Design of Real Time Echo Detector Using Fuzzy Logic for Voice over Internet Protocol (VoIP)

Durojaiye, Oyetade Adebowale
M. ENG/SEET/2007/1624

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING,
SCHOOL OF ENGINEERING AND ENGINEERING TECHNOLOGY,
FEDERAL UNIVERSITY OF TECHNOLOGY MINNA,
NIGER STATE.

NOVEMBER, 2011

Design of Real Time Echo Detector Using Fuzzy Logic for Voice over Internet Protocol (VoIP)

Durojaiye, Oyetade Adebowale
M. ENG/SEET/2007/1624

A THESIS SUBMITTED TO THE POSTGRADUATE SCHOOL,
FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF
MASTER OF ENGINEERING (M. ENG.) DEGREE IN ELECTRICAL AND
COMPUTER ENGINEERING (COMMUNICATION OPTION).

NOVEMBER, 2011

DECLARATION

I, **DUROJAIYE, OYETADE ADEBOWALE** hereby declare that this work titled "Design of Real Time Echo Detector Using Fuzzy Logic for Voice over Internet Protocol (VoIP)" was done by me in partial fulfilment of the requirements for the award of M. Eng. in Communication Engineering in the Department of Electrical and Computer Engineering, Federal University of Technology, Minna, Niger State.

Durojaiye, Oyetade Adebowale
(Student)

Duto 05/03/2012
Signature and Date

CERTIFICATION

...ign of Real Time Echo Detector Using Fuzzy Logic for Voice Protocol (VoIP): Durojaiye, Oyetade Adebowale ...007/1624) meets regulations governing the award of the degree of ...ineering of the Federal University of Technology, Minna and is approved ...ibution to scientific knowledge and literary presentation.

07/03/12
Signature and Date

...E. N. ONWUKA
(Supervisor)

(March 7 20
Signature and Date

ENGR. A.G. RAJI
(Head of Department)

03/12
Signature an...

ENGR. PROF. M. S. ABOLARIN
(Dean, SEET)

/12
Si...

PROF. MRS. S. N. ZUBAIRU
(Dean, Postgraduate School)

Dedication

I dedicate this work to my parents.

ACKNOWLEDGEMENTS

All praise and thanks are due to God Almighty the most exalted for making this work a reality. In putting together this work some individuals have contributed tremendously and I am grateful to them for their assistance. In this regard, my mentor Dr. E. N. Onwuka deserves a high level of appreciation for finding time to read through this work thoroughly, and for her suggestions which have assisted in improving the quality of this project. I would also like to acknowledge the contribution and encouragement of my Head of Department; Engr. A. G. Raji, the Dean, School of Engineering and Engineering Technology; Prof. M.S Abolarin, the Deputy Dean, School of Engineering and Engineering Technology; Engr. Dr. O. Chukwu, the Dean, Postgraduate School; Prof. Mrs. S. N. Zubairu, the Secretary, School of Engineering and Engineering Technology and my colleagues. May God bless you all.

I cannot express my gratitude in words to my parents for their concern and constant prayers, my beloved wife Gladys for all her love, support and encouragement.

Special thanks go to Engr. Mrs. C. Alenoghena, Engr. Dr. Y. Adediran , Engr. Dr. M. N. Nwohu whose counsel instilled courage in me to pursue this work to this end. Thank you all.

ABSTRACT

Voice will remain a fundamental communication media that cut across people of all walks of life. It is therefore important to make it very affordable. VoIP has to do with digital telephony system that uses IP protocol for voice communication and has been increasingly popular in recent times due to its affordability; however, poor reliability and voice quality remain important factors that limit the widespread adoption of VoIP systems. Good voice quality is a key factor for users transiting from the Public Switched Telephone Network (PSTN) to VoIP networks. It has been shown that line echo is one of the key factors that deteriorate voice quality in VoIP systems. Several non-real-time algorithms have been developed in literature to estimate various aspects of voice quality in VoIP systems. But there is no real-time algorithm that estimates the echo content of a VoIP conversation, which could enable the operator take some corrective actions to improve the quality while the call is in progress. In this thesis, we propose a real-time fuzzy algorithm by engaging fuzzy rules to estimate the strength of the line echo component of the voice quality in VoIP networks. The results obtained shows that the algorithm is able to track and estimate echo content of a live VoIP traffic in real-time. This algorithm could be embedded in VoIP systems to enable operators monitors calls in real-time.

TABLE OF CONTENTS

ABBREVIATIONS

ACOM	Combined loss
CDR	Call detail record
dB	decibel (dB) is a logarithmic unit for the ratio of a physical quantity (sound pressure level)
DSP	Digital signal processing
E-Model	European Telecommunications Standards Institute (ETSI) Computation Model
ERL	Echo return loss
ERLE	Echo return loss enhancement
ETSI	European Telecommunications Standards Institute
Hz	Hertz SI unit of frequency defined as the number of cycles per second
ISDN	Integrated Services Digital Network
IP	Internet Protocol
ITU	International Telecommunication Union
ITU-T	The Telecommunication Standardization Sector of the ITU
MG	Media Gateway
MOS	Mean Opinion Score
ms	milliseconds
PAMS	Perceptual Analysis/Measurement System
PCM	Pulse-code modulation
PESQ	Perceptual Evaluation of Speech Quality
PSQM	Perceptual Speech Quality Measure

PSTN	Public switched telephone network
TDM-IP	Time-division multiplexing over IP network
QoS	Quality of Service
kbps	kilobits per second
Kbps	Kilobytes per second
VoIP	Voice over IP

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background to the study

Voice quality is essential in any communication system that is based on speech transmission. Voice over the Internet Protocol (VoIP) a method of transmitting voice / multimedia service on an IP network have been increasingly popular in the past few years and will continue to spread both in the carrier and enterprise sectors. In fact, current projections estimate that the total market value for services using VoIP is forecast to grow almost tenfold over the next five years (Lee and Cho, 2010). It is clear that VoIP will evolve from being a replacement service for the public switched telephone network (PSTN) market to providing truly converged services to the home and business.

Voice is one of the hardest services to provide on an IP (Internet Protocol) network. The PSTN was built to provide an optimal service for time-sensitive voice applications, with low delay, low jitter, and constant but low bandwidth. IP networks on the other hand have been built to support non-real-time data applications such as email or file transfer. These applications are characterized by bursty traffic, with occasional peaks in demand for high bandwidth, but are not sensitive to delay. (Cisco, 2009) (How stuff works, 2010)During a conversation, humans have little tolerance to delays, jitter, echo (which is a direct consequence of the delay in VoIP networks) and noise (which, for instance, can be introduced during low bit rate voice coding that is commonly implemented by VoIP systems). In addition to the degrading factors introduced in the PSTN, VoIP networks include additional factors such as latency, delay jitter and packet loss Bhatia *et al.*, (2007). In order to provide a good Quality of Service (QoS) for VoIP networks, the existence of an embedded module that assesses the voice quality in each live call is necessary. This embedded module is the main concern of this work.

The performance of a VoIP network can be determined by a variety of parameters such as the availability of the network and dial tone, call setup request processing performance, call completion, call drop rate, one-way voice transport delay and voice quality during the call. Table 1.1 briefly discusses the service requirements that are taken into consideration when evaluating the performance of a VoIP network (Bhatia *et al.*, 2007) (Thompson, 2010).

Table 1.1: Service requirements that are taken into consideration when evaluating the performance of a VoIP network

Service requirements	Parameters
Service requirements before call setup	Availability of dial tone. Availability of computing and network resources for honouring call processing requests.
Service requirements during call setup	Total amount of time to setup a call (can vary from 500 ms to 10 s, depending on availability of the network). The number of simultaneous calls that can be handled without any per call wait.
Service requirements during a VoIP session	Voice coding and processing delay, Voice packet loss. Echo and Jitter
Service requirements after a VoIP session is complete	Maintenance of a complete call log and call detail record (CDR).

Providing good QoS in VoIP networks is of major importance for the transition from the PSTN to VoIP networks. The evaluation of the QoS for a VoIP system or network depends on a set of parameters and requirements that contain those described in Table 1.1. (Bhatia *et al.*, 2007) (Hwang et al., 2010) (Octasic, 2010).

More specifically, in this work we will be interested in evaluating the service requirements during a VoIP session, that is, the voice quality over VoIP networks. One of the main components of the voice quality parameters is the amount of echo present in the conversation; the higher the echo, the lower the voice quality. We intend to use a fuzzy algorithm to evaluate the echo component of voice and also see how efficient and precise the algorithm is.

We can evaluate the quality of voice over IP networks in three different perspectives: the network quality, the objective quality, and the subjective quality, as illustrated by Figure 1.1.

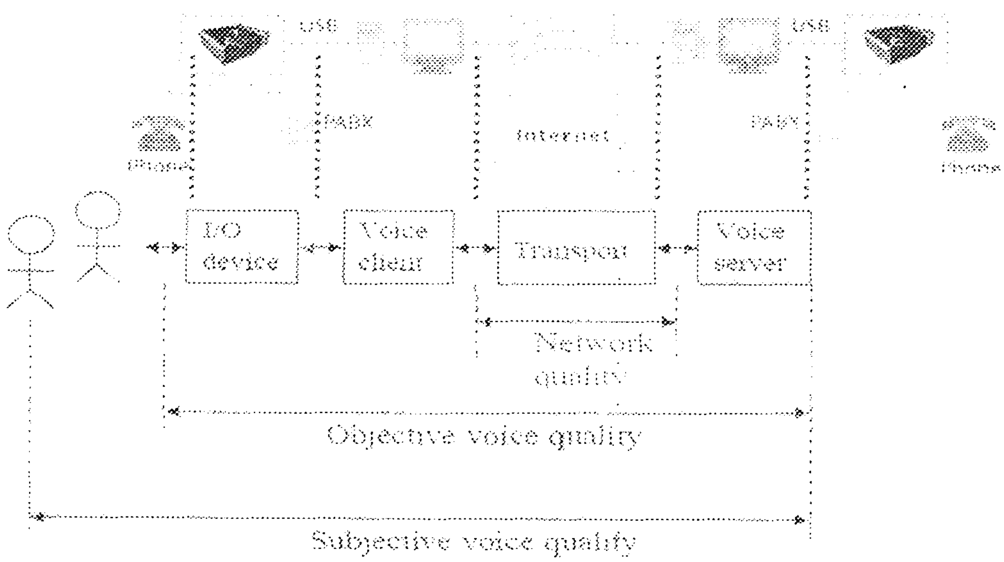

Figure 1.1: Different perspectives for voice quality evaluation in a VoIP network.
Source: Bhatia *et al.*, (2007), Oh and Kim (2010)

The network quality reflects the provider's perspective. The objective and subjective quality reflect the customer's perspective. The network quality can be relatively easily measured by network parameters, such as the packet loss rate or packet delay or jitter. Subjective quality is generally more meaningful than network quality, as it relates directly to user-perceived quality. Assessing subjective voice quality, however, requires listening tests with a large number of test subjects. For this reason, objective quality measures that predict subjective quality are typically employed in the evaluation of voice transmission systems (ITU-T P.800) (Oh and Kim, 2010).

