



Faculty of Electronics and Computer Engineering

REAL-TIME IMPLEMENTATION OF LPC-10 CODEC ON TMS320C6713 DSP

Wissam Talib Alshammari

Master of Electronic Engineering (Telecommunication System)

2012

**REAL-TIME IMPLEMENTATION OF LPC-10 CODEC ON
TMS320C6713 DSP**

WISSAM TALIB ALSHAMMARI

A thesis submitted

**In fulfillment of the requirements for the degree of Master of Electronic Engineering
(Telecommunication System)**

Faculty of Electronics and Computer Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2012

DECLARATION

I declare that this thesis entitle “Real-Time Implementation of LPC-10 Codec on TMS320C6713 DSP” is the result of my own research excepted as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature:

Name: Wissam Talib Alshammari

Date:

APPROVAL

“I hereby declare that I have read through this project report, and in my opinion this thesis is sufficient in terms of scope and quality for the award for the degree Master of Electronic Engineering (Telecommunication System)”.

Signature:

Professor Dr. Muhammad Ghazie Bin Ismail

Supervising Professor

Signature:

Associate Professor Dr. Abdul Rani bin Othman
Dean,
Faculty of Electronics and Computer Engineering

DEDICATION

Those who are raising hands

in the depths of the night asking Allah to blessing me.

ABSTRACT

During last two decades various speech coding algorithms have been developed. The range of toll speech frequency is from 300 Hz- 3400 Hz. Generally, human speech signal could be classified as non-stationary signal because of its fluctuation randomly over the time axis. One important assumption made to make the analysis of such signal even easier by assuming the speech signal is quasi-stationary over short range (frame). The frames of speech signal can be classified further into Voiced or Unvoiced, where the voiced part is quasi-stationary while the unvoiced part as an AWGN.

The quality of the synthesized signal is degraded significantly due to the excitation of voiced part not equally spaced within the frame and the excitation of the unvoiced part is not exact AWGN. This assumption produced a non-natural speech signal but with high intelligible level. One more reason is that the frame could have voiced plus unvoiced parts within the same frame, and by classifying this frame as voiced or unvoiced due to rigid decision would drop the level of quality significantly.

Speech compression commonly referred to as speech coding, where the amount of redundancies is reduced, and represent the speech signal by set of parameters in order to have very low bit rates. One of these speech coding algorithms is linear predictive coding (LPC-10).

This thesis implements LPC-10 analysis and synthesis using Matlab and C coding. LPC-10 have been compared with some other speech compression algorithms like pulse code modulation (PCM), differential pulse code modulation (DPCM), and code excited linear prediction coding (CELP), in term of segmental signal to quantization noise ratio SEG-SQNR and mean squared error MSE using Matlab simulation. The focus on LPC-10 was implemented on the DSP board TMS320C6713 to test the LPC-10 algorithm in real-time. Real-time implementation on TMS320C6713 DSP board required to convert the Matlab script into C code on the DSP Board. Upon successfully completion, comparison of the results using TMS320C6713 DSP against the simulated results using Matlab in both graphical and tabular forms were made.

ABSTRAK

Semasa dekad terakhir pelbagai algoritma pengekodan pertuturan telah dibangunkan. Pelbagai kekerapan pertuturan tol adalah dari 300 Hz - 3400 Hz. Secara amnya, isyarat pertuturan manusia boleh diklasifikasikan sebagai isyarat bukan pegun kerana turun naik secara rawak sepanjang paksi masa. Satu andaian penting dibuat untuk membuat analisis isyarat itu lebih mudah dengan menganggap isyarat pertuturan adalah seakan-akan bergerak dalam julat pendek (frame). Rangka isyarat ucapan boleh diklasifikasikan lebih jauh ke dalam Suara atau yg tak disuarakan, di mana sebahagian bersuara seakan-akan pegun manakala bahagian yg tak disuarakan sebagai AWGN.

