An Impedance Glottograph

A dissertation
submitted in partial fulfillment of the requirements
for the degree of

Master of Technology

By

Anil Luthra
(Roll No. 02307413)

under the supervision of

Prof P. C. Pandey

Department of Electrical Engineering
Indian Institute of Technology, Bombay

July 2004

INDIAN INSTITUE OF TECHNOLOGY, BOMBAY

M. Tech Dissertation Approval

Dissertation entitled, "An Impedance Glottograph", submitted by Anil Luthra (Roll No. 02307413) is approved for the award of degree of Master of Technology in Electrical Engineering with specialization in Electronic Systems

Supervisor : _____ (Prof. P. C. Pandey)

Internal examiner : _____ (Prof. R. Lal)

External examiner : _____ (Prof. L. R. Subramanyan)

Chairman : _____ (Prof. G. G. Ray)

Date : July 12, 2004

Anil Luthra / Prof P C Pandey (Supervisor), "An Impedance Glottograph", *M.Tech. dissertation*, Department of Electrical Engineering, Indian Institute of Technology, Bombay, July 2004.

ABSTRACT

Impedance glottography measures the time variation of the degree of contact between the vibrating vocal folds during voice production. It is useful for estimation of voice pitch, diagnosis of voice disorders and as a speech training aid for the hearing impaired. The objective of this project is to develop an impedance glottograph instrument. It entails hardware and software development for signal acquisition, display and analysis. An instrument has been developed to pass a high frequency (400 kHz), low intensity (~ 1 mA) current through the central discs of a pair of electrodes held in contact with the skin across the thyroid cartilage. The impedance variations caused by varying contact area between the vocal folds results in amplitude modulated voltage waveform. This waveform is demodulated to get impedance glottogram. A Windows based application software has been developed to acquire the impedance glottogram along with the speech waveform through the stereo line input of the PC sound card, and plot the global pitch and pitch histograms.

ACKNOWLEDGEMENTS

I would like to express my deep sense of gratitude to my supervisor, Prof. P. C. Pandey, for his invaluable help and guidance which effectively contributed in successful completion of the project. He was instrumental in providing technical and moral support. The regular discussions we had were a source of encouragement and guidance.

I would like to thank Prof. R. Lal for the invaluable suggestions offered during the various stages of the project which enabled me to improve the work.

I would also like to thank all lab mates and friends for all the assistance provided.

IIT Bombay Anil Luthra
July 2004

CONTENTS

LIST OF SYMBOLS

Symbols	Explanation
F_o	Fundamental frequency or pitch
Fx	Pitch plot
Lx	Laryngogram
Z_S	Source impedance
Z_L	Load impedance
Z_g	Fixed impedance across thyroid cartilage
z_{gv}	Variable impedance across thyroid cartilage due to movement of vocal folds
Z_{C1}	Impedance of tissue between CE1 and thyroid cartilage
Z_{R1}	Impedance of tissue between RE1 and thyroid cartilage
Z_{C2}	Impedance of tissue between CE2 and thyroid cartilage
Z_{R2}	Impedance of tissue between RE2 and thyroid cartilage
Z_{CR1}	Impedance across skin between CE1 and RE1
Z_{CR2}	Impedance across skin between CE2 and RE2
$x(n)$	IGG waveform
$a(n)$	IGG waveform average
$p(n)$	IGG waveform peak
$v(n)$	IGG waveform valley
$y(n)$	Hysteresis comparator output
$x_H(n)$	Upper threshold for hysteresis comparator
$x_L(n)$	Lower threshold for hysteresis comparator

LIST OF ABBREVIATIONS

Abbreviation	Explanation
IGG	Impedance glottogram
EGG	Electro glottogram
PC	Personal computer
PCI	Peripheral components interconnect
PCB	Printed circuit board
PTH	Plated through holes
LED	Light emitting diode
CMRR	Common mode rejection ratio
CE1	Centre electrode1
RE1	Ring electrode1
CE2	Centre electrode2
RE2	Ring electrode2

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Overview

Lungs, larynx, and vocal tract are the main organs related with generation of speech. The lungs are the source of airflow. The vocal tract is an accoustic enclosure and acts as acoustic filter shaping the spectrum of the generated sound. The mechanism of generation of speech is known as phonation. The source of most speech occurs in the larynx. There are two folds of the muscular bundle known as vocal folds inside the larynx. These vocal folds obstruct the airflow from the lungs and produce audible vibrations that provide excitation to the vocal tract for voiced speech. The vibrations consists of three phases namely contact phase, separation phase and open phase [1][2][3][4].

The larynx due to its position in the throat, is difficult to study. A wide spectrum of diseases impair laryngeal function, and result in changes in speech [21]. Measurement of vocal fold movement has, thus, assumed significance. Vocal fold movement can be measured by four methods namely laryngoscopy, transglottal illumination, air flow measurement and impedance glottography [2].

