

Ain Shams University
Faculty of Engineering
Electronics and Communication Eng. Dept.

A Method For Nonuniform Sampling of Signals with Application to Speech Encoding

By

Eng. Shawky M. Galal Foda

A Thesis

Submitted to the Faculty of Engineering in partial fulfillment of the
requirements of Master Degree in Electrical Engineering
[Electronics & Communications Engineering, Dept.]

Supervised By

Prof. Dr. Salwa El-Ramly
Electronics and Communication Eng. Dept.
Faculty of Engineering
Ain Shams University

Prof. Dr. Hadia El-Henawy
Electronics and Communication Eng. Dept.
Faculty of Engineering
Ain Shams University

Dr. Mohammed H. El-Shafie
Computer & Systems Eng. Dept.
Faculty of Engineering
Ain Shams University

Cairo 1997



Examiners Committee

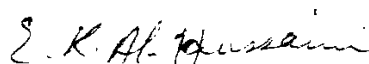
Thesis Title : A Method For Nonuniform Sampling of Signals
with Application to Speech Encoding
Presented by : Shawky M. Galal Foda

Approved By

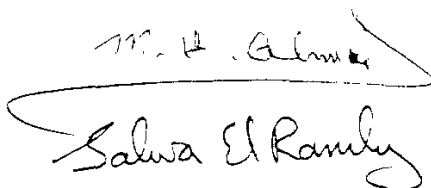
Name and Title

Signature

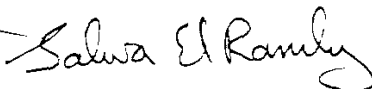
1- Prof. Dr. Emad El-deen El Hosseni



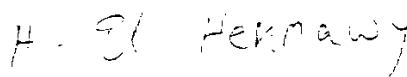
2- Prof. Dr. Mahmoud Hanfi Ahmed



3- Prof. Dr. Salwa H. El- Ramly



4- Prof. Dr. Hadia S. El-Henawy





Statement

This thesis is submitted to Ain Shams University in Partial fulfillment of the degree of M.Sc. in Electrical Engineering.

The Work included in this thesis was carried out by the author in the Department of Electronics and Communication Engineering, Faculty of Engineering, Ain Shams University.

No part of this thesis has been submitted for a degree or qualification at any other university or institute.

Name : Shawky M. Galal Foda
Signature :
Date :

Curriculum Vita .

Name of Researcher : Shawky Mohammed Galal Foda.
Date of Birth : 26/8.1968.
Place of Birth : Cairo - Egypt.
First University Degree : B.Sc. in Electrical Engineering
University : Ain Shams University.
Date of Degree : 1990

Name of Researcher : Shawky Mohammed Galal Foda.

Signature :

Date :

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Acknowledgment

I would like to express my deep feelings and appreciation to all the people who contributed to the fulfillment of this work. No words can express my sincere gratitude to all of them.

I would like to thank my supervisor, Prof. Dr. Salwa El-Ramly Prof. Of Communications, Faculty of Engineering, Ain Shams University, for her supervision, her helpful comments, her encouragement, and her unlimited assistance and patience.

Also I would like to thank Prof. Dr. Hadia El-Henawy Prof. Of Electronics and communications, Faculty of Engineering, Ain Shams University, for her supervision and her encouragement.

I wish to express my deep gratitude and respect to Dr. Mohammed H. El-Shafie, System engineering and computer Dept. Faculty of Engineering, Ain Shams university.

I would like to thank all the members of the Electronics and Communication Engineering Dept., Ain Shams University, for providing me with the laboratory equipment and academic experiences and also the unlimited assistance during this work.

Finally, let me thank my family, my parents and let me thank my friends, who helped me along the way.

Shawky Galal Foda

Abstract

Shawky M. Galal Foda. "A Method For Nonuniform Sampling of Signals with Application to Speech Encoding.", Master Degree. Ain Shams University. Faculty of Engineering. Electronics and communications Eng. Dept. 1997.

In digital communication a uniform sampling of signals is usually used where the inter-samples intervals are constant and equal to the inverse of the Nyquist sampling frequency which is calculated to be twice the maximum frequency component in the signal. Some signals have the nature of being nonstationary, where its statistical characteristics is varying with time, and one of these characteristics is the maximum frequency component in the signal.