1.2 Objectives

The objective of this work is to develop a simple, objective, low computational complexity using fuzzy inference system to evaluate the echo content of the voice quality over VoIP network in real-time.

In achieving this aim, the relationship between echo control and voice quality in VoIP network is reviewed. An algorithm based on fuzzy logic is developed at the outputs obtained from a standard echo canceller putting in mind the computational complexity of existing non real-time algorithms. Finally, VoIP calls with known MOS (mean opinion score) are tested with the developed fuzzy algorithm and results are shown.

CHAPTER 2

2.0 LITERATURE REVIEW

2.1 Background

Voice over Internet Protocol (Voice over IP, VoIP) is a family of technologies, methodologies, communication protocols, and transmission techniques for the delivery of voice communications and multimedia sessions over Internet Protocol (IP) networks, such as the Internet. Other terms frequently encountered and often used synonymously with VoIP are IP telephony, Internet telephony, voice over broadband (VoBB), broadband telephony, and broadband phone.

VoIP systems employ session control protocols to control the set-up and tear-down of calls as well as audio codecs which encode speech allowing transmission over an IP network as digital audio via an audio stream. The codec used is varied between different implementations of VoIP (and often a range of codecs are used); some implementations rely on narrowband and compressed speech, while others support high fidelity stereo codecs (Wikipedia, 2011).

Voice over Internet Protocol (VoIP) is growing rapidly as a new communication tool (Periakarruppan *et al.*; 2006).Originally, voice traffic was not specifically designed to run on IP networks. The original PSTN system for voice is circuit based while the IP network is packet based.. Because of this, factors like delay, packet loss and jitter set in due to the bursty nature of IP traffic (Jiuchun *et al.*; 2010).

- Transition from PSTN to VoIP has made it difficult to predict and obtain the quality of a VoIP call. A couple of research has been made on how best to evaluate the quality of a VoIP call but popularly accepted method have specific pitfalls. The subjective methods are not real-time, and they are time consuming, The objective methods are better but most of them are not real-time. they are also computationally complex. (Paglierani and Petri, 2009)

2.1.1 Problems of Voice Traffic

A modern real-life telephony network very often consists of local Public Switched Telephone Networks (PSTNs) that collect voice traffic from users; it is connected to a long-distance network optimized for data transport by means of the Internet Protocol (IP). The interconnection of the PSTN to the IP network is implemented through a Media Gateway (MG). which can convert the PSTN signal into the format required by the IP, and vice versa (Manjunath, 2009). According to Paglierani and Petri (2009) in addition to suitable interfaces to the two networks and all the required signalling capabilities. the MG must also provide a number of speech processing functionalities that can considerably influence speech quality: echo control, speech coding, delay variation compensation (jitter buffering), packet loss concealment, and silence suppression but this is not enough.

The main reason for the delayed success of the VoIP might be that the Internet was designed to be a fault tolerant data exchange/transmission medium and the traffic re-routing is the primary target in the case of Internet server additions and removals. The delivery time or transmission delay was never the primary goal in the design phase although the Internet Protocol (IP) has a support for real-time transmission (Ren Jiuchun *et al.*; 2010). According to Periakarruppan *et al.*; (2006) the common problems that occur when a VoIP network is utilized are echo, delay, jitter and loss of packets. All this problems are closely interlinked.

i. Echo

In most cases our everyday conversations take place in the presence of echoes. We hear echoes of our speech waves as they are reflected for instance from the floor and the walls. However, if the reflected waves arrive shortly after we speak them, we do not perceive them as echo but as some reverberation. On the other hand if the reflected wave takes 20 or 30 milliseconds (ms) to come back to us, we will identify it as an annoying echo (Adaptive Digital Technologies, 2010).

Echo is mainly dependent on the amount of delay present in the circuit or network. Most callers will hear echo of their own voice if the circuit contains as little as 30 milliseconds of round-trip delay. If the round-trip delay approaches 50 ms, virtually all callers will complain of echo if it is left uncontrolled.(Falk and Chan, 2008)(Periakarruppan *et al.*; 2006)

ii. Delay

Delay is introduced into the telecommunications network primarily by transmission facilities and transmission equipment. Negligible delay is introduced into the telecommunications network by some types of transmission equipment, such as a digital switch. Other transmission equipment, such as low bit rate voice encoders, often introduces significant delays. Depending on the network topology, and the type of transmission equipment used in the network, 30 ms of roundtrip delay can occur in connections that are across country or just across town(Lakaniemi *et al.*; 2001).

The International Telecommunication Union (ITU) has a guide for the amount of delay introduced by specific transmission medium as shown in Table 2.1.

Table 2.1: Some typical transmission facility delays

Transmission Facility	Delay per 100 miles
T1 carrier over copper	1ms
Fiber optic cable	1ms
Microwave Radio	0.7ms

Source: ITU-T G.165

iii. Jitter

Jitter is another problem encountered by voice calls; it is caused by retransmission of lost packet which is directly linked to delay in the IP network. Delay can be caused by different factors already mentioned in section ii above.

iv. Packet Loss

Packet Loss can occur for a variety of reasons; these include link failure; traffic congestions; and misrouted traffic. In an IP environment, the packets are retransmitted or re-routed and this causes delay (Jiuchun *et al.*; 2010)(Periakarruppan *et al.*; 2006)

2.2 Voice Quality Measurement

As shown in Figure 1.1, the voice quality of a VoIP call can be measured using different methods mainly by obtaining the network quality, subjective quality and objective quality.

2.2.1 Network Quality

In general, poor network quality decreases the performance of a VoIP system. In VoIP applications, delay, jitter and packet loss are the main network impairments that affect perceived voice quality. Jitter can be partially compensated for by using a playout buffer at the receiving end, but this introduces further delay and additional packet loss (Yackoski and Shen, 2010) (Narbutt and Davis, 2007) (Hwang *et al.*, 2010).

As mentioned in section 2.1.1, there are several components (logical and physical) in the IP network that cause delay, jitter and packet loss.

Network reliability is an important component that introduces delay and packet loss, especially in the backbone of IP networks. There are two important scenarios that can directly influence the network reliability: (a) link failures, and (b) routing reconfiguration.

(a) Link failure:

There are many reasons that can lead to link failures such as fibre cuts, router crashes and maintenance operations. In fact, long outage durations are typically attributed to a link failure in the IP network backbone (Bhatia *et al.*, 2007).

(b) Routing reconfiguration:

It is typical for a routing protocol to require around five seconds to converge to a new configuration when a link goes down and around fifteen seconds when a link goes up During this reconfiguration period, forwarding may be disrupted and voice packets may be lost. All the network behaviour described above can influence the amount of delay and packet loss present in a VoIP system. When this happens, the IP network can exhibit undesirable characteristics, such as large delay spikes, periodic delay patterns and packet loss on one or more paths. All these lead to poor VoIP performance.

2.2.2 Subjective Voice Quality

The MOS is a subjective voice quality assessment method. It is considered by many researchers as the best evaluation method for assessing voice quality because its result is based on the human direct ears. The MOS is a subjective rating system that is defined in ITU-T P.800. It is based on the opinions of several testing volunteers who listen to a sample of voice traffic and rate the quality of that transmission. The volunteers listen to a variety of voice samples and are asked to consider factors such as loss, noise and echo. The volunteers then rate the voice samples by giving a score in range from 1 to 5 as described in Table 2.2. The MOS score is calculated as an average of scores given by all listeners (ITU-T P.800).

While MOS represents the true perceptual assessment of speech quality, it has obvious limitations. It is a time consuming process, it is not an automated method and it cannot be applied to estimate the quality of a call in a real-time environment.

It is interesting to note that even using this time consuming MOS methodology, most experiments can only indicate the speech quality of unidirectional connections as stated by Bhatia *et al.*,(2007), Ngamwongwattana and Thompson (2010). For instance, the MOS test does not indicate how the increased delay degrades the final QoS due to decreased interactivity when long transmission delays are introduced.

Table 2.2: Description of MOS scores

MOS Score	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Source: Kang *et al.*, (2001)

2.2.3 Objective Voice Quality

One of the advantages of objective voice quality algorithms over subjective voice quality algorithms is that objective algorithms can be automated and may not require any human intervention at all. There are two main classes of objective voice quality algorithms: active and passive algorithms.

Objective voice quality monitoring, whether active or passive, has recently gained ground among VoIP providers. In active monitoring, a network analyzer injects traffic patterns that resemble a VoIP application into the network; the analyzer then observes the overall voice quality by comparing the impaired voice with the original voice sample using a perceptual model. Although this scheme can provide useful input for optimization and network dimensioning, it uses network resources, provides non real-time results, and can't concretely determine the causes of degradation.

A passive monitoring scheme, on the other hand, can operate in real-time, and lets VoIP applications take corrective action when Quality of Service is unacceptable. For these reasons, the algorithm we propose in this work is a passive, objective voice quality algorithm.

Several active, objective algorithms have been proposed to automate the voice quality assessment of a call. The most successful two methods are the PAMS – Perceptual Analysis / Measurement System - and the PSQM - Perceptual Speech Quality Monitor (Falk and Chan, 2008). According to Paglierani and Petri (2007), Qiao *et al.*; (2008) in both methods a reference speech sample representing the transmitted speech signal is passed through degradation producing the degraded speech sample representing the received signal. Signal analysis is performed both on the time and frequency domains of the two speech samples and an estimate of the MOS score is provided. This requires a large amount of computing power which means it cannot be embedded into a VoIP network.

PSQM as originally conceived was not developed to account for network Quality of Service perturbations common in Voice over IP applications, items such as packet loss, delay variance (jitter) or non-sequential packets. (Qiao *et al.*; 2008) (Manjunath, 2009). According to Jiuchun *et al.*; (2010) other voice quality measurement algorithms like the E-Model (European Telecommunications Standards Institute (ETSI) Computation Model) developed by an ETSI work group chosen by ITU is different from other methods because it represents also a network simulation tool. However, the effect of delay is not considered in the E-MODEL.

2.3 Echo Effect

With respect to Table 2.1, there is a limit for the amount of delay that will be tolerated by an average user. The ITU-T Recommendation G.114 provides limits for one-way transmission time (delay) on connections with adequately controlled echo (Jiuchun *et al.*; 2010) (Kang *et al.*, 2001).

Researchers have worked on evaluating the voice quality of a VoIP call by improving on previous methods mainly PAMS, PESQ and the E-Model. They have proposed modified E-Model (Jiuchun *et al.*; 2010), they also proposed an introduction of a Packet based Echo Canceller and others but none has been real-time. Figure 2.1 shows how echo affects voice calls.