Laryngoscopy involves illumination and simultaneous viewing of vocal folds. It may be direct or indirect, that is, viewing direct by a scope or indirectly by interposing a mirror. In either case it requires inserting a scope or a mirror via a nostril. Thus, normal phonation is not possible during laryngoscopy. In transglottal illumination, source of light is introduced via one nostril and the naso-pharynx and the light obtained on other side of vocal folds is measured by a photocell placed externally on the exterior wall of the throat. Again normal phonation is not possible due to introduction of light source. Air flow measurement can be used for measurement of vocal fold vibration as air flow from lungs is associated with vocal fold vibration and results in a pulsating component. This pulsating component can be used for deriving laryngeal function by an inverse filtering technique applied to pressure waveform of speech itself. Impedance glottography involves measurement of impedance variation due to vocal fold movement. It provides a

practical and non-invasive means of measuring vocal fold movement during normal phonation [2][3][5].

For impedance glottography, two electrodes are held in contact with the skin across the thyroid cartilage. A time varying current of high frequency is injected through the two electrodes. When the vocal folds vibrate, the impedance between the vocal folds changes. Since the current is constant, the voltage across the two electrodes changes in accordance with the impedance between the vocal folds and we get an amplitude modulated voltage waveform, across the electrodes, which is varying in direct proportion to the variation in the impedance due to the movement of the vocal folds. The device used is called impedance glottograph or electroglottograph (EGG) or laryngograph [4][6].

The impedance between the vocal folds is a function of tissue path length. When the vocal folds are open, the tissue path length is maximum and hence the impedance is maximum. When the vocal folds are closed, the tissue path length is minimum and hence the impedance is minimum. Impedance glottograph measures this impedance variation.

Laryngoscopy, transglottal illumination and air flow measurement give information about separation of the vocal folds and very little indication of the nature of their contact. Impedance glottography gives information about the contact phase of vocal folds [7]. However, it does not provide information during the separation phase.

The method was first reported by Fabre (1957) and other significant contributions have been made by Fourcin (1971) and Frokjaer-Jensen (1968) [8]. Commercially available devices have been produced by Laryngograph Ltd. [9], F-J Electronics [10] and Glottal Enterprises [11].

1.2 Objectives

This project is a continuation of development work done earlier at IIT Bombay as part of student projects by Bhagwat (1990) [12], Sriram (1991) [13], Thajudin (1994) [14], Mahajan (1995) [15], Chitnis (1998) [16] and Patil (2000) [17]. The circuit developed earlier had a bandwidth from 50 Hz to 2.4 kHz and also the sensitivity of the circuit was low. The aim of this project is to

(a) Increase the sensitivity of the instrument keeping the noise level low, and increase the bandwidth to 5 kHz, with a good phase response for more faithful acquisition of the waveform.

(b) Study the electrode configurations thoroughly to select the most appropriate one.

(c) To develop a Windows based software for processing and analyzing the output waveform.

1.3 Outline of the dissertation

The second chapter describes the speech production mechanism, basics of impedance glottography and its applications in laryngeal function assessment and in speech analysis, development of the instrumentation at IIT Bombay and the various commercially available and patented instruments. Chapter 3 describes the hardware modules: Wien bridge oscillator, voltage-to-current converter module, high impedance indicator module, instrumentation amplifier module, demodulator module, and the filter module. Chapter 4 discusses the Windows based signal acquisition and analysis package which has been developed. Chapter 5 describes the design, fabrication and assembly of the unit and gives results for various tests carried out on the circuit and analysis of recordings from some subjects. Chapter 6 summarizes the work alongwith some suggestions for further work. Appendices provide supplementary information: instrumentation specification, component list, PCB layouts, panel and cabinet design, operation manual for the software and simulator design.

Chapter 2

IMPEDANCE GLOTTOGRAPHY

2.1 Speech production

Speech signal in the form of pressure waves travel from the speaker to the listener. The signal comprises of variation in pressure as a function of time. Speech is usually produced as air is exhaled. For each sound, there is a definite position for each of the vocal tract articulators.

The speech organs can be divided into three main groups: lungs, larynx and vocal tract as shown in Fig 2.1. Here vocal tract refers to the portion above the larynx and includes oral cavity, nasal cavity, tongue and pharynx [1]. Speech comprises of two mechanical functions namely phonation and articulation. Phonation is carried out in the larynx and articulation in the vocal tract [18]. The larynx acts as a vibrator. The vibrating element is the vocal folds. The basic structure of the larynx is as shown in Fig 2.2.

Lungs are the source of airflow. The air from the lungs passes the larynx and the vocal tract. The air flow is basically achieved by the contraction and relaxation of the diaphragm. The contraction of the diaphragm expands the lungs and relaxation of the diaphragm allows the elastic recoil of the lungs and ribs to expel air. Air flow patterns are, however, different for speaking and for normal breathing. Sound amplitude is individually related to air flow rate.

During normal breathing, air passes the larynx and vocal tract unobstructed and creates no or little audible sound. Sound occurs when the path is constricted or totally occluded interrupting the air flow to create turbulent noise or pulses of air. The source of most speech occurs in the larynx when vocal folds partially or completely obstruct air flow from the lungs [1]. The vocal folds are in the process set into vibration. The cause of vibration of the vocal folds can be explained when the vocal folds are adducted and air is expired, pressure of air from below first pushes the vocal folds apart and air flows rapidly. The rapid flow of air then immediately creates a partial vacuum between the vocal folds because of the Bernoulli principle, which pulls them once again towards each other. This stops the air flow, building pressure again which again opens the vocal folds

and, thus, a vibratory motion is set in [18]. Fig 2.3. gives the position of vocal folds in a horizontal section and Fig 2.4. gives the vocal chord vibration sequence in a vertical section.

