of the 3-bit shift register using the run-of-three coincidence algorithm. For clock rates of 16 kHz and below, the 3-bit algorithm is well suited. For clock rates greater than 16 KHz, the 4-bit algorithm is preferred.

The syllabic filter acts as a low-pass filter for the output signal from the overload algorithm. The step-function response of the syllabic filter is related to the syllabic rate of speech, is independent of the sampling rate, and is exponential in nature. When the overload algorithm output is true, a charging curve is applicable. When this output is false, a discharging curve is applicable. The output signal from the syllabic filter and the digital signal from the comparator (± 1) are multiplied and fed to the reconstruction integrator. The syllabic filter output signal determines the amplitude of the reconstruction integrator input signal and the digital signal from the comparator is the polarity control that determines the direction. The reconstruction integrator produces an analog feedback signal to the comparator that is an approximation of the analog input signal. Typical values for the syllabic and reconstruction integrator time constants are 5 to 12 ms and 0.5 to 1.5 milliseconds, respectively. D_{min} and D_{max} are chosen depending on the dynamic range, the maximum frequency and sampling frequency of the input signal. A low pass reconstruction filter at the receiving circuit output eliminates most of the quantizing noise.

The significant feature of MCVSD is the energy detector block which uses reconstructed output power level for changing the exciting input to the syllabic integrator. This provides a significant improvement in the dynamic range. This is because at lower signal levels, signal envelope controls (provides a bearing) on maximum step size and thus better tracking of input slope gets possible.

III. SIMULATION

MCVSD algorithm has been simulated using Matlab to verify its performance. First floating point simulation is performed, then fixed point simulation of algorithm is carried out and

results compared. Ideally, all algorithms should be implemented with floating-point processors; in that way the rounding error after each operation will be very small, and there is no issue of numeric overflow. But floating-point processors are relatively expensive, due to the increased size of the processor's chip needed to operation support the more complex operations; also, power consumption is higher when compared to a fixed-point processor. For cost and power sensitive consumer appliances (e.g., the cellular handset), the fixed-point processor is almost essential.

Fig.2. depicts comparative performance of CVSD and MCVSD scheme for speech signal. The figure of merit used is the SNR at the decoder output. From the figure, it is clear that MCVSD scheme has better SNR performance even at low speech signal levels and it has only slightly less SNR for higher speech signals. Thus it provides an overall improvement in dynamic range.

Fig.3. shows time domain plot of the performance improvement achieved by MCVSD for low level speech signal. MCVSD scheme provides better performance compared to CVSD for low speech levels at the expense of slightly poor SNR at higher signal levels. Henceforth, it results in improvement in overall dynamic range.

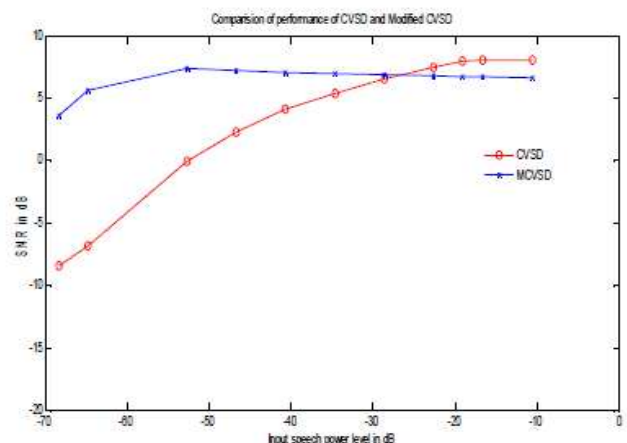


Fig.2. Comparison of output SNR performance of CVSD and MCVSD.

Although SNR is the most used method to objectively quantify performance of speech coding algorithms, segmental SNR (SSNR) is considered a better perceptual model since it evaluates the quantization noise with respect to signal energy in each underlying speech segment. SSNR measure is used to cater for dynamic nature of non-stationary signals such as speech. SSNR is the average of SNR values obtained for isolated frames, where the frames are block of samples. It is defined as