Table 2.3: ITU-T limits for one-way transmission time (delay) with echo control

One-way transmission time	User acceptance
0 to 150ms	Acceptable for most users
150 to 400ms	Acceptable but has impact
400ms and above	Unacceptable

Source: Choudhry *et al.*; (2010), Wikipedia, (2010)

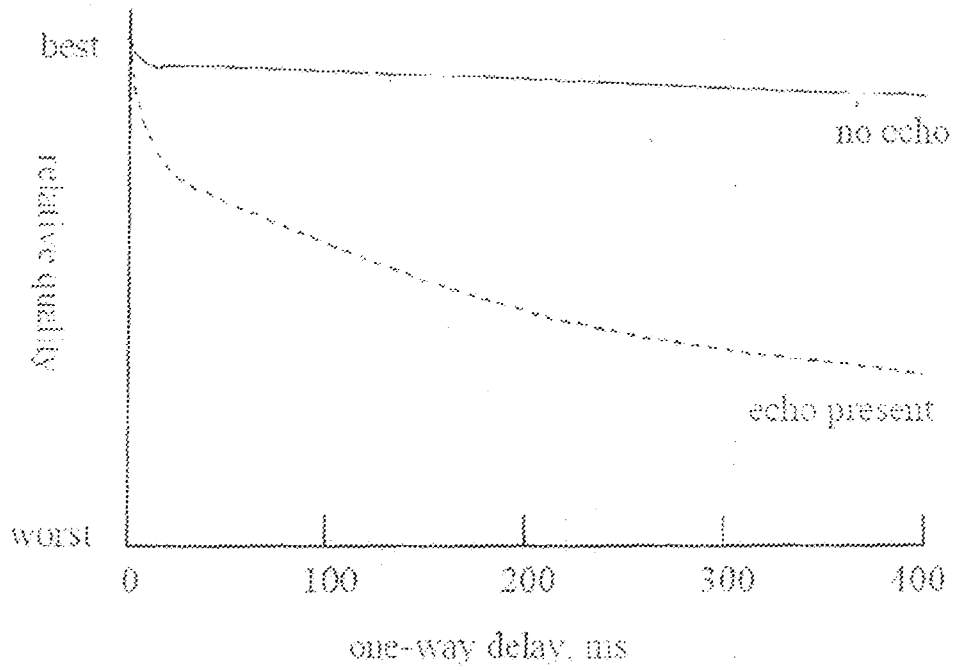

Figure 2.1: Impact of delay on call quality with and without echo
Source: Periakarruppan *et al.*; (2006)

It is clear from Figure 2.1 that echo is a determinant component in decreasing the quality of the call. Echo can be loud and it can be long. The louder and longer the echo, the more annoying it becomes.

2.4 Echo Canceller

An echo canceller monitors speech from the far end that passes through its receive path and uses this information to compute an estimate of the echo that is then subtracted from its send path. If the estimation is good, the echo is cancelled and only the near end speech is sent to the far end. Good echo cancellation is essential for the quality of the voice in the network.

Echo cancellation occurs between the send-in and send-out ports, reducing the echo present in the send path. The total amount of echo attenuation that an echo canceller provides is called echo return loss enhancement (ERLE). ERLE is the difference in the echo level between the send- in and send-out ports and it is measured in dB (Octasic, 2010). The Block diagram of a line echo canceller is shown in Figure 2.2.

The two key parameters used by echo cancellers which are included in our proposed algorithm are: Echo return loss (ERL), which is the amount of echo attenuation provided by the hybrid, that is, the attenuation of the signal from the Rout port to the Sin port of the echo canceller. The ERL is measured in dB.

Combined loss (ACOM), which is defined by the sum (in dB) of the ERL, the attenuation provided by the adaptive filter (cancellation loss) and the attenuation provided by the nonlinear processor (nonlinear processing loss). (Octasic, 2010)(ITU-T G.165)

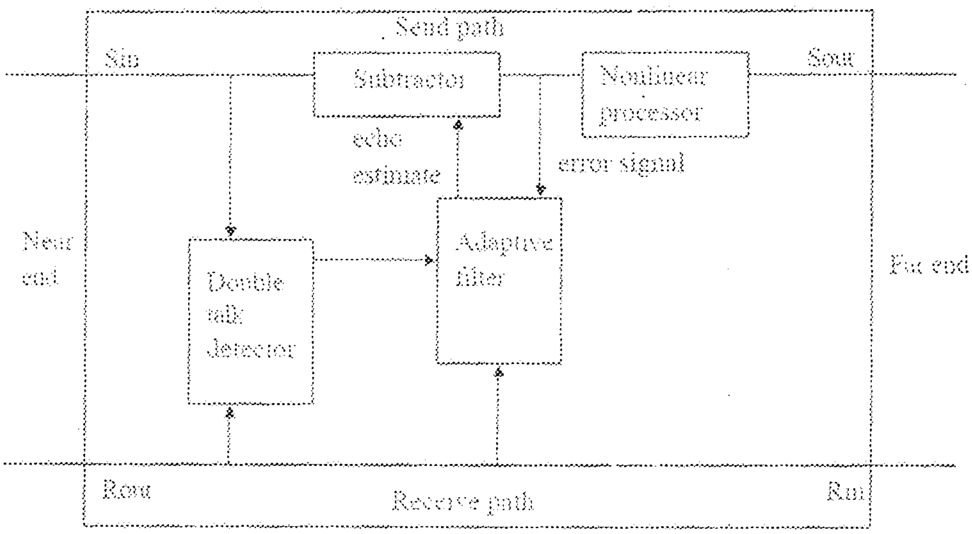

Figure 2.2: Block diagram of a line echo canceller
Source: Adaptive Digital Technologies, (2010)

In the foregoing, we have been able to explain the relationship between echo and delay, the algorithm we propose is designed to make voice quality measurement real-time and less computationally complex by using fuzzy logic to vary the input parameters. It means it can be embedded into existing VoIP systems. This is achieved by using parameters obtained from existing echo cancellers in the system.

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Fuzzy Logic

Fuzzy logic starts with and builds on a set of user-supplied human language rules. The fuzzy systems convert these rules to their mathematical equivalents. This simplifies the job of the system designer and the computer, and results in much more accurate representations of the way systems behave in the real world (Ibrahim, 2004) (Wikipedia, 2010) In other words, a fuzzy model of a system is a set of fuzzy rules by which the behavior of the system is approximately emulated

Fuzzy logic requires some numerical parameters in order to operate such as what is considered significant error and significant rate-of-change-of-error, but exact values of these numbers are usually not critical unless very responsive performance is required in which case empirical tuning would determine them. For example, a simple temperature control system could use a single temperature feedback sensor whose data is subtracted from the command signal to compute "error" and then time-differentiated to yield the error slope or rate-of-change-of-error, hereafter called "error-dot". Error might have units of degs F and a small error considered to be 2F while a large error is 5F. The "error-dot" might then have units of degs/min with a small error-dot being 5F/min and a large one being 15F/min. These values don't have to be symmetrical and can be "tweaked" once the system is operating in order to optimize performance. Generally, fuzzy logic is so forgiving that the system will probably work the first time without any tweaking (Emily, 1993).

3.1.1 Advantages of Using Fuzzy Logic

Key benefits of fuzzy logic are its simplicity and its flexibility. Fuzzy logic can handle problems with imprecise and incomplete data, and it can model nonlinear functions of arbitrary complexity. "If you don't have a good plant model, or if the system is changing, then fuzzy will produce a better solution than conventional control techniques (Imperial College London, 2008).

By selecting the number of fuzzy representative sets, there is a way of adjusting the precision level of a solution. If more fuzzy sets are used in design, systems will require more memory and faster CPUs. At the limit, the number of fuzzy sets becomes equal to the number of crisp data points. That represents the most precise and costly solution.

Two important characteristics of successful fuzzy systems are:

1. The fuzzy systems are simple in terms of their objective and structure - which we call a fuzzy inference system.
2. The fuzzy systems employ solutions articulated in daily language by means of IF-THEN fuzzy rules (Ross, 2004) (Wikipedia, 2010).

A successful fuzzy system is robust, has adjustable precision and when compared with traditional systems of computation they are more practical and cost effective.

3.1.2 Fuzzy Logic and Embedded Systems

An algorithm to evaluate the voice quality in a VoIP system or network should be a real-time algorithm in order to give operators the precise current voice quality in their network and a chance to react as fast as possible when the quality drops. There are two distinct options of where such algorithm could run:

1. It could run inside the embedded system that processes the call.
2. It could run in a network server.

In the first option, there is a disadvantage that embedded systems normally have limited processing power, which is used for high priority tasks like call control, speech compression and echo cancellation. It is common to have a situation where such embedded systems are working very close to their processing capacity.

An advantage of this approach is that each embedded system can take action based on the real-time results of the algorithm and try to improve its performance without having to rely on decisions made by a remote server that may even be offline for some reason (Ibrahim, 2004).

On the other hand, option two seems to relieve the embedded system of such high processing requirement, once the algorithm would be running in a server somewhere in the network. This approach has a tremendous disadvantage of requiring all the embedded systems in the network to send information about each of their calls to this centralized server (which could be one or more servers).

In this case, the bandwidth of the network is compromised. It also has the disadvantage of removing from the network device the ability of monitoring its own voice quality, generate alarms or even try some self- fixing action. It should also be noticed that even in this approach there is some extra processing required from the embedded systems, once they will have to code the required information and access the network in order to send it to the server.

In this work, we focus on the first option and we use the ideas of fuzzy logic to develop an algorithm that requires low processing power from the embedded system that carries the VoIP application.

3.2 Objective Evaluation of the Voice Quality

In Section 1.1 we described the several parameters that can be used to evaluate the performance of a VoIP system or network (Table 1.1) and showed three different perspectives for voice quality evaluation in a VoIP network in Section 2.2. In this work we are focusing on measuring the voice quality in a VoIP system or network using echo as the key factor.

In this section we develop an algorithmic tool for evaluating the echo content of a voice traffic. We used the objective evaluation technique for the following reasons:

1. Contrary to the subjective quality method, an objective method can be automated since it does not require human intervention or feedback. If well designed, such a method can also estimate the quality in real- time.
2. Contrary to the network quality method, an objective method takes into consideration the user's perception of the call and not only parameters that qualify the performance of the IP network.

As is described by Ditech Networks (2007), there are three classes of objective voice quality evaluation metrics: (i) network-parameter based metrics, (ii) psycho-acoustic metrics, and (iii) elementary metrics.

i. Parameter-based metrics do not consider the actual voice signal. Instead, these metrics sum impairment factors that characterize the individual components of the communication system. For instance, in the E-model the packet loss and delay in a VoIP system are translated into impairment factors.

   Parameter-based metrics such as the E-model hold promise for predicting subjective voice quality but still requires extensive refinements and verifications.

ii. Psycho-acoustic metrics transform voice signals to a reduced representation to retain only perceptually significant aspects. These metrics aim to predict the subjective quality over a wide range of voice signal distortions. One example of such metric is the PESQ algorithm

iii. Elementary objective voice quality metrics rely on low-complexity signal processing parameters and techniques to predict subjective voice quality. Elementary metrics generally have smaller correlations with subjective voice quality than highly complex psycho-acoustic metrics and do not provide the perception modeling needed for psycho-acoustic coder algorithm development. However, elementary metrics represent a good engineering tradeoff for communication and networking system researchers and developers in that they allow for fairly detailed conclusions about voice quality while having low computational complexity.