The rate of vibration is called the fundamental frequency or the pitch (F_o). The fundamental period between successive vocal chord closures, $T_o = 1/ F_o$, has an average value that varies with the size of the speakers vocal folds. Typically the vocal folds are 15 mm long in men, 13 mm in women [18]. Thus, the pitch of women is higher than that of men. Typical pitch range for male and female speakers are 100-250 Hz and 200-450 Hz respectively. Children have a higher pitch [1].

The structures in the vocal tract that move in production of different sounds are called articulators. The three major organs for articulation are the lips, the tongue and the soft palate. Each articulator moves in response to neural command to produce speech.

2.2 Basics of impedance glottography

Impedance glottography is an electrical impedance measurement technique for measuring the electrical impedance variations of the larynx using a pair of plate electrodes held in contact with the skin on both sides of the thyroid cartilage. The electrodes are usually mounted on a flexible band whose length may be adjusted to hold the electrodes in a steady position and to still allow the subject to comfortably speak and breathe naturally. Sometimes the electrodes are mounted on a small holder which is pressed against the throat by hand. The output is amplitude modulated by the varying tissue impedance due to the vibrating vocal folds. This is then demodulated by a detector circuit. The percentage of amplitude modulation in this signal is directly proportional to the percentage change in the tissue impedance. The output thus obtained is a function of the lateral area of contact between the vocal folds [19].

Impedance glottography has been also known as electroglottography and laryngography. The instruments are known as electroglottograph and laryngograph and the output voltage corresponding to impedance variation has been known as electroglottogram (EGG) and laryngogram (Lx). In this dissertation, we will use the term impedance glottogram (IGG) for the output waveform.

2.3 Impedance glottogram

Impedance glottogram gives information of the contact phase, open phase and separation phase of the vocal folds which is very useful for laryngeal function assessment. Fig. 2.5 shows speech and IGG waveform for sustained phonation of vowel /*a*/, as given by Marasek [20]. Fig. 2.6 shows a model of impedance glottogram as given by Marasek [20].

When the vocal folds are open and it is ensured that there is no lateral contact between the vocal folds, the impedance is maximum and peak glottal flow occurs corresponding to segment (e) in Fig. 2.6. The waveform in this segment is flat, with small fluctuations. Then, the upper margins of the vocal folds make initial contact corresponding to segment (f) in Fig. 2.6.

In the next phase of the movement, denoted as (a), the lower margins come into contact and the vocal folds as a whole continue to close zipper-like. If the vocal folds close very rapidly and along their whole length, the phases (f) and (a) become indistinguishable and consequently the slope of the closure phase ((f)+(a)) becomes steep. The presence of this knee is typical for low to normal voice intensities and the slope of segment (f) is more gradual than the slope of (a). Glottal closure is achieved during phase (a). During the next phase, indicated as (b) in Fig. 2.6, the vocal folds remain in contact and the airflow is blocked.

The opening and the open phase are described analogously. In the process of vocal fold separation, the contact between the folds starts to diminish and subsequently the lower margins of the vocal folds begin to separate, initializing the opening. Lower margin separation proceeds gradually during phase (c) (Fig. 2.6). Then the upper margins also begin to separate, resulting in an acceleration in the growth of impedance (phase (d)) until the full opening is reached.

The time derivative of the impedance glottogram is usually used in the determination of the periodicity of the signal. It can also be used to identify the distinguishable changes in the slopes during the phases of increasing and decreasing impedance which correspond to those of the simplified model of the IGG as shown in Fig. 2.7.

The positive peak of the derivative serves as an indication for the instant of the glottal closure. This is a robust and reliable marker of the pitch period for various voice qualities and intensities. The negative peak of the IGG derivative is regarded as an

indicator of glottal opening. During the opening phase the glottal area increases monotonically until it reaches its maximum. The movements of the vocal folds that are reflected in the IGG have two distinct phases. First, the IGG decreases monotonically, reflecting the decreasing lateral contact between the vocal folds. During this interval the IGG waveform is convex. Then, as the upper margin separates the waveform of the IGG changes to concave.

IGG, thus, reflects the contact area between the vocal folds. Assuming that the depth of contact does not change during vocal fold vibration, the lateral contact area depends on the length of contact area along the upper margins of the vocal folds. There is, thus, a strong correlation between the IGG and the length of the vocal folds contact [20][21][22].

2.4 Applications of IGG

2.4.1 Laryngeal function assessment.

Laryngeal function assessment can be done by impedance glottography as impedance glottogram depicts the variation of impedance due to movement of vocal folds, which depends on pathological condition of vocal folds [2][23][24]. This can be established by observing the impedance glottogram of a normal voice, whispery voice, creaky voice, breathy voice, and tense voice as shown in Fig. 2.8.

For a normal voice, there are three distinct parts of the waveform. First, a relatively sharp rise denoting the rapid closing of vocal folds. Second, a gradual fall denoting the parting of vocal folds and third, a flatter base denoting the interval during which the vocal folds are open. The signal duty ratio is about 50%. For a whispery voice, the increase and decrease in impedance is faster. The duty ratio is much lower than in normal voice. For a creaky voice, the shape of waveform is triangular with rounded corners. Rise in impedance is very fast and fall is gradual. So it is difficult to define the signal duty cycle as fall is gradual and no knee can be identified. In breathy voice, peak amplitude is lower, which may be due to poor contact between vocal folds or their incomplete closure. The maximum contact phase is extremely short. Tense voice is produced with a strongly increased muscular tension in the larynx. The shape of waveform is almost rounded. Peak to peak amplitude is lower. Increase in impedance is

gradual. Maximum contact phase is relatively long. Also the decrease in impedance is gradual. The duty ratio is comparable to the normal voice.