$$SSNR = \frac{1}{N} \sum_{m=1}^N SNR_m$$

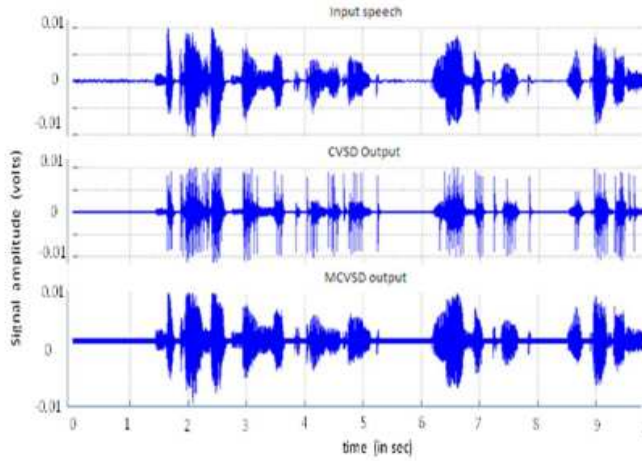


Fig.3. Time domain plot of Input speech, CVSD and MCVSD output

This measure compensates for under emphasis of weak signal performance in conventional SNR measure. SSNR value is compared for floating point and fixed point simulations taking 60 segments with segment size of 160. The value of SSNR is found comparable for floating and fixed point cases with MCVSD having better performance.

Table 1. Comparison of segmental SNR for CVSD and MCVSD

Parameter	Floating Point (CVSD)	Floating Point (MCVSD)	Fixed Point (CVSD)	Fixed Point (MCVSD)
Segmental SNR	8.43	11.2	7.63	10.6

PESQ (Perceptual Evaluation of Speech Quality) is an ITU algorithm, widely used to emulate Mean opinion Score (MOS), which is a listening test for speech codec [4]. The objective is to automate and ensure reproducibility of measurement of and obtain results that closely correlate with MOS human listening tests. We have utilized the PESQ software from ITU for evaluating PESQ.

PESQ values can be used to predict MOS quiet accurately using the mapping function given below:

$$y = 0.999 + \frac{(4.999 - 0.999)}{(1 + \exp(-1.4945 * x + 4.6607))}$$

where x =PESQ value, y = Predicted MOS value.

Table2. PESQ value and predicted MOS results of MCVSD scheme

Tests (for different input speeches)	Floating Point Model (PESQ/Predicted MOS)	Fixed Point Model (PESQ/Predicted MOS)
Test1	2.985 / 2.80	2.613 / 2.28
Test2	2.900 / 2.68	2.418 / 2.05
Test3	3.102 / 2.98	3.017 / 2.85

IV. HARDWARE DETAILS

MCVSD algorithm has been implemented on Texas Instrument C64x+ DSP Processor. The C64x+ is a fixed point DSP primarily intended for audio and video applications [3].

A simplified block diagram of the DSP internal architecture is shown in Fig.4. C64x+ is a VLIW processor with clock rate of 594 MHz. It has 32 general-purpose registers of 32-bit word length and eight highly independent functional Units; two multipliers for a 32-bit result and six arithmetic logic units. The C64x+ uses a two-level cache-based architecture and has a powerful and diverse set of peripherals. The Level 1 Program cache (L1P) is a 32-KB direct mapped cache and the Level 1 data (L1D) memory of 80-KB with 32-KB 2-way set associative cache. The Level 2 memory/cache (L2) consists of 64 KB memory space that is shared between program and data space. L2 memory can be configured as mapped memory, cache, or a combination of both. The peripheral set includes (Multichannel Buffered Bidirectional serial Port) McBSP, general-purpose timers; I^2C , GPIO and SPI interface.

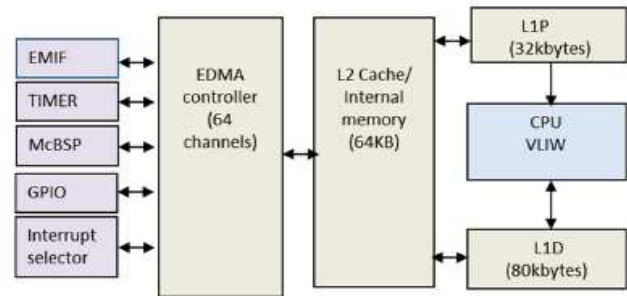


Fig. 4. Internal architecture of C64x+ DSP

Enhanced Direct Memory Access (EDMA) controller relieves the DSP from the duty of data transfer. The EDMA controller takes incoming audio data directly from ASP and places it in a memory buffer. It also takes data from a memory buffer and sends it to McBSP to generate the audio output. Separate EDMA channels are used to transmit and receive audio data.