Speech signal is a nonstationary signal and its maximum frequency component is varying with time. The conventional uniform sampling rate of speech signal is equal to 8 kHz where the maximum frequency component in the speech signal is considered to be 4 kHz. This maximum frequency component, in fact, occurs in short intervals of time and during most of the speech signal interval the maximum frequency component is less than this value. So during these intervals of low maximum frequency component (low bandwidth) the 8 kHz sampling frequency is considered to be over-sampling frequency and results in a high amount of redundant information which degrades the efficiency of the speech coder. many efforts are done to overcome this problem by using a variable or nonuniform sampling rate.

In this thesis a survey of nonuniform sampling rate methods is introduced and a simple technique for continuous variable sampling rate is proposed. The sampling rate of the proposed method is continuously varied according to the amount of information (not only the rate of change) in the speech signal. The proposed method is studied using a sinusoidal test signal and a real speech signal. The quality of the decoded speech signal is compared with that of the uniform sampling method using a sampling frequency equal to the resulting average nonuniform sampling frequency in order to maintain the same average bit rate.

Contents

<i>Chapter (1): Nature Of Speech Signal</i>	1
1-1 Introduction.....	1
1-2 Application.....	1
1-3 Compression of Speech signal.....	2
1-4 Waveform coders versus vocoders.....	3
1-5 Characteristics of speech waveform.....	4
1-5-1 Amplitude versus time characteristics.....	4
1-5-2 Probability Density function.....	6
1-5-3 Probability density of speech signal.....	6
1-5-4 Advanced time frequency representation for speech.....	7
1-6 Statistical characteristics of speech.....	9
1-6-1 Non-stationarity of speech signals.....	9
1-6-2 Autocorrelation Function.....	9
1-6-3 Power Spectral Density Function.....	12
1-7 Sampling of Signals.....	12
1-8 Pre-sampling filter.....	15
1-9 Quantization.....	16
1-9-1 Uniform quantization.....	16
1-9-2 Logarithmic quantization.....	18
1-10 Conclusion.....	20
<i>Chapter (2): Waveform Speech coding</i>	21
2-1 Introduction.....	21
2-2 Analog Waveform Coding.....	21
2-3 Digital waveform coding.....	22
2-3-1 Pulse Code Modulation (PCM).....	22
2-3-1-1 Quantization error in PCM.....	23
2-3-1-2 Dynamic range of PCM.....	24
2-3-2 Delta Modulation.....	24
2-3-3 Differential PCM.....	26
2-3-3-1 DPCM versus PCM and prediction error.....	26
2-3-3-2 Closed loop DPCM with feedback around quantizer.....	27
2-3-3-3 Reconstruction error in DPCM.....	27
2-3-4 Linear Prediction of Order $N=1, 2$	29
2-3-4-1 First order predictor.....	29
2-3-4-2 Second order predictor.....	30
2-3-4-3 Linear predictors of order N	31
2-3-4-4 General Expressions for Optimum Predictor.....	31
2-4 Adaptive DPCM (ADPCM).....	33
2-4-1 Adaptive Quantization.....	33
2-4-2 Adaptive prediction.....	35
2-4-2-1 Forward Adaptation.....	35
2-5 Sub-band Coding.....	37