2.4.2 Use of IGG in speech analysis

(a) *Spectral analysis*. Spectral estimation techniques employ short-time fourier transform with windows such as Hanning, Hamming, Kaiser etc. The analysis window which is arbitrarily placed, causes errors when it includes glottal open phase regions and differing amounts of excitation epochs. The problems are significantly reduced when the analysis window is restricted to a maximum of one pitch period in length and restricted to encompass only the closed glottal region. This information is available from time aligned IGG signal, resulting in proper frame positioning for speech analysis for vocal tract related parameters [21].

(b) *Voiced-unvoiced classification*. With IGG signal, classification of speech into voiced, unvoiced, mixed and silent regions is relatively simple and accurate. For voiced-unvoiced classification, thresholding of IGG signal is sufficient as for unvoiced regions IGG is almost zero and for voiced regions IGG is periodic and non zero. For unvoiced-silent or voiced-mixed decisions, the algorithms are hybrid speech IGG in nature [2][21].

(c) *Estimation of fundamental frequency*. Fundamental frequency or pitch can be estimated from speech signal by zero crossing rate, auto correlation method etc but these are not accurate. Using the IGG, pitch can be calculated with simplicity and with high accuracy. Pitch value is calculated based on either zero crossings or the distance between minima in the differentiated IGG [21].

(d) *Synthesis applications*. IGG aided analysis methods play a key role in synthesis of high quality speech. This is because naturalness and intelligibility of synthesized speech depends upon accuracy of vocal tract modeling, voiced-unvoiced classification and pitch detection. And these parameters can be accurately obtained using IGG [21].

(e) *Recognition applications*. The success of any speech recognition system depends upon the accuracy of speech analysis. IGG aided speech analysis can give the various parameters of speech accurately and, thus, IGG is useful for recognition applications [21].

2.5 Available instruments

Impedance glottograph, also called electroglottograph was first developed by Fabre in 1957 and significant contributions have been made by Frokjaer Jensen in 1968 and Fourcin in 1971[8]. Some of the commercially available instruments are

(a) Laryngograph, developed by Fourcin et al (University college, London) around 1979, marketed by Laryngograph Limited of UK [9]. The design has a US patent [25].

(b) EG2-PC, developed by Rothenberg et al (Syracuse University, US) around 1990, marketed by Glottal Enterprises of US [11]. The design has a US patent [26].

(c) EG 90, developed by Frokjaer Jensen, is marketed by F J Electronics of Denmark [10].

2.6 Impedance detection methods

Glottal impedance variation due to vocal fold vibration can be detected using one of the several methods. Fig 2.9 shows Childer's scheme [21]. A pair of electrodes is held in contact with the skin on both sides of the thyroid cartilage. An RF voltage source is applied using a transformer. The high frequency RF signal is given to a pair of electrodes. The model of the same can be given as in Fig 2.10. The impedance between the electrodes can be modeled as a fixed impedance, Z_g in series with a time varying impedance z_{gv}. Z_s is source impedance and Z_L is load impedance due to input impedance of the detecting circuit. Thus,

$$V_O = V_S \frac{Z_L}{Z_L + Z_S + Z_g + z_{gv}}$$

$$\approx V_S \left\{ \frac{Z_L}{Z_L + Z_S + Z_g} \right\} \left\{ 1 - \frac{z_{gv}}{Z_L + Z_S + Z_g} \right\} \qquad (2.1)$$

Thus, when the vocal folds vibrate, z_{gv} varies and the output gets amplitude modulated. As z_{gv} is small compared to $(Z_L + Z_s + Z_g)$, the resulting modulation index will be small.

The impedance detection scheme used by Rothenberg is as shown in Fig 2.11 [26]. A pair of electrodes is held in contact with the skin on both sides of the

thyroid cartilage. An RF voltage source is applied using a transformer. The high frequency RF voltage is applied across the electrodes, in parallel with the two transformers. The scheme can be modeled as shown in Fig 2.12. The impedance between the electrodes can be modeled as a fixed impedance, Z_g in series with a time varying impedance z_{gv}. Z_s is source impedance and Z_L is load impedance due to input impedance of the detecting circuit. Thus,

$$V_O = Vs\frac{Z_L \parallel (Z_g + z_{gv})}{Z_S + Z_L \parallel (Z_g + z_{gv})}$$

$$\approx V_S\left\{\frac{Z_L}{Z_L + Z_S + Z_L Z_S / Z_g}\right\}\left\{1+\frac{z_{gv}}{Z_g}\right\} \qquad (2.2)$$

Thus, when the vocal folds vibrate, z_{gv} varies and the output gets amplitude modulated.

