International Journal of Emerging Engineering Research and Technology
Volume 2, Issue 2, May 2014, PP 91-95

Implementation of Speech Compression Using Linear Predictive Coding (LPC) with Tms320c6713dsk and Comparison with Other Platforms

Uday Mithapelli , Suraj P. Patil
Dept. of E & Tc, KJEI's Trinity College of Engineering & Research,
Pune, India

Abstract: *In recent multimedia system, people consider speed, power, and size very important. People want the process to have very high performance and high quality, while consume only small amount of power. Beside the development on better algorithm, special hardware design techniques can also help in increasing performance. This project mainly focuses on the linear predictive coding algorithm on speech. The algorithm is analyzed and implemented on DSP processor TMS320C6713 .*

The implementation is done using Code Composer Studio(CCS) for which we also have a chance to evaluate the effectiveness of the software in implementing the encoder from a hardware level perspective. This LPC coding is having wide application in wireless communication and also used for accurate estimation of Speech Properties.

Keywords: *Linear Predictive Coding(LPC), DSP processor TMS320C6713,Code Composer Studio(CCS 3.1),Segmental Signal To Noise Ratio (SEGSNR) and Power Signal To Noise Ratio (PSNR)MATLAB platform.*

1. Introduction

Linear Predictive Coding (LPC) is one of the methods of compression that models the process of speech production. Specifically, LPC models this process as a linear sum of earlier samples using a digital filter inputting an excitement signal.

An alternate explanation is that linear prediction filters attempt to predict future values of the input signal based on past signals. LPC "...models speech as an autoregressive process, and sends the parameters of the process as opposed to sending the speech itself" . It was first proposed as a method for encoding human speech by the United States Department of Defense in federal standard 1015, published in 1984. Another name for federal standard 1015 is LPC-10 which is the method of Linear predictive coding that will be described in this paper.[1]

Speech coding or compression is usually conducted with the use of voice coders or vocoders. There are two types of voice coders: waveform-following coders and model-base coders. Waveform following coders will exactly reproduce the original speech signal if no quantization errors occur. Model-based coders will never exactly reproduce the original speech signal, regardless of the presence of quantization errors, because they use a parametric model of speech production which involves encoding and transmitting the parameters not the signal. LPC vocoders are considered model-based coders which means that LPC coding is lossy even if no quantization errors occur.

Speech coder that is developed analyzed using both subjective and objective analysis. Subjective analysis will consist of listening to the encoded speech signal and making a judgment on its quality. The quality of played back signal will solely depends on opinion of listener. Objective analysis will be introduce to technical access the speech quality and to minimize human bias.[4]

The Objective Analysis will be performed by computing between the original and coded speech signal . Also speed of operating platform in terms of second is also analyzed.

2. Key Features and Basic Operation of 6713 Processor

This project is implemented on TMS3206713 DSK kit. A Texas Instruments TMS 320C6713 DSP operates at 225 MHz It has 16 Mbytes of synchronous DRAM

and512 Kbytes of non-volatile flash memory (256 Kbytes usable in default configuration). Digital signal Processing is one of the most powerful technologies that will shape science and engineering in the twenty first century.

Revolutionary changes have already been made in a broad range of fields: communications, medical imaging, radar & sonar, high fidelity music reproduction, and oil prospecting, to name just a few. Each of these areas has developed a deep DSP technology, with its own algorithms, mathematics, and specialized techniques. The DSK is designed to work with TI's Code Composer Studio development environment and ships with a version specifically tailored to work with the board. Code Composer communicates with the board through the on-board JTAG emulator or USB connector.

3. LPC IN GENERAL

LPC technique will be utilized in order to analyze and synthesize speech signals.

This method is used to successfully estimate basic speech parameters like pitch, formants and spectra. A block diagram of an LPC vocoder can be seen in Figure 3.1. The principle behind the use of LPC is to minimize the sum of the squared differences between the original speech signal and the estimated speech signal over a finite duration. This could be used to give a unique set of predictor coefficients [3].

For LPC-10 P = 10. This means that only the ten coefficients of the predictor are transmitted to the LPC synthesizer.