Kualiti isyarat yang disintesis dihina dengan ketara disebabkan pengujaan bahagian bersuara yang tidak sama rata dijarakkan dalam jangka dan pengujaan bahagian yang tak disuarakan tidak tepat AWGN. Andaian ini menghasilkan isyarat bukan semula jadi ucapan tetapi dengan tahap difahami tinggi. Salah satu sebab lagi ialah bahawa bingkai dapat telah menyuarakan ditambah dengan yg tak disuarakan bahagian dalam bingkai yang sama, dan dengan mengklasifikasikan rangka ini sebagaimana yang disuarakan atau yg tak disuarakan kerana keputusan tegar akan turun tahap kualiti ketara.

Mampatan ucapan biasanya dirujuk sebagai pengekodan pertuturan, di mana jumlah berlebihan dikurangkan, dan mewakili isyarat pertuturan oleh satu set parameter untuk mempunyai kadar bit yang sangat rendah.

Salah satu-ucapan algoritma pengkodan pengkodan ramalan lurus (LPC-10). Tesis ini melaksanakan LPC-10 analisis dan sintesis menggunakan Matlab dan C pengkodan. LPC-10 telah ada berbanding dengan beberapa algoritma pemampatan lain ucapan seperti kod pemodulatan denyut (PCM), modulasi kod denyut pembezaan (DPCM), dan pengkodan ramalan kod teruja linear (CELP), dalam jangka isyarat segmen hingar pengkuantuman SEG-SQNR, dan min MSE ralat kuasa dua menggunakan simulasi Matlab. Tumpuan kepada LPC-10 telah dilaksanakan di papan DSP TMS320C6713 ke menguji algoritma LPC-10 ini dalam masa nyata. Real-time pelaksanaan pada TMS320C6713 DSP lembaga yang diperlukan untuk menukar skrip Matlab kepada kod C di Papan DSP. Apabila selesai dengan jayanya, perbandingan keputusan menggunakan TMS320C6713 DSP terhadap keputusan simulasi yang menggunakan Matlab dalam bentuk grafik dan jadual telah dibuat.

ACKNOWLEDGEMENT

I would like to thank my thesis supervisor Prof. Dr. Muhammad Ghazie bin Ismail, for his continuous advises, guidance, and support. The outcome of this thesis would have been impossible without the open discussion that we had during the courses of this thesis. Many thanks to digital signal processing laboratory staff, and their support is highly considered.

Last but not least, many thanks to my family members; in particular, Hosam and Husham, and my good friends living in various parts of the world for their continuous support in particular, Salem and Muntadhar.

LIST OF TABLES

Table 2.1: ITU-T Narrowband Speech Coding Standards	23
Table 2.2: ETSI Speech Coding Standards For GSM Mobile Communication	23
Table 2.3: TIA Speech Coding Standards For North America	24
Table 2.4: DoD Speech Coding Standards	24
Table 2.5: Speech Coders Classification According To Bit-Rate	28
Table 3.1: Implementation Requirement of A-Law And μ -Law	42
Table 3.2: Average Results of The Selected Speech Coders In Term of SEG-SQNR And MSE	58
Table 3.3: Statistical Review of The Selected Speech Coders Each In Term of SEG-SQNR And MSE For 25 Frames	63
Table 3.4: Speech Coders Classification According To Bit-Rate, Quality, And Compression Rate	69
Table 4.1: Processing Cycles of Triple Buffering System	84
Table 4.2: Statistical Data In Term of SEG-SQNR And MSE	91

LIST OF FIGURES

Figure 2.1: Anatomy of Human Vocal System	6
Figure 2.2: Time Waveform of The Sentence “The Wife Helped Her Husband” By Male Speaker	8
Figure 2.3: V/UV Parts Within Speech Signal	9
Figure 2.4: (Left) Voiced Frame With Corresponding Periodgram on The (Right)	9
Figure 2.5: (Left) Unvoiced Frame With Corresponding Periodgram on The (Right)	10
Figure 2.6: Periodic Similarity Plot Superimposing Speech Signal	11
Figure 2.7: (Left) Error Signal Obtained From Speech Waveform, (Right) Peakiness Measure of Error Signal	12
Figure 2.8: (Left): Sample Speech Voice Has Both Voiced/ Unvoiced Parts, (Right) Zero-Crossing Response of Speech Under Test	13
Figure 2.9: Spectrum Tilt Superimposing Speech Signal	14
Figure 2.10: Pre-Emphasis Spectrum Ratio Superimposing The Speech Signal	15
Figure 2.11: Low-Band To Full-Band Energy Ratio Superimposing The Speech Signal	16
Figure 2.12: AMDF And Speech Signal: Minima of AMDF Point Out The Pitch Values Are Indicated By The Circles	18
Figure 2.13: Normalized Autocorrelation and Speech Signal: Maxima of The Normalized Autocorrelation Corresponding To Pitch Value Are Indicated By Circles	19