The impedance detection scheme used by Fourcin is as shown in Fig 2.13 [25]. A pair of electrodes is held in contact with the skin on both sides of the thyroid cartilage. Each electrode has a central disc and a ring. An RF voltage is applied using a transformer across the central disc and ring on one side. Other electrode is used to sense the output voltage between its central disc and ring on the other side. So electrode on one side acts as a transmitter and the other acts as a receiver. The scheme can be modeled as shown in Fig 2.14. The impedance across the thyroid cartilage can be modeled as a fixed impedance, Z_g in series with a time varying impedance z_{gv}. Z_s is source impedance and Z_L is load impedance due to input impedance of the detecting circuit, Z_{C1} is the impedance of the tissue between CE1 and thyroid cartilage, Z_{R1} is the impedance of the tissue between RE1 and thyroid cartilage, Z_{C2} is the impedance of the tissue between CE2 and thyroid cartilage, Z_{R2} is the impedance of the tissue between RE2 and thyroid cartilage, Z_{CR1} is the impedance across the skin between CE1 and RE1, Z_{CR2} is the impedance across the skin between CE2 and RE2. The equivalent circuit can be modeled to the one in Fig 2.15.

$$V_O = V_S'\left\{\frac{Z_g + z_{gv}}{Z_s' + Z_{C1} + Z_{R1} + Z_g + z_{gv}}\right\}\left\{\frac{Z_L'}{Z_L' + Z_{C2} + Z_{R2}}\right\}$$

$$\approx V_S'\left\{\frac{Z_L'}{Z_L' + Z_{C2} + Z_{R2}}\right\}\left\{\frac{Z_g}{Z_s' + Z_{C1} + Z_{R1} + Z_g}\right\}\left\{1 + z_{gv}\left(\frac{Z_S' + Z_{C1} + Z_{R1}}{Z_g(Z_S' + Z_{C1} + Z_{R1} + Z_g)}\right)\right\} \qquad (2.3)$$

Thus, when the vocal folds vibrate, z_{gv} varies and the output gets amplitude modulated.

The impedance detection scheme used in the present project is as shown in Fig 2.16. A pair of electrodes is held in contact with the skin on both sides of the thyroid cartilage. A high frequency RF voltage is converted to current and it flows through the two electrodes. The equivalent circuit of the scheme being used for impedance detection in the present project is as shown in Fig 2.17.The impedance between the electrodes can be modeled as a fixed impedance, Z_g in series with a time varying impedance z_{gv}. Z_s is source impedance. When the vocal folds vibrate, z_{gv} varies and the output gets amplitude modulated. Thus,

$$V_O = \{(Z_g + z_{gv}) \| Z_S\} I_S$$
$$\approx (Z_S \| Z_g)\left\{1 + \frac{z_{gv}}{Z_g}\right\} I_S \qquad (2.4)$$

In this instrument being designed, a transformer less version is used. The frequency being used is 400 kHz.

Thus, we see that out of these methods, the modulation of the output voltage is most directly linked to glottal impedance variation in case of excitation by current source. This method has been selected for instrument to be developed in this project.

2.7 Instrument development

Development work for impedance glottograph instrument has been carried out at IIT Bombay as part of student projects by Bhagwat (1990) [12], Sriram (1991) [13], Thajudin (1994) [14], Mahajan (1995) [15], Chitnis (1998) [16] and Patil (2000) [17]. The circuits developed by Bhagwat (1990) to Mahajan (1995) had noise problems and impedance change of less than 1 ohm could not be detected properly. Chitnis in 1998 was successful in developing a glottal impedance sensor, which could detect impedance change of less than 1 ohm. He also developed a stand alone data acquisition and display unit for acquiring impedance variation waveform and analyzing it. Patil in 2000 made design changes to the filter stage of the sensor. The filter earlier used by Chitnis was a cascade of first order high-pass filter with 3 dB cut off frequency of 10 Hz and two first order low-pass filters with cut off frequency of 10 kHz. The frequency response deteriorated due to cascade of two low-pass filters. Patil replaced the filters with second

order Butterworth high pass (3 dB cut off frequency of 50 Hz) and low pass (3 dB cut off frequency of 2.4 kHz) filter followed by a first order bandpass filter (3 dB cut off high pass of 10 Hz and 3 dB cut off low pass of 22 kHz). The block diagram of impedance glottograph developed by Patil is given in Fig. 2.18.

The filter stage has been modified in the present design. Also an instrument amplifier has been used to remove common mode noise pickup. The output obtained from electrodes is passed through the instrumentation amplifier. The output of the instrumentation amplifier is then demodulated and passed to the filter stage. Filter stage comprises a notch filter to reduce 50 Hz pick up, followed by a passive first order high pass filter with a cut off frequency of 2 Hz to remove the dc offset, followed by a fifth order elliptic low pass filter with a cut off frequency of 5 kHz with a roll off of 40 dB in an octave. The band width has thus been increased to 2 Hz −5 kHz as compared to the previous design of 50 Hz - 2.4 kHz. The high impedance indicator module was not working in the previous design. It has been redesigned. Resistance values have been chosen to take into account the bias current effects. Op amps for various circuit blocks have been selected after examining and testing various options.

The signal acquisition and analysis software was developed by Lehana [4] in C language on DOS platform. The signal acquisition and analysis software has been completely redesigned, based on Windows platform, with added features.

The hardware of the system developed is described in Chapter 3, while the software is described in Chapter 4.