The two most commonly used methods to compute the coefficients are the covariance and the autocorrelation methods. In this study, the

autocorrelation formula is preferred because it is superior to the covariance method in the sense that the roots of the polynomial in the denominator of the above equation is always guaranteed to be inside the unit circle, hence guaranteeing the stability of the system, H(z).

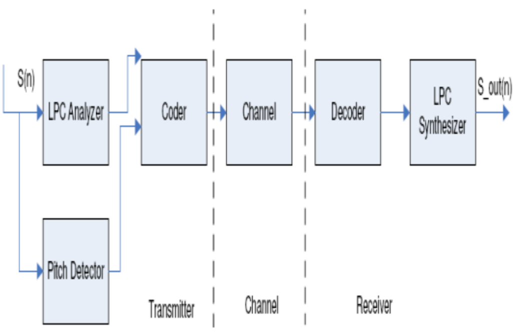
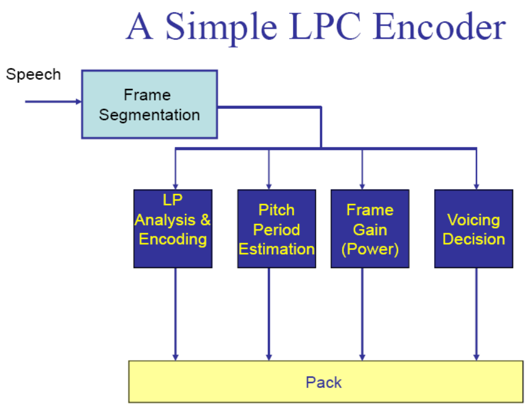


Fig: 3.1

4. LPC ENCODING

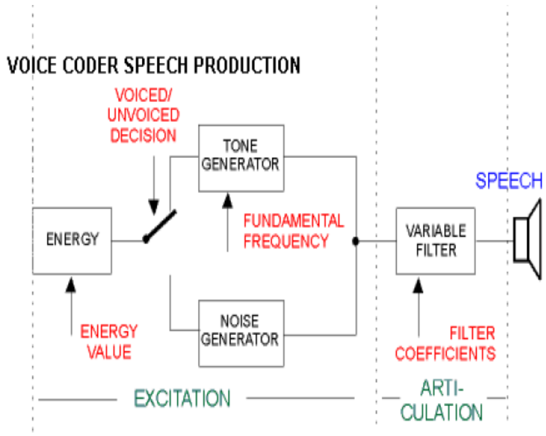
These predictor coefficients are normally estimated every frame, which is normally 20 ms long. The predictor coefficients are represented by ak where k = {1, 2, . . . , P} and P is the predictor order. Another important parameter is the gain, G. The transfer function of the time-varying digital filter is given by:

$$H(z) = \frac{G}{1 - \sum_{k=1}^{P} a_k z^{-k}}$$



5. LPC SYNTHESIS/DECODING

The process of decoding a sequence of speech segments is the reverse of the encoding process. Each segment is decoded individually and the sequence of reproduced sound segments is joined together to represent the entire input speech signal. The decoding or synthesis of a speech segment is based on the 54 bits of information that are transmitted from the encoder. The speech signal is declared voiced or unvoiced based on the voiced/unvoiced determination bit. The decoder needs to know what type of signal the segment contains in order to determine what type of excitement signal will be given to the LPC filter. Unlike other speech compression algorithms[3].



This combination of voice/unvoiced determination and pitch period are the only things that are need to produce the excitement signal. Each segment of speech has a different LPC filter that is eventually produced using the reflection coefficients and the gain that are received from the encoder. 10 reflection coefficients are used for voiced segment filters and 4 reflection coefficients are used for unvoiced segments. These reflection coefficients are used to generate the vocal tract coefficients or parameters which are used to create the filter. The final step of decoding a segment of speech is to pass the excitement signal through the filter to produce the synthesized speech signal.

6. IMPLEMENTATION PLATFORM

6.1 Code Composer Studio.(CCS)

CCS provides an IDE to incorporate the software tools. CCS includes tools for code generation, such as a C compiler, an assembler, and a linker. It has graphical capabilities and supports realtime debugging. It provides an easy-to-use software tool to build and debug programs.