Figure 2.14: (A) Speech Signal Of $T_0 = 34$ -Sample ($F_s = 8$ KHz), (B) Magnitude Spectrum, (C) Zero-Crossing Spectrum, And (D) Spectral Autocorrelation	21
Figure 2.15: (Top) Speech Signal Male Voice (Bottom) Corresponding Periodgram “Time-Frequency”	22
Figure 2.16: MOS Score Vs. Bit Rate	25
Figure 2.17: Speech Coders According To Coding Techniques	26
Figure 2.18: Evaluation of Speech Coders According To Coding Techniques	28
Figure 2.19: LP Synthesis Model	29
Figure 3.1: Transfer Characteristic $Q(X)$ of Uniform Quantizer	34
Figure 3.2: Quantization Error $E(X)$ of Uniform Quantizer	34
Figure 3.3: (A) 1Hz Sine Wave Signal (B) Quantized Signal At 3-Bits (C) Error Signal	37
Figure 3.4: (A) 1 Hz Sine Wave Signal (B) Quantized Signal At 8-Bits (C) Error Signal	38
Figure 3.5: (Top) Normalized Sine Wave Signal (Middle) Compressed Signal Using μ -Law Compressor (Bottom) Characteristic Curve of μ -Law Compressor	40
Figure 3.6: (Top) Normalized Sine Wave Signal (Middle) Compressed Signal Using A-Law Compressor (Bottom) Characteristic Curve of A-Law Compressor	41
Figure 3.7: Comparison of A-Law And μ -Law Characteristic	42
Figure 3.8: Differential Pulse Code Modulation (A) Transmitter (B) Receiver	43
Figure 3.9: (Top Downwards) The Original Signal, Error Signal, Quantized Error Signal, Reconstructed Signal, And Lastly The Reconstructed Filtered Signal	45
Figure 3.10: (Top) Original Sine Wave Signal With Corresponding Binary Code (Bottom) Error Signal With Corresponding Binary Code	46
Figure 3.11: CELP Encoder Analysis-By-Synthesis Loop	47
Figure 3.12: (Left) Formant Synthesis Filter (Right) Associated Formant Analysis Filter	48

Figure 3.13: (Top) Male Speech Signal (Middle-Left) Frequency Envelop Of Voiced Frame With Frequency Response of Synthesis Filter (Middle-Right) Frequency Envelop of Unvoiced Frame (Bottom Left To Right) Is The Voiced And Unvoiced Frames Respectively	49
Figure 3.14: Magnitude Response of Pre-Emphasis Filter	50
Figure 3.15: (Top) Speech Signal Record Male Voice (Bottom) Corresponding Spectrogram	51
Figure 3.16: (Left) Selected Voiced Frame 240 Samples (Right) Corresponding Periodgram	52
Figure 3.17: Analysis-Synthesis Filters Response of Voiced Frame	52
Figure 3.18: Synthesis Filter Response Matching The Formants of Voiced Frame	53
Figure 3.19: (Top) Pitch Periodicity of Voiced Frame (Bottom) Corresponding Pitch Lag Using Autocorrelation Method	54
Figure 3.20: (Left) Selected Unvoiced Frame 240 Samples (Right) Corresponding Periodgram	54
Figure 3.21: (Top Downwards) Original Signal, Quantized Signal, And Error Signal of PCM Codec	56
Figure 3.22: (Top Downwards) Original Signal, Quantized Signal, And Error Signal of DPCM Codec	57
Figure 3.23: (Top Downwards) Original Signal, Quantized Signal, And Error Signal of CELP Codec	57
Figure 3.24: (Top Downwards) Original Signal, Quantized Signal, And Error Signal of LPC-10 Codec	58
Figure 3.25: Average SEG-SQNR Comparison of The Selected Speech Codecs	59
Figure 3.26: Average MSE Comparison of The Selected Speech Codecs	60