2-6 Comparison between waveform coders	39
2-7 Conclusion	40
<i>Chapter (3) Analysis By Synthesis Speech Coding</i>	41
3-1 Introduction	41
3-2 Analysis by synthesis predictive coding	41
3-3 General model for analysis by synthesis speech coding	42
3-3-1 Short term predictor	43
3-3-2 Long term predictor	45
3-4 The excitation signal	48
3-4-1 Multi-pulse and regular-pulse excitation	50
3-4-2 Code excited linear prediction (CELP)	51
3-5 Conclusion	52
<i>Chapter (4) Quality Evaluation of Speech Coding</i>	53
4-1 Introduction	53
4-2 Objective measurement of coder performance	53
4-2-1 Signal-to-noise ratio, SNR	54
4-2-2 Segmental Signal to Noise Ratio, SEGSR	54
4-2-3 Frequency weighted SNR and reconstruction error	55
4-2-4 Maximum signal to noise ratio SNRMAX	57
4-2-5 Signal-to- Uncorrelated Noise Ratio SNRu	57
4-3 Subjective measurement of coder performance	58
4-3-1 Quality Measurements	58
4-3-1-1 Binary Decision Tasks	58
4-3-1-2 Subjective SNR	58
4-3-1-3 Mean Opinion Score (MOS)	59
4-3-1-4 Speech Reception Threshold(SRT)	60
4-3-2 Intelligibility tests	62
4-4 Conclusions	63
<i>Chapter (5) Variable Rate and Nonuniform Sampling Methods</i>	64
5-1 Introduction	64
5-2 Nonuniform Sampling Delta Modulation NS-DM	64
5-2-1 Nonuniform sampling delta Modulation Algorithm	65
5-2-2 Signal Reconstruction - Demodulation	67
5-3 Sampling at the Maxima and Minima and its Application on Speech Coding	67
5-3-1 Implementation of sampling at maxima and minima	68
5-3-2 Extreme points detection algorithm	69
5-3-3 Reconstruction	69
5-3-4 Amplitude and interval coding	71
5-4 Adaptive control of sampling rate	71
5-4-1 Time Domain Sampling Theorem	72
5-4-2 Theorem	72
5-4-3 Design of adaptive sampling rate system	73

5-5 Variable bit rate ADPCM on ATM networks.....	74
5-5-1 Quality of Conventional Fixed Rate Coders.....	74
5-6 Variable sampling rate PCM (VSRPCM).....	76
5-7 Conclusion.....	79
 <i>Chapter(6) :The Proposed Nonuniform Sampling Method.....</i>	<i>80</i>
6-1 Introduction.....	80
6-2 Justification of the variable sampling rate technique.....	90
6-3 The proposed triangle method for nonuniform sampling.....	83
6-4 Application on a sinusoidal signal.....	85
6-5 Application on speech signal.....	88
6-6 Segmental signal to noise ratio.....	91
6-7 Estimating the compression ratio C.....	95
6-8 Time domain analysis of the proposed nonuniform sampling method.....	99
6-9 Segmental signal to noise ratio versus compression ratio.....	101
6-10 Subjective quality measuerments.....	103
6-11 Comparision between the proposed method and other variable sampling rate method.....	103
6-11-1 Comparison with sampling at maxima and minima method.....	103
6-11-2 Comparison with the variable sampling rate VSRPCM.....	106
6-12 Optimum values of A, N.....	110
6-13 Implementation of the proposed nonuniform sampling method.....	116
6-13-1 Analog circuit.....	116
6-13-2 Digital circuit.....	117
6-14 Conclusion.....	117
 <i>Chapter (7) Conclusion and Future work.....</i>	<i>118</i>
<i>References.....</i>	<i>117</i>
 Appendices.....	
Appendix A: Matlab Source Code for Ovr2non.m.....	A-1
Appendix B: Matlab Source Code for Cfbuan.m.....	B-1
Appendix C: Matlab Source Code for Bestan.m.....	C-1
Appendix D: Matlab Source Code for Ctubuan.m.....	D-1
Appendix E: Matlab Source Code for Nonstat.m.....	E-1