The C compiler compiles a C source program with extension .c to produce an assembly source file with extension.*asm*.The assembler assembles an.*asm* source file to produce a machine language object file with extension.*obj*. The linker combines object files and object libraries as input to produce an executable file with extension.*out*. This executable file can be loaded and run directly on the C6713 processor. A linear optimizer optimizes this source file to create an assembly file with extension .asm (similar to the task of the C compiler).

A number of debugging features are available, including setting breakpoints and watching variables; viewing memory, registers, and mixed and assembly code; graphing results; and monitoring execution time. Key statistics and performance can be monitored in real time. Through the joint team action group (JTAG), communication with on-chip emulation support occurs to control and monitor program execution. The C6713 DSK board includes a JTAG interface through the USB port.

6.2 Mat lab

Same Algorithm is also implemented in Matlab. Matlab is very efficient tool for such type of processing .It has given a validation of algorithm which we are going use for actual DSP implementation at simulation level.

7. PERFORMANCE ANALYSIS

7.1 Throughput

The throughput of this implementation can be calculated based on the input rate. Assume input rate is 8Ksamples/s, and 8bit /samples; thus input rate is equal to 64Kbps. Since a non-overlapping window contains 180 samples, the LPC coder will give a group of code every (8Ksamples/s) /(180). The output of the LPC coder is “pitch(n)”7 bit, “gain (5 bits)” a voices unvoiced decision (1 bit) for synchronization (1 bit) and LPC Coefficient(40bits). Therefore, The total throughput of the output is

(8KS/s) / 180 * [54 bits] =2.4Kbps.

7.2 SNR Measurement

To find the implementation quality, we will usesegmented SNR. The measure is defined as follows

$$SegSNR = \frac{1}{NrFrames} \times \sum_{i=1}^{NrFrames} SNR(i)$$

And

$$SNR(i) = 10 \times \log\left(\frac{\sum_{n=1}^{NSF} s_i(i)^2}{\sum_{n=1}^{NSF} \left(s_i(n) - s_p(n)\right)^2} \right)$$

With $S_i(n)$ the input speech samples and $S_p(n)$ the synthesized output speech samples. First, we calculate SegSNR for the speech coder algorithm. This compares the input speech samples si(n) with floating point output samples sp(n)[6].Since we found that the average value of the original voice is different from the average value of the synthesized floating-point implementation, we thus first normalize all two samples so that they have equal amplitude.

After normalizing, we found that:

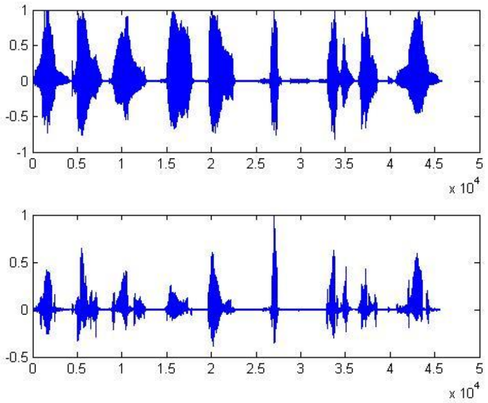
SegSNR_A = 0.1752

These means that the “noise”, which means the difference of two samples here, is smaller than the original signal.

This compares the floating-point output samples sp(n) with the fixed-point output samples.

We found that: SegSNR_Q = -1.3065 It seems very bad, but when we listen to the real samples, we could still understand the synthesized output.Second, we calculate SegSNR_Q for the speech coder implementation

Therefore, we can conclude that the fixed-point implementation did not distort the floating implementation by a large amount. The signal waveform is shown in Fig



The top waveform is the original voice, and the second one is the floating-point implementation. We see that there is some difference between the floating point and the original.

7.3 Power Signal To Noise Ratio(PSNR)

P=10 log10 {{Max[A]/MSE}}

Where A is samples of original signal[6].