Figure 3.27: MSE Against SEG-SQNR of PCM Codec In Semiology Graph	61
Figure 3.28: MSE Against SEG-SQNR of DPCM Codec In Semiology Graph	61
Figure 3.29: MSE Against SEG-SQNR of CELP Codec In Semiology Graph	62
Figure 3.30: MSE Against SEG-SQNR of LPC-10 Codec In Semiology Graph	62
Figure 3.31: The Response of SEG-SQNR of PCM Codec With Its Average Values	64
Figure 3.32: The Response Of MSE of PCM Codec With Its Average Values	65
Figure 3.33: The Response of SEG-SQNR of DPCM Codec With Its Average Values	65
Figure 3.34: The Response of MSE Of DPCM Codec With Its Average Values	66
Figure 3.35: The Response of SEG-SQNR of CELP Codec With Its Average Values	66
Figure 3.36: The Response Of MSE Of CELP Codec With Its Average Values	67
Figure 3.37: The Response Of SEG-SQNR Of LPC-10 Codec With Its Average Values	67
Figure 3.38: The Response of MSE of LPC-10 Codec With Its Average Values	68
Figure 3.39: Evaluation of The Selected Speech Codecs According To The Subjective Test	70
Figure 4.1: TI DSP Processors Family	71
Figure 4.2: TMS320C6713 DSP Board Layout	72
Figure 4.3: TMS320C6713 DSP Functional Blocks Diagram	73
Figure 4.4: Data Path In C6713 DSP	75
Figure 4.5: Sample-By-Sample Real-Time System	76
Figure 4.6: Single Buffer Timing System	77
Figure 4.7: Broken System Scenario	78
Figure 4.8: Ping-Pong Timing System	79
Figure 4.9: Receive And Transmit Buffers	80
Figure 4.10: Ping-Pong Toggling Between 0 And 1	81
Figure 4.11: Ping-Pong Process With Related Hardware	81

Figure 4.12: Flow Chart Of Triple Buffering Technique	82
Figure 4.13: The Three Buffers With Its Pointers In Initial Condition	83
Figure 4.14: Host-Target Jtag Connection	87
Figure 4.15: Under Test Speech Signal Of Male Voice	88
Figure 4.16: Segmental-Sqnr Versus Mean Squared Error For Matlab And C6713 Dsp	89
Figure 4.17: Curve Fitting Of SEG-SQNR Vesus MSE; Left: C6713 DSP Right: Matlab	90
Figure 4.18: Single Side Fft Plot; Left: TMS320C6713 DSP Implementation Right: Matlab Simulation	90
Figure 4.19: Matlab Simulation And Dsp Implementation Comparison In Term Of SEG- SQNR	92
Figure 4.20: Matlab Simulation And Dsp Implementation Comparison In Term Of MSE	92

LIST OF ABBREVIATIONS

Acronym	Definition
AbS	Analysis by Synthesis
ACELP	Algebraic code-excited linear prediction
ACF	autocorrelation function
ADC	Analog-to-digital converter,
ADPCM	Adaptive differential pulse-code modulation
A-Law	companding algorithm
AMDF	Average Magnitude Difference Function
AR	Autoregressive
ARMA	autoregressive moving average
AWGN	Additive white Gaussian noise
CCS	code composer studio
CELP	code excited linear prediction
CSR	codec sample rate
DAC	Digital-to-analog converter
DMA	Direct Memory Access
DPCM	differential pulse code modulation
DRT	Diagnostic Rhyme Test
DSP	digital signal processing digital signal processor
EDMA	Enhanced Direct Memory Access
FIR	Finite impulse response