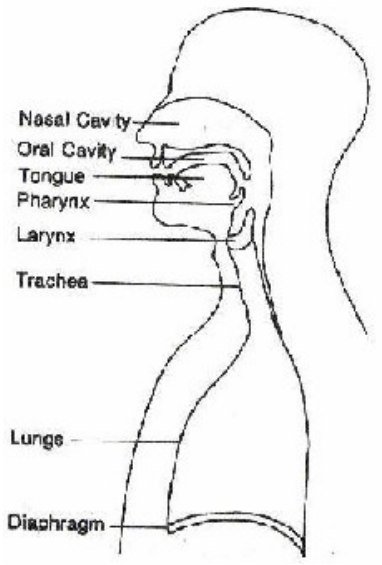


Fig 2.1. Speech organs. Adapted from [1].

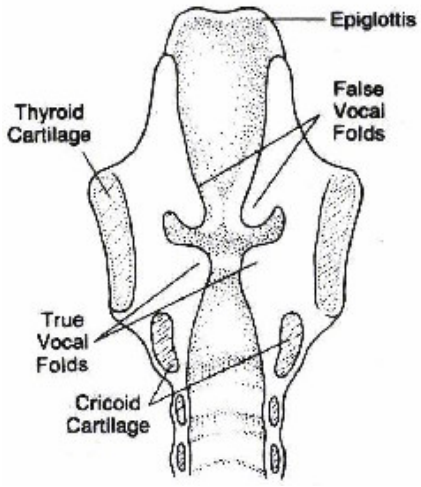


Fig 2.2. Vertical cross section of larynx. Adapted from [1].

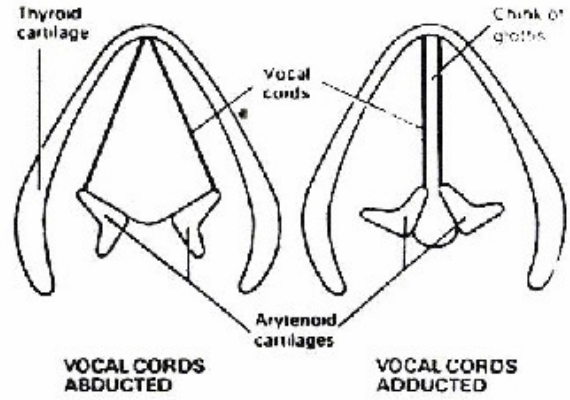


Fig 2.3. Horizontal cross section of larynx. Adapted from [1].

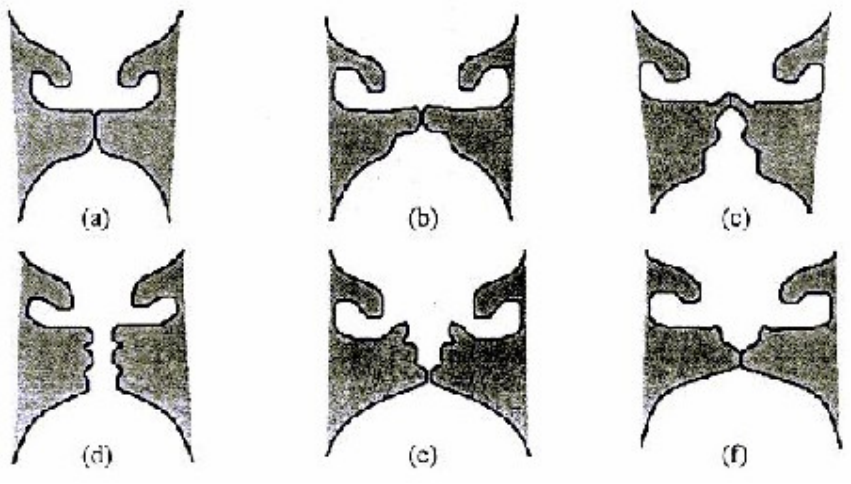


Fig 2.4. Vocal chord vibration sequence in a vertical section. Adapted from [1].