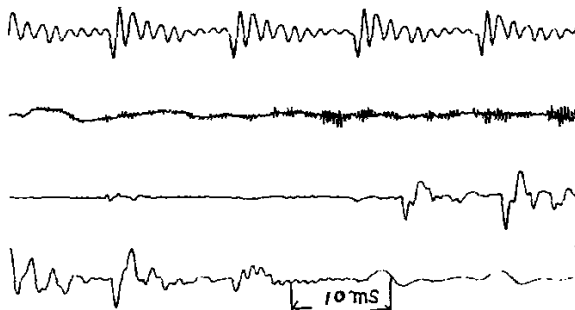
List of tables

Table(1-1) Comparison between waveform and parametric coders.....	4
Table(1-2) Four model pdf's with a mean value of zero	7
Table(1-3) Values of acf for 8 kHz-samples speech.....	11
Table(2-1) Objective SNR measurements for some of speech waveform coders.....	39
Table(4-1) SEGSNR and Conventional SNR	55
Table(4-2) Frequency bands of equal contribution to the Articulation Index (all entries are in Hz).....	57
Table(4-3) Five-point adjectival scales for quality and impairment, and associated number scores	60
Table(5-1) The Logic of MIF	66
Table(5-2) Function of F1,F2,F3	66
Table(5-3) Coding rate control criteria	75
Table(6-1) Average improvement factor η_{avr} for different compression ratios C and different messages.....	92
Table(6-2) A comparison between the actual and estimated C	96
Table(6-3) The values of A,N to get $\eta_{avr}=20$ and the corresponding improvement	97
Table(6-4) The values of A,N to get $\eta_{avr}=30$ and the corresponding improvement	97
Table(6-5) The values of A,N to get $\eta_{avr}=50$ and the corresponding improvement in segmental signal to noise ratio	98
Table(6-6) Calculated improvement ratio in time domain versus compression ration C for different messages.....	99
Table(6-7) Measured MOS values for different speech signals with the proposed nonuniform triangle method and the uniform with the same average sampling frequency.....	101
Table(6-8) A comparison between the segmental signal to noise ratio of the proposed nonuniform triangle method and sampling at maxima and minima method.....	106
Table (6-9) : Segmental signal to noise ratio for different bit rates and different coding system	109
Table(6-11): The optimum values of A and N and the corresponding improvement in the SNR for the English utterance " add the sum to the product of these three".....	110
Table(6-12): The optimum values of A and N and the corresponding improvement in the SNR (for the English sentence " Paper jam".....	111
Table(6-13): The optimum values of A and N and the corresponding improvement in the SNR for a sound of peep.....	112
Table(6-14): The optimum values of A and N and the corresponding improvement in the SNR (for the English utterance " add paper to printer".....	113
Table(6-15): The average optimum values of A,N and the improvement in SNR from tables (6-11), (6-12), (6-13), (6-14).....	116

List of Abbreviation

A-b-S	Analysis-by-Synthesis
acf	Auto-correlation function
ADM	Adaptive Delta Modulation
ADPCM	Adaptive Delta Pulse Code Modulation
APB	Adaptive Prediction Backward
APF	Adaptive Prediction Forward
AWGN	Additive Gaussian Noise
BER	Bit Error Rate
CCITT	International Telegraph and Telephone Consultative Committee
CELP	Code Excited Linear Prediction
CVSD	Continuously Variable Slope Delta Modulation
DAC	Digital to Analog Converter
DFT	Discrete Fourier Transform
DPCM	Differential Pulse Code Modulation
DPCM-AQB	Differential Pulse Code Modulation-Adaptive Quantizer Backward
DPCM-AQF	Differential Pulse Code Modulation-Adaptive Quantizer Forward
DSI	Digital Signal Interpolation
DSP	Digital Signal Processing
FFT	Fast Fourier Transform
ISDN	International Service Digital Network
MP-LPC	Multi-Pulse Linear Predictive Coder
MSE	Mean square Error
PCM	Pulse Code Modulation
RAM	Random Access Memory
RFLP	Regular Excited Linear Prediction
RMS	Root Mean Square
ROM	Read Only Memory
SBC	Sub-Band Coders
SEGSNR	Segmental Signal to Noise Ratio
SNR	Signal to Noise Ratio
TDMA	Time Division Multiple Access
VBR	Variable Bit Rate
VQ	Vector Quantization

The signals shown in figure(1-1) illustrate the great variety of the character of speech wave. Sometimes periodic or quasi periodic, other times a mixture of periodic and random-like signals, and some times the waveform appears like random noise. A 10 ms time interval is shown in figure(1-1), a speech coder operating for example, at 4kb/s must be able to describe any such 10 ms segment (80 samples) using only 40 binary digits in such a way that the segment will be reproduced with an accuracy sufficient to ensure that it will sound very close to the original[20].