Acronym	Definition
FR	Full Rate
I/O	Input/Output
IEEE	Institute Of Electrical And Electronics Engineers
ITU	International Telecommunications Union
KHz	kilo hertz
LD-CELP	low-delay code excited linear prediction
LP	linear predictive
LPC	linear predictive coding
LTP	Long Term Predictor
MELP	Mixed-excitation linear prediction
MIPS	Million instructions per second
MOS	Mean Opinion Score
MSE	Mean Squared Error
PCM	Pulse Code Modulation
QCELP	Qualcomm code-excited linear prediction
RTDX	real time data exchange
RTS	Real-Time System
SDRAM	Synchronous dynamic random access memory
SNR	signal to noise ratio
STFT	short-time Fourier transform
TI	Texas Instruments
TIA	Telecommunication Industry Association
UQ	Uniform Quantization
V/UV	Voiced/ Unvoiced
VOCODER	Voice Encoder
VSELP	Vector sum excited linear prediction

TABLE OF CONTENT

ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENT	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS	xii
CHAPTER 1	1
INTRODUCTION TO WIRELESS COMMUNICATIONS AND SPEECH CODERS.....	1
1.0 Introduction	1
1.1. Low Bit Rate Communication.....	2
1.2. Organization of the thesis	3
CHAPTER 2	5
OVERVIEW OF SPEECH CODING	5
2.0 Phonetics and Theory of Speech Production.....	5
2.1. Characteristics of Speech Signal	7
2.1.1. Voiced/ Unvoiced Sounds.....	7
2.1.2. Voiced/ Unvoiced Decision Algorithms	10
2.1.2.1. Periodic Similarity	10
2.1.2.2. Peakiness of Speech.....	11
2.1.2.3. Zero-Crossing	12
2.1.2.4. Spectrum Tilt	13
2.1.2.5. Pre-emphasized Energy Ratio.....	14
2.1.2.6. Low-band to Full-band Energy Ratio	15
2.2. Pitch Period	16
2.3. Time Domain Pitch Estimation Approaches	17
2.3.1. Average Magnitude Difference Function (AMDF) Method.....	17

2.3.2.	Autocorrelation Function (ACF) Method	18
2.4.	Frequency Domain Pitch Estimation Approaches	19
2.5.	Time-Frequency Pitch Estimation	19
2.6.	Formants	21
2.7.	Major Speech Coder Standards	22
2.8.	Classification of Speech Coders According to Coding Techniques.....	25
2.8.1.	Waveform Coding	26
2.8.2.	Parametric (Vocoders) Coding.....	26
2.8.3.	Hybrid Coding.....	26
2.9.	Classification of Speech Coders According to Bit Rate.....	28
2.10.	Linear Prediction Coding	28
2.10.1.	Fundamentals of Linear Prediction	28
2.10.2.	Levinson–Durbin Algorithm.....	30
2.11.	Speech Quality Assessment	31
2.11.1.	Objective Measurement	31
2.11.2.	Subjective Measurement	32
CHAPTER 3		32
VOICE CODING TECHNIQUES		32
3.0	Introduction	32
3.1.	Pulse Code Modulation (PCM)	33
3.2.	Uniform Quantization (UQ)	33
3.3.	Scalar Quantization Overview.....	35
3.3.1.	Partitions	36
3.3.2.	Codebooks.....	36
3.4.	Non-Uniform Quantization	39
3.4.1.	μ -Law Quantization	39
3.4.2.	A-Law Quantization.....	40
3.5.	Comparison of A-Law And μ -Law Quantizers	41
3.6.	Differential Pulse Code Modulation (DPCM).....	42
3.7.	Comparison of PCM and DPCM.....	45
3.8.	Code Excited Linear Predictive Coding (CELP).....	46
3.8.1.	Analysis By Synthesis (AbS)	47
3.8.2.	Perceptual Weighting	47