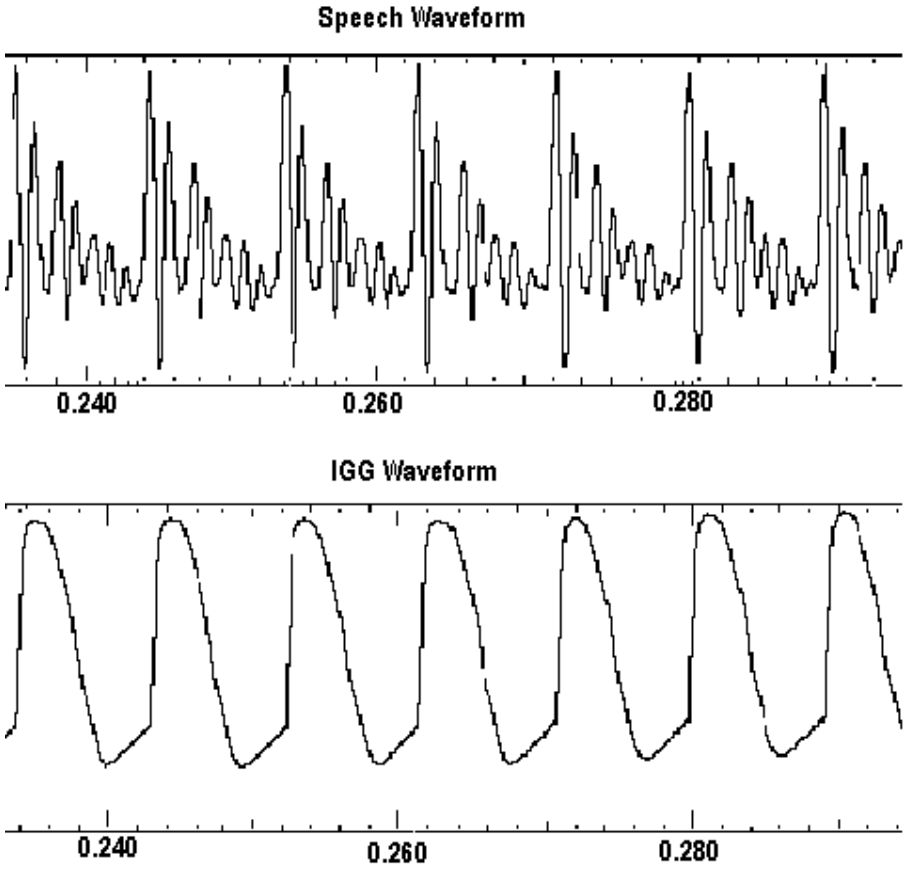


Fig 2.5. Speech and IGG waveform. Adapted from [20].

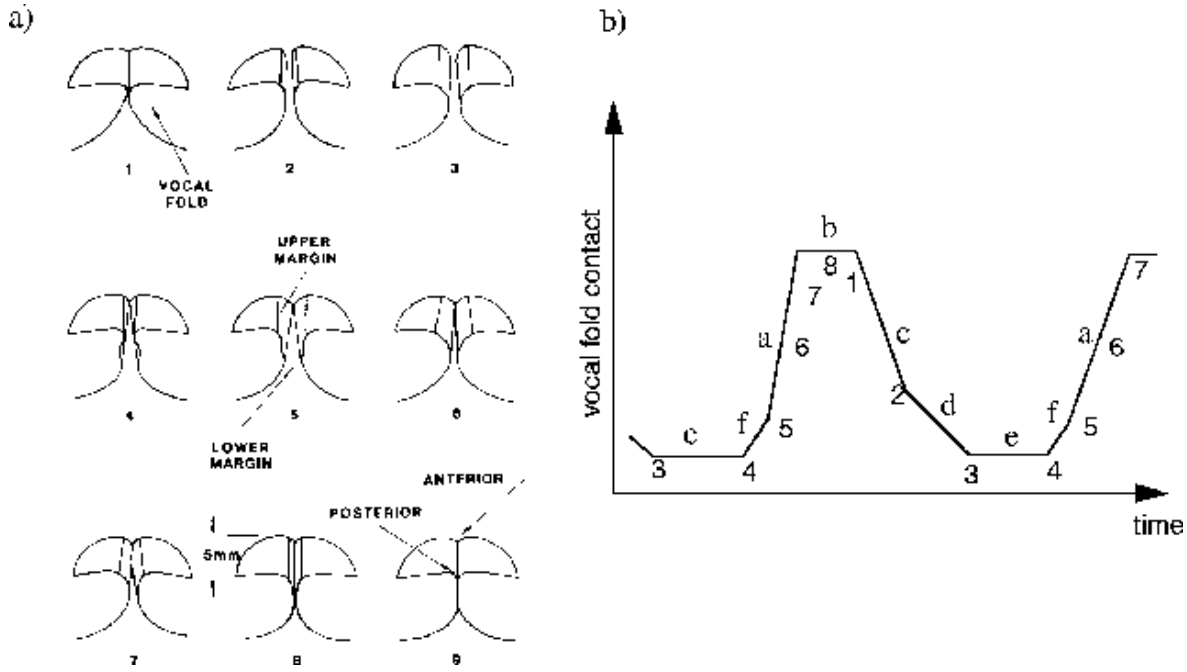


Fig 2.6. Representation of the vocal fold vibratory cycle. Adapted from [20].

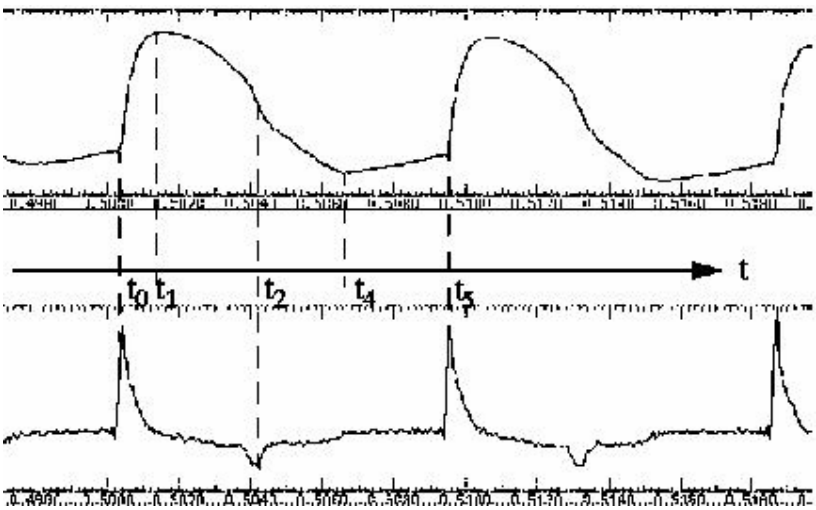


Fig 2.7. IGG signal and its time derivative. Adapted from [20].

