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 lelebihan dikurangkan, dan mewakili isyarat pertuturan oleh satu set parameter untuk mempunyai kadar bit yang sangat rendah.

Salah satu-ucapan algoritma pengekodan pengekodan ramalan lelurus (LPC-10). Tesis ini melaksanakan LPC-10 analisis dan sintesis menggunakan Matlab dan C pengekodan. LPC-10 telah ada berbanding dengan beberapa algoritma pemampatan lain ucapan seperti kod pemodulatan denyut (PCM), modulasi kod denyut pembezaan (DPCM), dan pengekodan 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 stuff, 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

LIST OF FIGURES

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

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.