3.8.3.	Pitch Estimation	48
3.9.	Linear Prediction Coding (LPC-10)	50
3.9.1.	Pre-Emphasis Filter	50
3.9.2.	Speech Analysis	50
3.10.	Analysis of the Results for the Comparison of Speech Codecs PCM, DPCM, CELP, and- LPC-10.....	55
3.11.	Summary	69
CHAPTER 4.....		71
IMPLEMENTATION OF LPC-10 USING TMS320C6713 DSP		71
4.0	TMS320 FAMILY.....	71
4.1.	TMS320C6713 DSP ARCHITECTURE	72
4.2.	Fixed-Point Versus Floating-Point Solutions	73
4.3.	DMA and CPU Data Access	74
4.4.	Real-Time System (RTS) Definition.....	75
4.5.	Buffering Technique.....	76
4.5.1.	Single Buffer Timing System.....	76
4.5.2.	Broken system.....	78
4.5.3.	Double Buffer Timing System (Ping-Pong Buffering System)	79
4.5.3.1.	Implementing a Double Buffer System	80
4.5.4.	Triple Buffering Technique.....	82
4.6.	Implementation of LPC-10 Algorithm	84
4.6.1.	Segmentation.....	84
4.6.2.	Windowing	85
4.6.3.	Levinson–Durbin algorithm	85
4.6.4.	Residual signal	86
4.6.5.	Speech synthesis.....	86
4.6.6.	Accumulation and buffering	86
4.7.	Real-Time Data Exchange (RTDX)	86
4.7.1.	How does RTDX work.....	87
4.7.2.	Target to Host Transfer	87
4.8.	Result	88
CHAPTER 5		93
CONCLUSION AND FUTURE WORK		93
BIBLIOGRAPHY		95

APPENDIX A	97
APPENDIX B.....	103

CHAPTER 1

INTRODUCTION TO WIRELESS COMMUNICATIONS AND SPEECH CODERS

1.0 Introduction

Even though humans are well equipped for analog communication, analog transmission remains practically inefficient due to the complexity of separating the real voice signals from noise. Amplifying the analog signal will amplify the noise too, and eventually it would be very difficult to restore the original analog signal from the noisy environment. Digital signals has only two states “One-bit” and “Zero-bit”, and it could be separated from noise easily, and they can be amplified without noise corruption. However, the fast growth of wireless digital communication systems in the last two decades, the need for low bitrate speech coders have increased rapidly for both civilian and military applications. With the rise of optical fiber, bandwidth in wire-based communication has become low-cost. The growing need for bandwidth saving and improving privacy in wireless cellular and satellite communications has also increased (Andreas 1994). However, the field of speech processing is becoming more important, and could be subdivided into speech analysis, speech coding and speech synthesis. Speech analysis is defined in which parameters that define the speech signal are extracted. While speech coding could be defined in which the information bearing the speech signals is coded to remove redundancy. This helps decrease the transmission bandwidth requirements,

improve storage efficiency while still maintaining acceptable speech quality. Finally speech synthesis could be defined in which speech parameters (extracted from speech analysis) are used to re-generate the original speech signal (Ogunfunmi and Narasimha 2010).

Speech coders are typically characterized in three groups: waveform, Hybrid and Vocoders. Waveform coders quantize the original speech signal directly operating at high bit rates of between 16-64 kbps. Pulse Code Modulation (PCM) and Differential PCM are good example of waveform coders. Hybrid coders can be categorized into frequency domain and time domain, the former divides the speech spectrum frequency into bands (Sub-Band Coding, Adaptive Transform Coding, and Multi-Band Excitation). While latter employ the source filter usually with linear prediction as examples are CELP, MELP, and MP-LPC, Hybrid coders are mixed from waveform-based and model-based, and operates at 2.4 to 16 kbps. Finally Vocoders are model-based and operates at 1.2 to 4.8 kbps and LPC-10 is an example (Quatieri 2002).

1.1. Low Bit Rate Communication

Nyquist criterion determines the minimum sampling frequency for speech signals. Therefore, the quality of the signal required at the receiver is determined by the number of quantization levels. These two conditions determine the bit rate of PCM system at 8 bit per sample thus making the bit rate at 64kbits per second. Due to fast emergence of fixed and mobile telecommunication in the last two decades, the need to reduce PCM bit rate in speech compression and the digital coding of speech signals becomes an important field of research.