The various aspect of voice quality evaluation is summarised in Figure 3.1

$$
Voice\ Quality \begin{cases} Network\ Quality \\ Subjective\ Voice\ Quality \\ Objective\ Voice\ Quality \rightarrow \begin{cases} Parameter - based\ metrics\ (PAMS) \\ Psycho - acoustic\ metrics\ (PSQM) \\ Elementary\ metrics\ (Proposed\ Method) \end{cases} \end{cases}
$$

Figure 3.1: Classification of voice quality algorithms for VoIP systems

The proposed fuzzy algorithm estimates the echo quality factor of the voice quality and should serve as a building block for an objective, passive, voice quality algorithm based on elementary metrics.

As was mentioned in Chapter 1, the critical issues in delivering good voice quality over IP networks are: packet loss, delay, echo and jitter. These issues are all correlated, but there is a stronger correlation between jitter, delay and packet loss. Jitter in VoIP systems is normally compensated for by using a playout buffer at the receiving end, which introduces delay and additional packet loss. So we can imagine a fuzzy inference system that evaluates the voice quality in a VoIP network described by Figure 3.2:

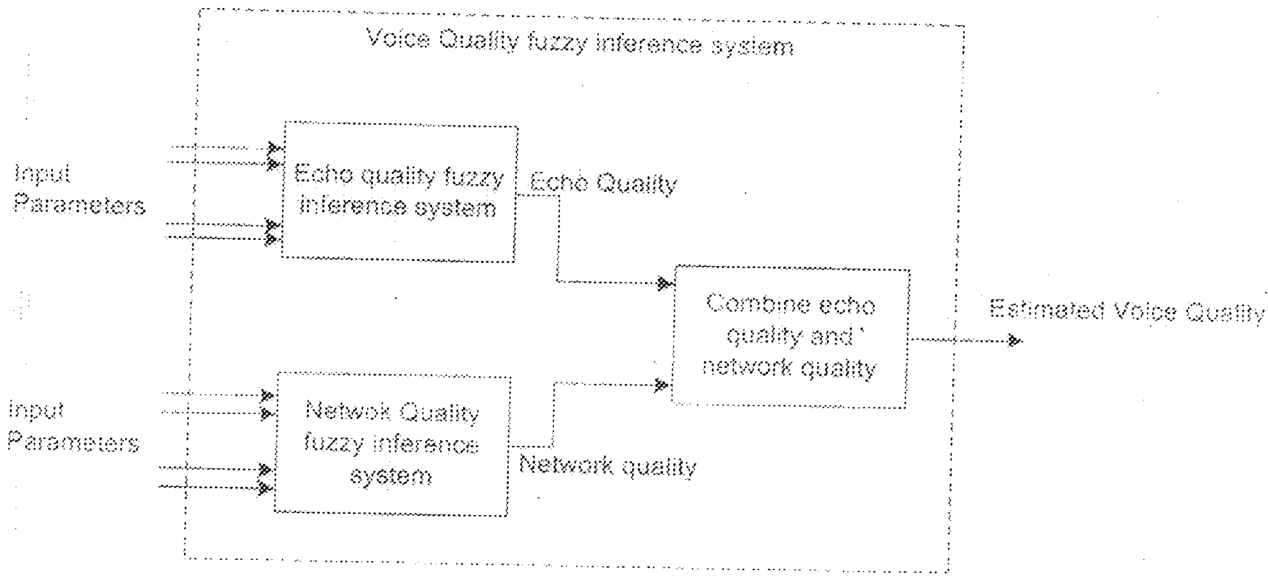

Figure 3.2: Proposed Fuzzy inference system to estimate the voice quality in a VoIP network

The features of the proposed algorithm are:

a. Obtain the echo quality component of an objective voice quality algorithm based on elementary metrics to estimate the voice quality in a VoIP system

b. The algorithm must have a computational complexity low enough such that it can run in an embedded module inside every VoIP channel in the VoIP system or network.

c. The algorithm must be able to run for live VoIP calls without the need of a reference signal, which is one of the limitations of several objective voice quality methodologies including the PSQM.

d. The algorithm must give a real-time estimation of the echo signal by outputting a few parameters (or scores).

3.3 Fuzzy Input and Output Membership Function for the Echo Component of the Voice Quality

In order to obtain a real-time, low computationally complex algorithm we chose to use as inputs to our fuzzy inference system parameters that are already being computed or estimated by the echo canceller, such as estimates for the Echo return Loss(ERL) and the Combined Loss(ACOM) (Section 2.4), speech power estimations and noise power estimations.

We should note that although the fuzzy logic implementation results in low computational complexity it has the disadvantage of not being precise. As we said before, it reflects approximate human reasoning and it will never be as good as the subjective MOS and it won't be as precise as the PSQM or PESQ.

Different echo cancellers may estimate a different set of parameters and in this case we need extra computations to estimate the required parameters for the algorithm proposed here. The input parameters that we will use are:

i. Echo return loss (ERL) - Section 2.4
ii. Combined loss (ACOM) - Section 2.4
iii. Receive speech power - an estimate of the speech power in the receive path (Figure 2.2)
iv. Receive noise power - an estimate of the noise in the receive path (Figure 2.2)
v. Transmit speech power - an estimate of the speech power in the send path (Figure 2.2) after the echo cancellation

vi. Transmit noise power - an estimate of the noise in the send path (Figure 2.2)

Table 3.1 shows the fuzzy sets associated with the input parameters propsed.

Table 3.1: Fuzzy sets associated to the input parameters

Fuzzy Set	Description
Good ERL	Represents values of ERL that will help the echo canceller to realize a good echo cancellation.
Bad receive speech Power	The receive speech powers in this set are either to low or to high, making it difficult for the echo canceller to generate the signal that must be subtracted in the send path.
Bad transmit noise Power	Represents values of the transmit noise that may disrupt the convergence of the adaptive filter.
Bad ACOM	With high probability, VoIP systems with ACOM values in this set will have echo problems and the voice quality will be bad.
Moderate ACOM	Represents values of ACOM that may indicate that the echo cancellation was not good enough and some echo may be leaked to the far end.
Good ACOM	VoIP systems with ACOM values in this set are able to cancel most of the echo in the calls.

We define the following three output membership functions which will give an estimate of the echo component of the voice quality.

a. The membership function for the fuzzy set "bad echo (be)"

$$\mu_{be}(x) = \begin{cases} 1 - 2x, 0 \leq x \leq {}^{1}/_{2} \\ 0, otherwise \end{cases} \quad (3.1)$$

b. The membership function for the fuzzy set "moderate echo (me)"

$$\mu_{me}(x) = \begin{cases} 2x, 0 \leq x \leq 1/2 \\ 2(1-x), 1/2 < x \leq 1 \\ 0, otherwise \end{cases} \tag{3.2}$$

c. The membership function for the fuzzy set "good echo (ge)"

$$\mu_{ge}(x) = \begin{cases} 2x-1, 1/2 \leq x \leq 1 \\ 0, otherwise \end{cases} \tag{3.3}$$

Graphically we have

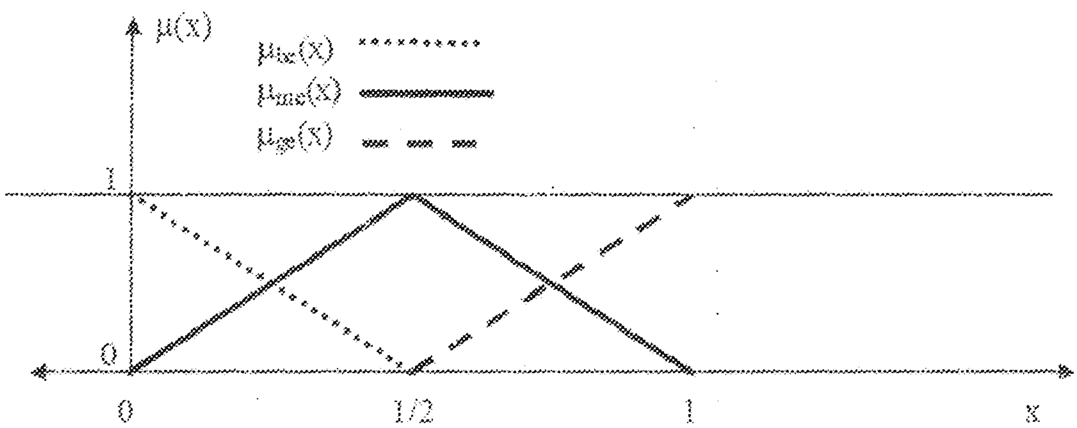

Figure 3.3: Output membership function for Echo

The approach we used is to define the input membership functions based on empirical reasoning and then we spent some time tuning those membership functions. But the tuning is done after the fuzzy inference system is designed because we need to use the output of the algorithm as a feedback for tuning. We will describe the input membership functions later in this section. First we will describe our proposed fuzzy inference system with the complete set of fuzzy rules, operations, and defuzzification.

The fuzzy rules together with the fuzzy membership functions are the main elements that reflect the empirical reasoning behind the proposed fuzzy inference system. Table 3.2 describes the fuzzy rules that we adopted for our proposed algorithm and the empirical reasoning behind each rule.

Table 3.2: Fuzzy rules to evaluate the echo component of the voice quality

Fuzzy Rule	Empirical reasoning
IF ACOM is bad THEN echo is bad.	The ACOM is a major parameter for estimating the quality of the echo signal (Section 2.4). If the ACOM is bad, most probably the user is perceiving echo.
IF ACOM is good, THEN echo is good.	With a good ACOM, some echo is being cancelled successfully, independently of the other parameters.
IF ACOM is moderate AND ERL is good THEN echo is moderate.	If ACOM is moderate, there is some uncertainty about the quality of the echo signal. So we use the ERL to better estimate it.
IF receive speech power is bad AND transmit noise is bad THEN echo is bad	The signal levels for transmit and receive speech as well as for transmit noise are all contributing to a bad echo signal.

The fuzzy implication operator that we chose for our proposed algorithm is Larsen Operator. The defuzzification method that we chose is the center of mass (centroid) method (Ross, 2004).

An advantage of using fuzzy logic is that we can first define the fuzzy input variables and elaborate the fuzzy rules and then we can tune the membership functions by running the algorithm for calls for which we know the MOS. That is exactly what we did in order to define the following membership functions for each one of the fuzzy inference input variables.

We used Matlab and its fuzzy logic toolbox to implement our proposed algorithm and ran a set of 16 calls. We implemented it in a way that we would give to Matlab 2008a different fuzzy inference systems at a time. The difference between the fuzzy systems was only in terms of the membership functions. All systems had the same input variables, fuzzy operations, and fuzzy rules, but different membership functions for the fuzzy variables. Then, after running all 16 calls for each one of the 5 fuzzy inference systems we could compare the result of the fuzzy algorithm to the expected MOS scores (after the echo cancellation).