Power Signal TO Noise Ratio Compares level of desired signal to the background noise. DSP implemented LPC sound is far better than matlab output although PSNR of DSP implemented algorithm gives barely positive PSNR =40 db DSP 6713

Matlab(PSNR)=38.6867db

7.4 Calculation of Execution Time

The Texas Instruments TMS320C6000 DSP

platform of high-performance digital signal processors (DSPs) now includes the TMS320C6713. The C6713 brings the highest level of performance in the C6000 DSP platform of floating-point DSPs. At the initial clock rate of 225 MHz, the C6713 can process information at a rate of 1.35 giga - floating-point operations per second

(GFLOPS).

Total Execution time of LPC=

(Clock frequency x the number of clock cycles taken to execute an algorithm)

Where, the clock frequency depends on the DSP board we are using. Here we are using TMS320C6713.The clock frequency of this TMS320C6713 DSP processor is 1/225 MHz in result Table the execution time for TMS320C6713 DSP processor and MATLAB have been indicated.

8. Results

Results for LPC compression of speech for both with DSP 6713 platform

and MATLAB for above parameters are noted in the result table shown below.

Parameters	TMS320C6713	MATLAB(Core i5- intel processor)
Execution Time(sec)	0.013312	1.846565
Seg SNR	0.2133	-1.3065
PSNR	40.1434	38.6249

9. Conclusion

This project, we used Code Composer Studio to implement a LPC coder which is then compared with implementation result of Mat lab on core i-5 Intel Processor.

It is observed that speed of execution in TI320C6713 DSP processor is very fast as compared with the speed of execution of MATLAB on I5 Intel processor. Also due to floating point operation on DSP Accuracy enhances and an intangibility of speech signal is maintained. While in case of MATLAB Accuracy is less and Speech Signal is less intelligible. Hence quality of DSP processor is far better . To ensure the quality of the speech implementation, Segmented SNR was used to define the quality of this coder. This concluded LPC coder implementation. Enhances and an intelligibility of speech signal is maintained. While in case of MATLAB Accuracy is less and Speech Signal is less intelligible. Hence quality of DSP processor is far better . To ensure the quality of the speech implementation, Segmented SNR was used to define the quality of this coder. This concluded LPC coder implementation.

References

[1] L. R. Rabiner, R. W. Schafer, “Digital Processing of Speech Signals.” Prentice Hall, Englewood Cliffs, New Jersey,

[2] http://www.datacompression.com/speech.html

[3] B. S. Atal, M. R. Schroeder, and V. Stover, “Voice- Excited Predictive Coding Systetm for Low Bit-Rate Transmission of Speech”, Proc. ICC, pp.30-37 to 30-40

[4] The newest breeds trade off speed,energy consumption, and cost to vie for an ever bigger piece of the action. BY JENNIFER EYRE Berkeley Design Technology Inc.

[5] Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs ITU-P862

[6] SPEECH COMPRESSION USING LINEAR PREDICTIVE CODING(LPC) Nikhil Sharma* International Journal of Advanced Research in Engineering and Applied Sciences ISSN: 2278-6252

[7] A 150KHz Linear Predictive Coding (LPC) Special Purpose Processor for Speech Processing Henry Kuo, Koon Lun Jackie Wong , Engineering Department, University of California, Los Angeles

[8] National Program on Technology Enhanced Learning (NPTEL) open Courseware(By Prof. Sengupta)

[9] The BDTImark2000™:A Measure of DSP Execution Speed Berkeley Design Technology, Inc.

[10] http://www.expertslinked.com/content/dsp-or-fpga-%E2%80%93-5-parameters-make-choice 5 point comparison for processors

AUTHOR'S BIOGRAPHY

Uday P Mithapelli received Bachelors of Engineering degree in 2010, from the University of Pune, Department of Electronics & Telecommunication Engineering. He received his MTech degree in 2013, from Vishwakarma Institute of Technology, Pune India.

Currently he is working as Assistant Professor in KJEI's Trinity College of Engineering & Research, Pune, India. His main interests lie in speech synthesis, coding, and recognition

Suraj Patil received Bachelors of Engineering degree in 2010, from the University of Pune, Department of Electronics & Telecommunication Engineering. He received his MTech degree in 2013, from Vishwakarma Institute of Technology, Pune India.

Currently he is working as Assistant Professor in KJEI's Trinity College of Engineering & Research, Pune, India. His main interests lie in speech synthesis, coding, and recognition.