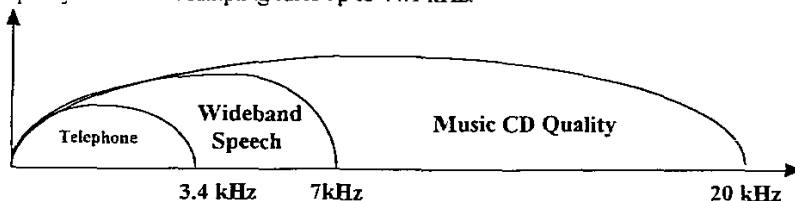


Figure(1-1): Example of speech waveform [20]

1.3-Compression of Speech signal

Compression of speech has been an ongoing area of research for several decades. In the last several years, there has been an interest and activity in this area with numerous applications in telecommunication and storage. High fidelity audio compression has also advanced rapidly in recent years, accelerated by the commercial success of consumer and professional digital audio products. The surprising growth of compression techniques is driven by the insatiable demand for voice communication, by the new generation of technology for cost effective implementation of digital signal processing algorithms, by the need to conserve band width in both wired and wireless telecommunication networks, and the need to conserve disk space in digital voice storage systems. As shown in figure(1-2) most of these efforts are focused on:

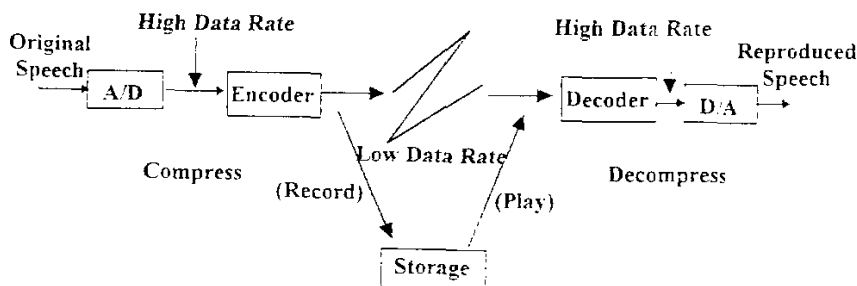
- 1-The usual telephone bandwidth of roughly 3.2 kHz (200 Hz to 3.4 kHz) with sampling rates of 8 kHz.
- 2- Wideband speech (7 kHz) for audio in video teleconferencing with sampling frequency of 16 kHz
- 3-Wideband audio signal (typically 20 kHz bandwidth) for transmission and storage of CD quality music with sampling rates up to 44.1 kHz.



Figure(1-2) Types of signals

Another new area of application is multimedia in personal computing where voice storage is becoming a standard feature.

Virtually all work in speech and audio compression involves *lossy compression* where the numerical representation of the signal samples is never recovered exactly after decoding (decompression). There is a wide range of tradeoffs between bit rate and recovered speech quality that are of particular interest in the coding of telephone speech, where users are accustomed to certain degrees of degradation. On the other hand, for wide band audio compression the quality is close to that of the compact disk (CD). Thus research in speech compression includes studies for distortion-rate tradeoffs motivated by different applications with different quality objectives.



Figure(1-3) Compression and Decompression of speech

Figure(1-3) is a schematic diagram illustrates the purpose of the encoding process which enables low data transmission rates and less storage capacity.

1.4-Waveform coders versus voice coders (vocoders)

Speech coding algorithm can be divided into two main categories

- 1- Waveform coders.
- 2- Vocoders (voice coders)

In *waveform coders*, the data transmitted from encoder to decoder specify a representation of the original speech as a waveform of amplitude versus time, so that the reproduced signal approximate the original waveform and, consequently, provides an approximate recreation of the original sound

Waveform coders (non-parametric) are used to compress telephone, speech and music bandwidth signal. It is designed to efficiently quantize any audio waveform by quantizing only new and unpredictable information. Such coders are robust for diverse signal sets, but don't accomplish high compression

In contrast, *Vocoders* do not reproduce an approximation to the original waveform, instead, parameters that characterize individual sound segments are specified and transmitted to the decoder, which then reconstructs a new and different waveform that will have a similar sound. *Vocoders* are sometimes called *parametric coders*, while *waveform coders* called *nonparametric coders*. Often these parameters characterize the short term spectrum of the sound. Alternatively the parameters specify a mathematical model of human speech production suited to a particular sound. In either case the parameters do not provide sufficient information to regenerate an