V. HARDWARE IMPLEMENTATION

Implementation scheme of MCVSD codec on C64x+ DSP Processor is shown in Fig.5. Implementation utilizes McBSP

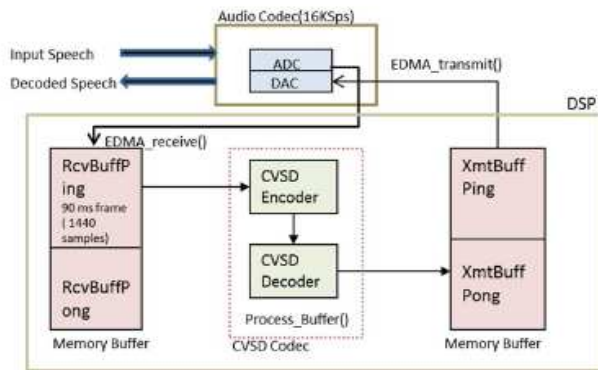


Fig.5. Implementation Scheme of MCVSD Codec

and EDMA to efficiently handle the data transfer without intervention from the DSP. Audio data is transferred back and forth from the codec through McBSP, a bidirectional serial port at a rate of 16 KHz. The EDMA controller takes incoming audio data directly from McBSP and places it in a memory buffer. It also takes data from a memory buffer and sends it to McBSP to generate the audio output. Separate EDMA channels are used to transmit and receive audio data. Ping-Pong buffers based data transfer and EDMA linked transfers are explained in following subsections.

A. Ping-Pong Transfers

At the highest level, the EDMA controller reads audio data from the McBSP port and places it in a buffer in memory. On the data receive side there are two logical buffers, receive PING and receive PONG. When the first data comes in, it is placed in the PING buffer. When it is full, new data is redirected to the PONG buffer and the DSP is free to process the PING data without fear of it being overwritten. When the PONG buffer fills up, the configuration is reversed. The ping-ponging data transfer continues indefinitely with one buffer always hosting the active transfer and one remaining stable for the DSP to operate on. If only one buffer is used, the DSP must process all of the data between the instant the buffer fills up and the next audio sample arrives. When ping-pong buffers are used, the DSP can take as long as the time it takes to fill an entire buffer to process the data, making it much easier to meet a real-time schedule.

Separate input and output buffers are used to decouple the receive side from the transmit side while data is being processed. The output buffers are also of the Ping-Pong variety so there are a total of four logical buffers, receive PING, receive PONG, transmit PING and transmit PONG.

B. EDMA Linked Transfers

EDMA linked transfer is a configurable inbuilt feature in C64x+ DSP. When the EDMA finishes with the PING side and needs to switch to the PONG side, the source and destination pointers need to be changed to point at the new buffer. When linked transfers are used, the new address can be stored in a link configuration structure and automatically loaded by the EDMA controller when the current transfer is complete. The reconfiguration can also be handled by the DSP in software interrupt service routine but using linked transfers removes the scheduling requirement that the software finish before the next sample is received to get uninterrupted audio.

Input speech frame of 90 ms (corresponds to 1440 speech samples) has been taken. The performance of MCVSD Codec has been listed below:

Table 1. MCVSD Performance results on C64x+ DSP

Codec	Execution Time (in msec)	PM (in KB)	DM (in KB)
MCVSD	23 msec	14	34

VI. CONCLUSION

Simulation of MCVSD algorithm is carried out in both floating and fixed point. The results of SNR vs. signal power, segmental SNR and PESQ measures are computed for CVSD as well as MCVSD and compared. Results clearly indicate the advantages of MCVSD over CVSD. Fixed point implementation results are having acceptable performance loss compared to floating point case. Fixed point implementation of MCVSD is carried out on Texas Instruments C64x+ DSP and implementation scheme have been discussed.

ACKNOWLEDGMENTS

The authors are thankful to Director, Defence Electronics Applications Laboratory, Dehradun for permission to publish the paper.

REFERENCES

- [1]. "Continuously Variable Slope Delta Modulation: A Tutorial", Application Doc. #20830070.001, MX-COM, Inc., Winston-Salem, North Carolina, 1997.
- [2]. A. Spanias, "Speech coding: a tutorial review," Proc. IEEE, vol. 82, pp 1541-1582, 1994.
- [3]. Data Sheet and user guide of TMS320DM6446 SoC.
- [4]. Antony Rix "End to End speech quality assessment of network using PESQ (P.862)" ITU-T SG12 workshop, oct,2001, © Psytechnics Ltd.
- [5]. V. D. Mytri, A. P. Shivaprasad, "Improving the dynamic range of a CVSD coder", Electronics Letters 10th April 1986 Vol. 22 No. 8.