We chose to use 5 different fuzzy inference systems in each tuning step because there are so many parameters that can be changed in a membership function that one easily gets lost if one tries to change several parameters at once. So, for instance, if we are tuning a specific triangular membership function that has a positive and a negative slope we would first tune the positive slope of the triangle and try it with say 3 to 5 different positive slopes.

As a result of the tuning process described above we derived the following membership functions for the fuzzy sets described in Table 3.1.

3.3.1 Echo Return Loss (ERL)

The membership function for the fuzzy set "good ERL (gerl)"

$$\mu_{gerl}(x) = \begin{cases} x - 20/_{10}, 20 \leq x \leq 30 \\ 1, x > 30 \\ 0, otherwise \end{cases} \qquad (3.4)$$

Graphically

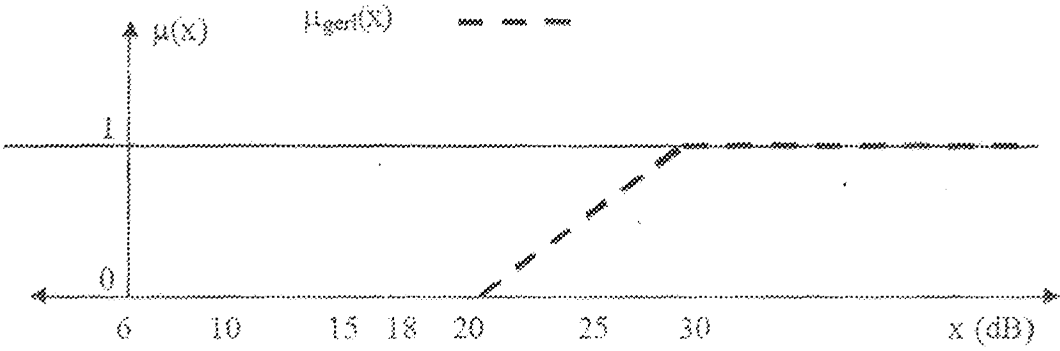

Figure 3.4: ERL fuzzy membership function

3.3.2 Combined Loss (ACOM)

The membership function for the fuzzy set "bad ACOM (bacom)"

$$\mu_{bacom}(x) = \begin{cases} {}^{23-x}/_{17}, 6 \le x \le 23 \\ 1, x < 6 \\ 0, otherwise \end{cases} \tag{3.5}$$

The membership function for the fuzzy set "moderate ACOM (macom)"

$$\mu_{macom}(x) = \begin{cases} {}^{x-12}/_{11}, 12 \le x \le 23 \\ {}^{36-x}/_{13}, 23 < x \le 36 \\ 0, otherwise \end{cases} \tag{3.6}$$

The membership function for the fuzzy set "good ACOM (gacom)"

$$\mu_{gacom}(x) = \begin{cases} {}^{x-23}/_{17}, 23 \le x \le 40 \\ 1, x > 40 \\ 0, otherwise \end{cases} \tag{3.7}$$

Graphically

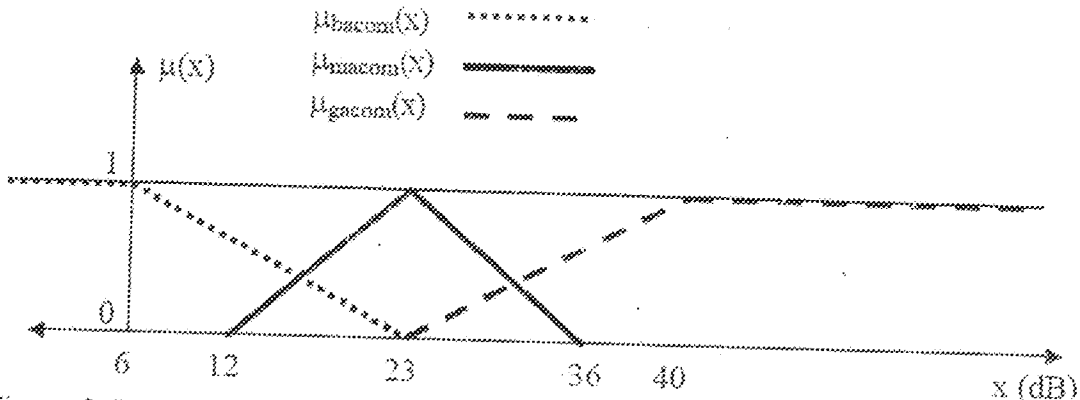

Figure 3.5: ACOM fuzzy membership function

3.3.3 Receive Speech Power

The membership function for the fuzzy set "bad receive speech power (brsp)"

$$\mu_{brsp}(x) = \begin{cases} -25 - x/_5, -30 \le x \le -25 \\ 0, -25 < x < -15 \\ x + 15/_{10}, -15 \le x \le -5 \\ 1, otherwise \end{cases} \qquad (3.8)$$

Graphically

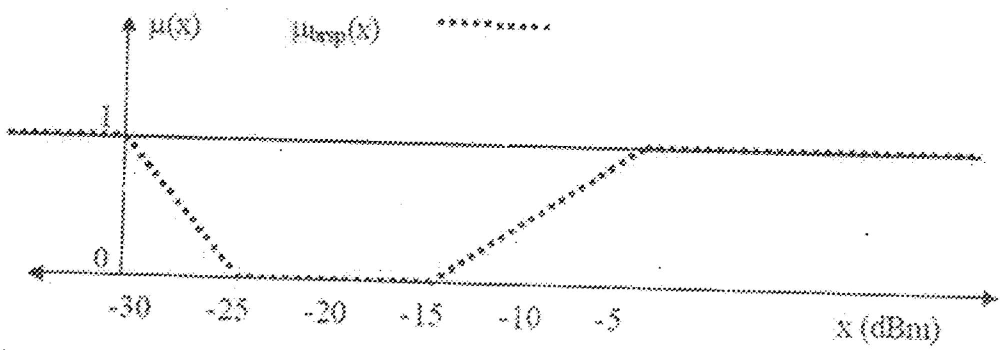

Figure 3.6: Receive speech fuzzy membership function

3.3.4 Transmit Noise Power

The membership function for the fuzzy set "bad transmit noise power (btnp)"

$$\mu_{btnp}(x) = \begin{cases} x + 45/_9, -45 \le x \le -36 \\ 1, x > -36 \\ 0, otherwise \end{cases} \quad (3.9)$$

Graphically

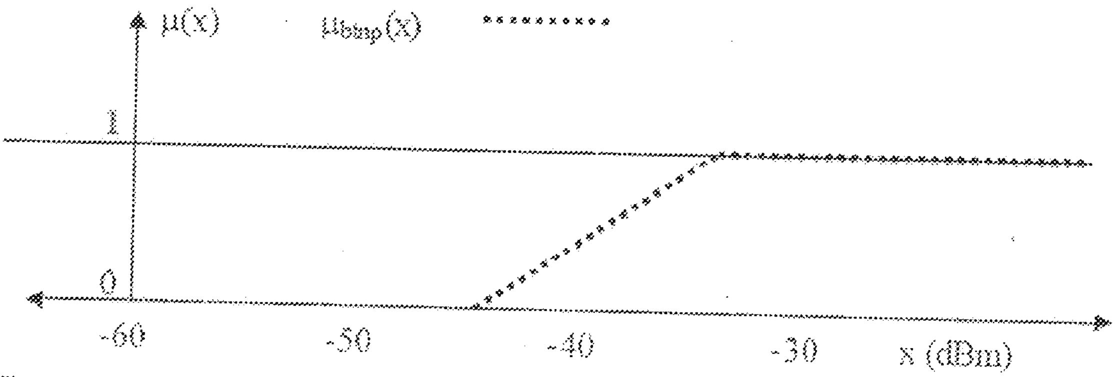

Figure 3.7: Transmit noise fuzzy membership function

3.3.5 Computational Example

Now we show an example of a fuzzy rule computation carried out by the fuzzy inference system proposed above. Suppose we have the following set of inputs (raw data)

ERL = 23 dB
ACOM = 28dB
Receive speech power = -27dBm
Transmit noise power = -50dBm

Suppose we want to compute the fuzzy rule "*IF ACOM* is moderate *AND ERL* is good *THEN* echo is moderate"

The fuzzy set moderate ACOM is described with the membership function $\mu_{macom}(x)$

The fuzzy set good ERL is described with the membership function $\mu_{gerl}(x)$

$$\mu_{macom}(x) = \begin{cases} x - 12/_{11}, 12 \le x \le 23 \\ 36 - x/_{13}, 23 < x \le 36 \\ 0, otherwise \end{cases} = \mu_{macom}(28) = 36 - 28/_{13} = 8/_{13} \quad (3.10)$$

$$\mu_{gerl}(x) = \begin{cases} x - 20/_{10}, 20 \le x \le 30 \\ 1, x > 30 \\ 0, otherwise \end{cases} = \mu_{gerl}(23) = 23 - 20/_{10} = 3/_{10} \quad (3.11)$$

The fuzzy implication operator (AND / intersection) that we are using is the Larsen operator, according to T. Ross (2004) the output of this rule is given by:

$$\mu_I(x) = \min\left(\mu_1(x), \mu_2(x), \dots, \mu_N(x)\right) \ for\ all\ x \in \Re$$

where N is the number of fuzzy sets, $\Re$ indicates real numbers and $\mu_N(x)$ are the membership functions

The membership function for the fuzzy set "moderate echo (me)"

$$\mu_{me}(x) = \begin{cases} 2x, 0 \le x \le 1/_2 \\ 2(1 - x), 1/_2 < x \le 1 \\ 0, otherwise \end{cases} \quad (3.12)$$

Therefore $\mu(x) = \frac{3}{10}\mu_{me}(x) = \begin{cases} \frac{3}{5}x, 0 \le x \le 1/_2 \\ \frac{3}{5}(1 - x), 1/_2 < x \le 1 \\ 0, otherwise \end{cases}$ (3.13)

Figure 3.8 is a graphical representation of the computation described above. This helps in tuning the membership functions.

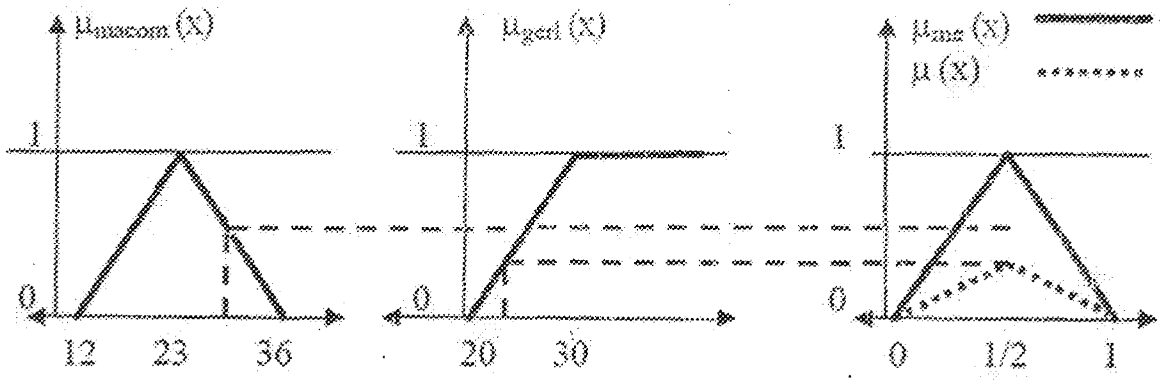

Figure 3.8: Graphical interpretation of the fuzzy rule

Once all the rules described in Table 3.2 are computed as shown above, the output membership functions ('consequences' of the fuzzy rules - which is $\mu(x)$ in the example above) are then aggregated and the output of the fuzzy inference system is computed using the center of mass defuzzification method.

The idea is to run such computation periodically throughout a VoIP call, getting instantaneous values of the quality of the echo signal in such calls. This was done in our simulations and the next chapter shows the output of the fuzzy algorithm for a call with a relatively bad echo signal. Figure 3.9 shows the flowchart for this fuzzy algorithm

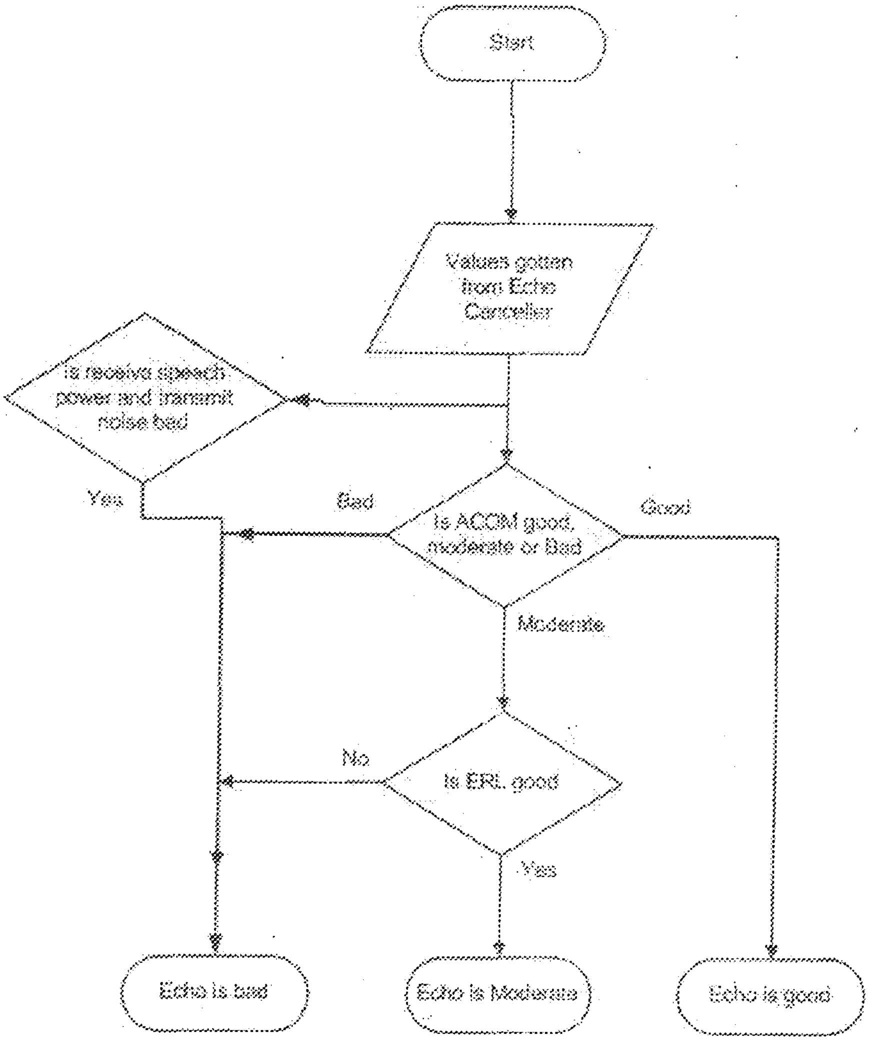

Figure 3.9: Flow chat for the Fuzzy Algorithm

CHAPTER FOUR

4.0 RESULTS

4.1 Simulation Results

We ran the proposed algorithm for 16 calls with different levels of voice quality. These calls were all previously generated during a live VoIP call and recorded in PCM format. We split the calls into two groups. The first group contains 8 calls where echo was effectively cancelled or no echo was created in that call. The second group contains calls with various levels of echo signals and background noise. Then, we ran these calls through a Texas Instruments DSP platform with a line echo canceller. We then collected the measurements from this echo canceller, analyzed these measurements and finally used them as inputs to our fuzzy algorithm.

As described in the previous section, the collected measurements from the echo canceller were the ERL, ACOM, transmit speech power, receive speech power, transmit noise power and receive noise power as shown in Table 4.1

For each simulated VoIP call we had the near end signal, the far end signal and the measurements provided by the echo canceller. As was described, the proposed algorithm gives an instantaneous estimation of the echo signal for the call. During our simulations we chose to compute this estimation every two seconds. The approach that we considered was to provide an average of the outputs of the fuzzy inference system throughout the call. Figure 4.1 shows a call with the duration of about 60 seconds.

Table 4.1: Average reading for 16 VoIP calls

Call Number	ERL (db)	ACOM (db)	Recive Speech power (dBM)	Transmit Noise power(dBM)	Fuzzy Output	Call Quality
1	20.00	21.80	-46.50	-11.60	0.441	Bad
2	20.30	22.30	-45.50	-11.61	0.383	Bad
3	20.39	22.33	-50.24	-15.55	0.412	Bad
4	20.39	22.34	-49.34	-13.67	0.413	Bad
5	20.41	22.37	-49.50	-13.10	0.423	Bad
6	20.96	21.19	-47.70	-14.14	0.402	Bad
7	24.38	28.55	-46.06	-18.55	0.562	Bad
8	21.98	21.52	-43.53	-13.20	0.470	Bad
9	25.90	34.90	-47.50	-15.70	0.780	Good
10	27.33	35.47	-48.15	-19.41	0.810	Good
11	27.23	35.34	-45.61	-18.52	0.803	Good
12	26.80	35.09	-49.64	-18.67	0.790	Good
13	27.90	35.67	-48.75	-20.55	0.820	Good
14	27.67	35.43	-48.60	-20.86	0.808	Good
15	27.38	35.28	-49.34	-19.14	0.800	Good
16	26.92	35.11	-49.19	-17.89	0.791	Good

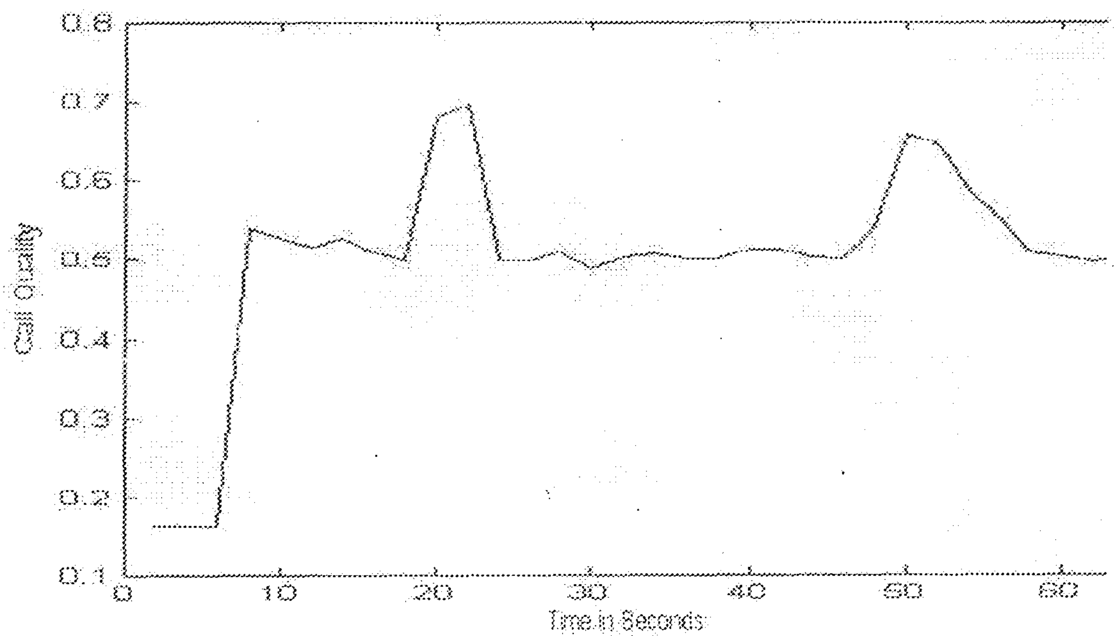

Figure 4.1: Call quality estimation for a simulated call

Figure 4.1 is a call with duration of about one minute. We computed the quality of the echo signal every two seconds. The higher the number for the estimation of echo quality, the better, it is as described by the membership functions of Figure 3.4

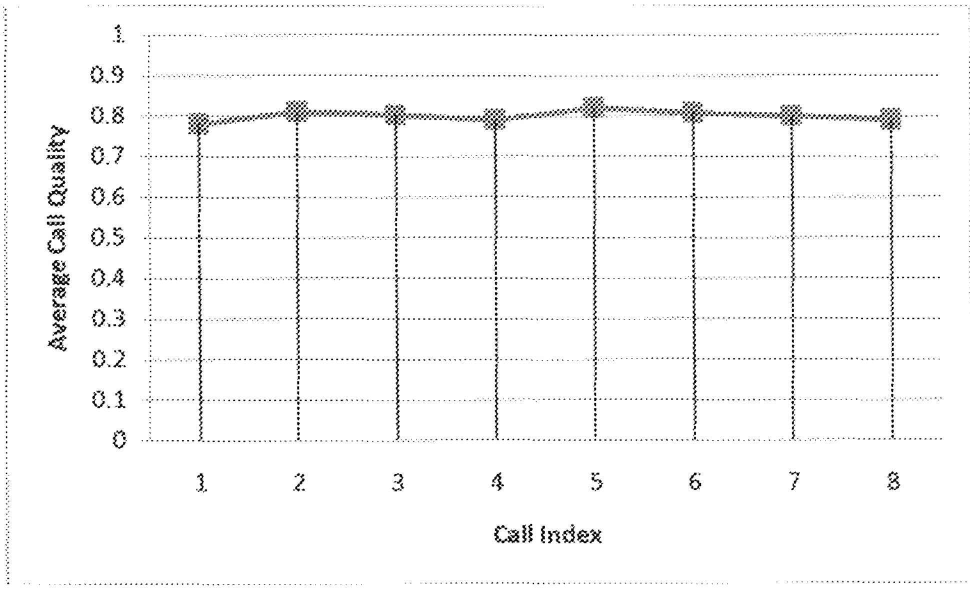

Figure 4.2: Average quality for calls with reduced echo component

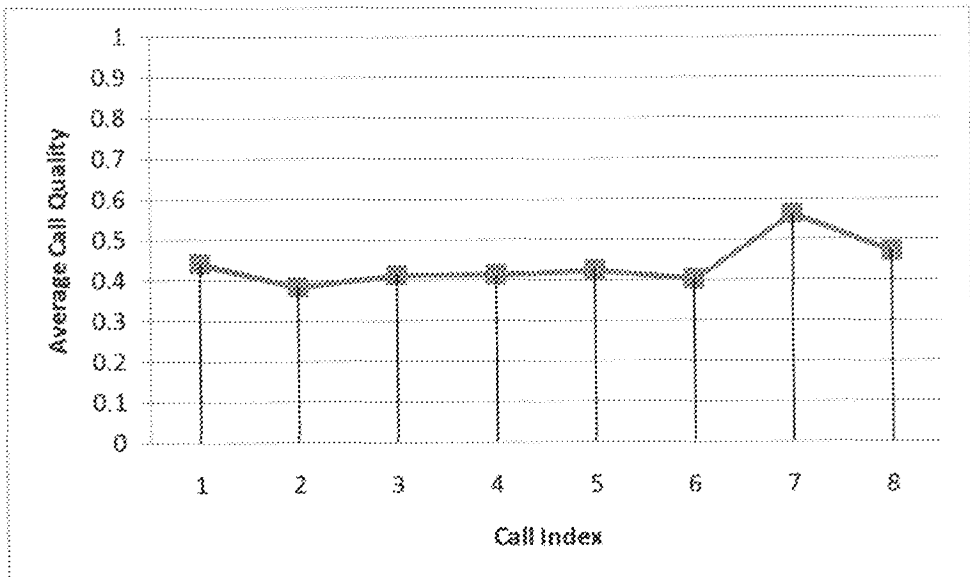

Figure 4.3: Average quality for calls with increased echo component

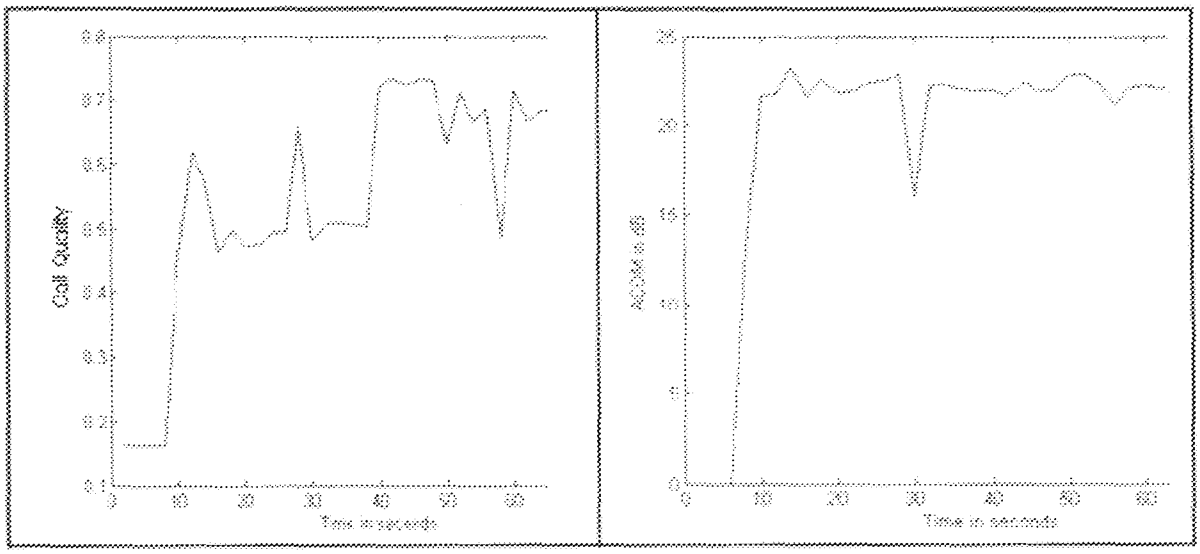

Figure 4.4: Comparing the estimated call quality with the estimated ACOM

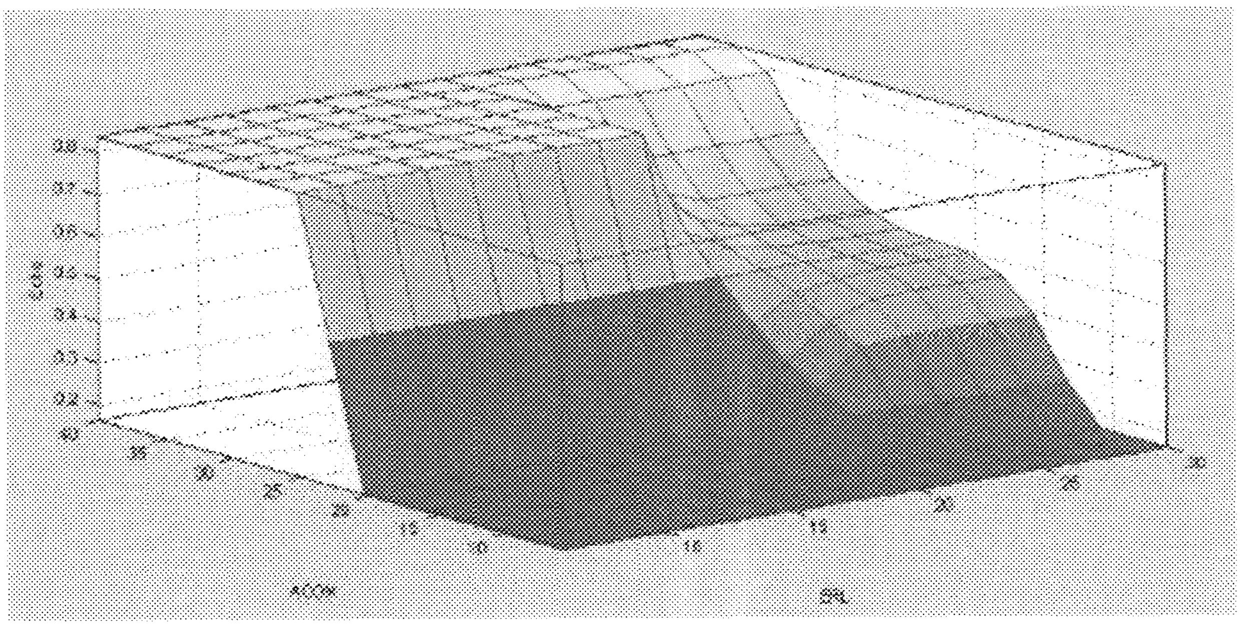

Figure 4.5 Mathlab surface view for Echo Content, ERL and ACOM

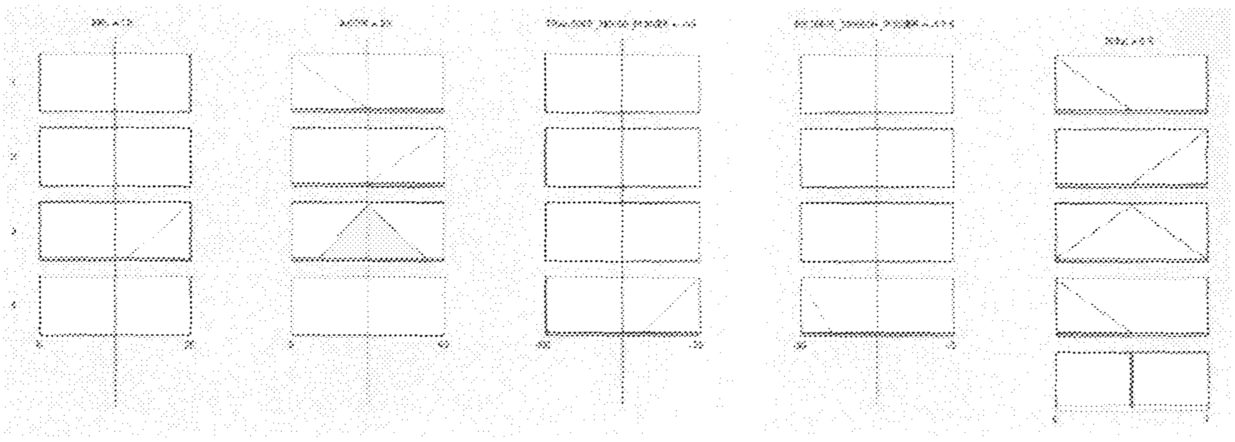

Figure 4.6 Mathlab Rule view for the Fuzzy Algorithm

CHAPTER FIVE

5.0 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 Discussion of Results

Analyzing Figure 4.2 in chapter 4, we note that the algorithm was really effective in estimating the echo component for these calls.

Figure 4.3 shows that echo was not effectively cancelled or there was some high noise or high difference level between the receive path signal and the send path signal. We see from this figure that the algorithm was able to declare several calls as having bad echo quality. However, some calls (like call number 7) obtained a much higher score than they really deserved.

Using the approach described above (finding the average of a completed call), we lose the detailed information about the call. For instance, if we get an average score for the voice quality in a given call we do not know if the call had a good quality for approximately half of the time and a bad quality in the remaining time or if the call quality was average during the whole call. This can be resolved by increasing the frequency of result samples taken for each call.

We conclude from these simulations that our proposed algorithm to evaluate the echo component of the voice quality in a VoIP call gave very good results for calls that had little or no echo. That is, the proposed algorithm was able to estimate with good accuracy that the final voice quality was not affected by echo in those calls (Figure4.2). We should note that due to our choices of defuzzification method and the output membership functions, even for calls with perfect quality the output wouldn't be 1.0. So the output shown in Figure 4.2 really reflects very good voice quality as we expected.

On the other hand, for calls where the voice quality was not so good due either to some high background noise, transmit/receive speech signal level disparity or presence of echo, the algorithm had a fair performance. As we can see in Figure 4.3, the algorithm was able to point out calls with voice quality problem but we still think that some calls got higher scores than expected considering our subjective analysis of the call.

5.2 Conclusions

In this work, we have been able to develop an algorithm that estimates echo content of a VoIP call in real-time. This algorithm is a building block (Figure 3.2) of an objective, passive, voice quality algorithm that can run in real-time and estimates the voice quality for live calls in a VoIP system or network. The use of fuzzy logic was motivated by its ability to give reasonably good result with low computational complexity; it is called soft computing method.

The proposed algorithm can run and give results in real-time in the embedded system that processes the VoIP calls, giving operators an almost instantaneous estimation of the quality of their network with respect to echo. Another advantage of the proposed algorithm is that unlike other popular methods, it does not require a reference signal, and this attribute makes it more suitable for real life calls.

On the other hand, the algorithm carries a tradeoff between precision and complexity. The main disadvantage of the proposed algorithm is that it is not as precise in its estimates as the PESQ, MOS or the E-model methods. Further work is needed to improve in the precision of the estimates while retaining the real-time feature.

5.3 Recommendations

In this section we discuss two main directions of future developments related to the algorithm proposed in this work.

i. The accuracy of the proposed algorithm could be improved.

ii. The proposed algorithm could be modified such that it can be used in different scenarios.

One possible way of improving the accuracy of a fuzzy inference system is to add more input fuzzy variables and more fuzzy rules by including other information like packet loss, delay and jitter. It would give a more precise estimate of the voice quality than taking only the echo component into consideration. Also, we should consider the possibilities of including fuzzy rules that incorporate knowledge of other speech parameters that can be affected by the IP network such as speech clarity and loudness.

Regarding different scenarios, there are few directions that can be followed in order to increase the range of scenarios that can take advantage of a similar algorithm. For instance, the main results of this work were developed for line echo present in a VoIP call. So, a modified algorithm could be used for similar purposes in a scenario where the echo present in the call has acoustic nature. It would be very useful for network operators to have a similar algorithm to evaluate the quality of an acoustic echo signal.

Another interesting direction is to adapt the algorithm for 802.11 or Wi-Fi networks, which are also becoming very popular. It seems that the chief challenge to it is that, relative to wired IP networks, packets are dropped at an excessive rate – in general 20% more packets are dropped. This can lead to distortion of the voice to the extent that the conversation is unintelligible and this must be taken into consideration when adapting our proposed algorithm to the wireless VoIP environment.

REFERENCES

Adaptive Digital Technologies. (2010). *The Echo "Phenomenon", Causes and Solutions.* Retrieved from http://www.adaptivedigital.com/product/echo_cancel/echo_explain.htm

Choudhry, U.I., JongWon K. and Hong K. K., (2006). *A Highly Adaptive Acoustic Echo Cancellation Solution for VoIP Conferencing System. Computer Systems and Applications.* IEEE International Conference, pp 433 – 436.

Cisco Systems. (2009). *Understanding Delay in Packet Voice Networks.* Retrieved from http://www.cisco.com/en/US/tech/tk652/tk698/technologies_white_paper09186a00800a8993.shml

David McMahon, (2007). *MATLAB Demystified*, The McGraw-Hill Companies

Ditech Networks. (2007). *Limitations of PESQ for Measuring Voice Quality in Mobile and VoIP Networks.*

Ditech Networks. (2010). *Echo Basics Tutorial including echo cancellers and echo's effect. on QoS* Retrieved from http://www.ditechnetworks.com/learningCenter/echoBasics.html

Emily, T. S, (1993). *Why the Japanese are going in for this 'Fuzzy Logic', Business Week*, pp. 39.

Falk, T.H. and Wai-Y. C. (2008). *Hybrid Signal-and-Link-Parametric Speech Quality Measurement for VoIP Communications, Audio, Speech, and Language Processing*, vol.16, no.8, pp.1579-1589.

How stuff works, (2010). How VoIP works, Retrieved from http://communication.howstuffworks.com/ip-telephony.htm

Farouk Y.M., Alkadhi, Zheng L., Min Y., Qiuhong W., and Jufeng D., (2010). *A Silence Duration Based Uplink Scheduling Algorithm for Multiple VoIP Users in*

M-WiMAX, Future Networks, ICFN 10. Second International Conference, pp 291 – 296.

Howon L. and Dong-Ho C. (2010). *Capacity Improvement and Analysis of VoIP Service in a Cognitive Radio System*, Vehicular Technology, vol. 59 pp 1646 – 1651.

Ibrahim, A. (2004). *Fuzzy Logic for Embedded Systems Applications*, Newnes. Imperial College London. Fuzzy logic and it uses. Retrieved from http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol2/jp6/article2.html

ITU-T G.165, (n.d.). *ITU-T Recommendation G.165 for Echo Cancellers, International Telecommunication Union.* Retrieved from, http://www.itu.int/en/ITU-T/publications/Pages/recs.aspx

ITU-T G.168, (n.d.). *ITU-T Recommendation G.168 for Digital Network Echo Cancellers, International Telecommunication Union.* Retrieved from http://www.itu.int/en/ITU-T/publications/Pages/recs.aspx

ITU-T P.800, (n.d.). *ITU-T Recommendation P.800, Methods for subjective determination of transmission quality, International Telecommunication Union.* Retrieved from http://www.itu.int/en/ITU-T/publications/Pages/recs.aspx

ITU-T P.862, (n.d.). *ITU-T Recommendation P.862, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality. International Telecommunication Union.* Retrieved from http://www.itu.int/en/ITU-T/publications/Pages/recs.aspx

Jin-suk K., Young-do K., In-ho N., Yeon-sung C. and Jang-hyung K., (2001). *A study of subjective speech quality measurement over VoIP network, Info-tech and Info-net.* Proceedings ICII 2001 - Beijing. vol. 5, pp. 311-316.

Jinho Hwang, Nakpo Kim, Yungha Ji, Jongseog Koh, (2010). *A mobile VoIP architecture over LTE & WLAN Networks,* Advanced Communication Technology (ICACT), 2010 The 12th International Conference, vol 1 pp 729 – 732

Justin Y. and Chien-C. S., (2010). *Managing End-to-End Delay for VoIP Calls in Multi-Hop Wireless Mesh Networks.* IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings

Harjit P. S., Sarabjeet S., Jasvir S., (2010). *Computer Modelling & Performance Analysis of VoIP under Different Strategic Conditions, Computer Engineering and Applications (ICCEA)*, Second International Conference, vol.1, pp 611 – 615.

Lakaniemi, A., Rosti, J. and Raisanen, V. I., (2001). *Subjective VoIP speech quality evaluation based on network measurements, Communications*. Proceedings of the ICC 2001. IEEE International Conference, vol.3, pp.748-752.

Lingfen S. and Emmanuel C. I. (2006). Perceived Speech Quality Prediction for Voice over IP-based Networks, Department of Communication and Electronic Engineering, University of Plymouth, Plymouth PL4 8AA, U.K.

Manjunath, T. (2009). *Limitations of perceptual evaluation of speech quality on VoIP systems*, Proceedings of the Broadband Multimedia Systems and Broadcasting, BMSB '09. IEEE International Symposium on communications, vol. 13-15 pp.1-6.

Manoj B., Jonathan D., Satish K., Sudipto M., James P., (2007). VoIP An In-Depth Analysis, Chapter 3 .Voice over IP Fundamentals, 2nd Edition, Cisco Press.

Miroslaw N. and Mark D. (2007). *Assessing the quality of VoIP transmission affected by playout buffer scheme, Communications Network Research Institute*. Dublin Institute of Technology.

Namboodiri, V. and Lixin G., (2010). *Energy-Efficient VoIP over Wireless LANs*, Mobile Computing, vol. 9, pp 566 – 581.

Ngamwongwattana, B. and Thompson, R. (2010). *Sync & Sense: VoIP Measurement Methodology for Assessing One-Way Delay Without Clock Synchronization*. Instrumentation and Measurement vol. 59, pp 1318 – 1326.

Octasic, M. (2010). Voice Quality Enhancement and Echo Cancellation, retrieved from http://www.octasic.com/en/tech/echo_cancel.php

Paglierani, P. and Petri, D. (2009) *Uncertainty Evaluation of Objective Speech Quality Measurement in VoIP Systems*, Instrumentation and Measurement, vol.58, no.1, pp.46-51

Paglierani, P. and Petri, D. (2007) *Uncertainty evaluation of speech quality measurement in VoIP systems*, Advanced Methods for Uncertainty Estimation in Measurement, vol. 22, pp.104-108

Periakarruppan, G.; Low, A.L.Y.; Azhar, H.; Rashid, A., (2006) *Packet Based Echo Cancellation for Voice Over Internet Protocol Simulated with Variable Amount of Network Delay Time*, TENCON 2006. 2006 IEEE Region 10 Conference, pp 1 – 4

Ren, Jiuchun; Zhang, ChongMing; Huang, WeiChao; Mao, Dilin, (2010) *Enhancement to E-Model on standard deviation of packet delay*, Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference, vol. 23-25, pp.256-259

Ross, T. (2004) Fuzzy Logic with Engineering Applications, Wiley, 2004

Songun Na and Seungwha Yoo, (2002) *Allowable Propagation Delay for VoIP Calls of Acceptable Quality*, Advanced Internet Services and Applications Lecture Notes in Computer Science, vol. 24, pp 469-480

Sung-Min Oh and Jae-Hyun Kim, (2010) *User-perceived QoS performance enhancement for VoIP services in IEEE 802.16 systems*, Pervasive Computing and Communications Workshops(PERCOM Workshops), 2010 8th IEEE International Conference

Vikram Chandrasekaran, Ram Dantu, Neeraj K Gupta Xiaohui Yang, Duminda Wijesekera, (2010) *Efficiency of social connection-based routing in P2P VoIP networks*, Communication Systems and Networks (COMSNETS), 2010, pp 1 – 6.

Wikipedia, (2010). *Fuzzy Logic*. Retrieved from http://en.wikipedia.org/wiki/Fuzzy_logic

Wikipedia, (2010). *Echo cancellation.* Retrieved from http://en.wikipedia.org/wiki/Echo_cancellation

Wikipedia, (2011). *Voice over IP. Retrieved from* http://en.wikipedia.org/wiki/Voice_over_IP

Zhang Xuejing, Li J. and Hao X. (2010). *Scheduling and Control of VoIP Video Telephone under DSP/BIOS.* Education Technology and Computer Science (ETCS), Second International Workshop, vol.2, pp 726 – 729

Zizhi Q., Lingfen S., Ifeachor, E., (2008). *Case study of PESQ performance in live wireless mobile VoIP environment,* Personal, Indoor and Mobile Radio Communications. PIMRC 2008. IEEE 19th International Symposium, vol. 15, pp.1-6

APPENDICES

Appendix A

Matlab code

```
/*******************************************/

[System]
Name='echo compoent'
Type='mamdani'
Version=2.0
NumInputs=4
NumOutputs=1
NumRules=4
AndMethod='min'
OrMethod='max'
ImpMethod='prod'
AggMethod='max'
DefuzzMethod='centroid'

[Input1]
Name='ERL'
Range=[6 30]
NumMFs=1
MF1='Good':'trimf',[20 30 30]

[Input2]
Name='ACOM'
Range=[6 40]
NumMFs=3
MF1='Bad':'trimf',[6 6 23]
MF2='Moderate':'trimf',[12 23 36]
MF3='Good':'trimf',[23 40 40]

[Input3]
Name='TRANSMIT_NOISE_POWER'
Range=[-60 -36]
NumMFs=1
MF1='Bad':'trimf',[-45 -36 -36]
```

```
[Input4]
Name='RECEIVE_SPEECH_POWER'
Range=[-30 -5]
NumMFs=2
MF1='Bad1':'trimf',[-30 -30 -25]
MF2='Bad2':'trimf',[-15 -5 -5]

[Output1]
Name='Echo'
Range=[0 1]
NumMFs=3
MF1='Bad':'trimf',[0 0 0.5]
MF2='Moderate':'trimf',[0 0.5 1]
MF3='Good':'trimf',[0.5 1 1]

[Rules]
0 1 0 0, 1 (1) : 1
0 3 0 0, 3 (1) : 1
1 2 0 0, 2 (1) : 1
0 0 1 1, 1 (1) : 1
/********************************************/
```