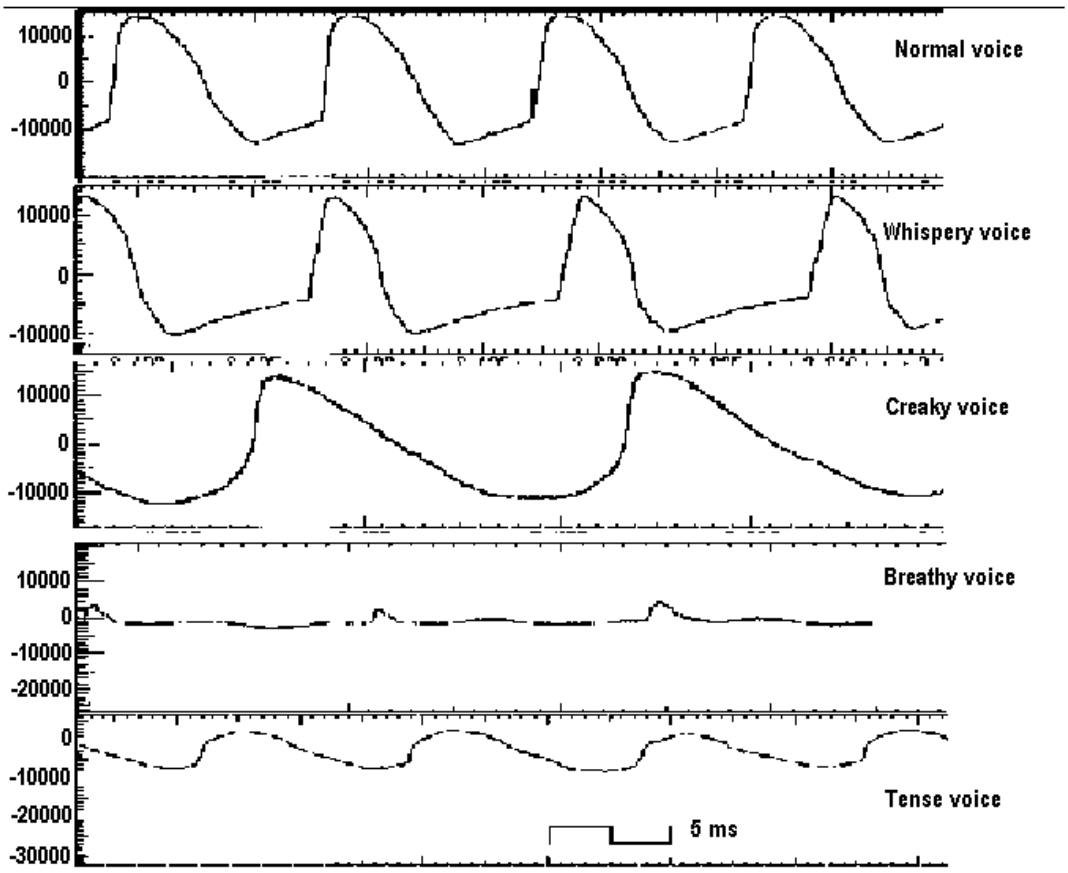


Fig 2.8 Impedance glottogram of a normal, whispery, creaky, breathy and tense voice. Adapted from [23]. x-axis is time, y-axis is in arbitrary units.

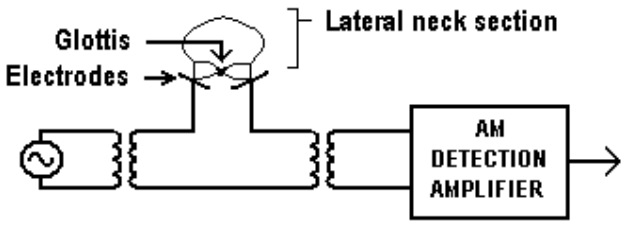


Fig 2.9 Block diagram of an impedance detection method described by Childer. Adapted from [19].

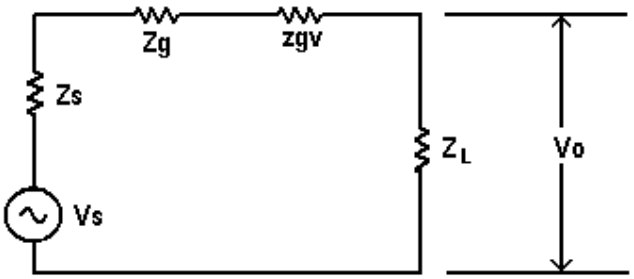


Fig 2.10 Equivalent circuit of impedance detection method described by Childer.

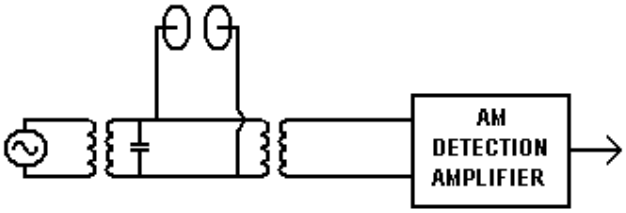


Fig 2.11 Impedance detection method used in instrument patented by Rothenberg. Adapted from [26].

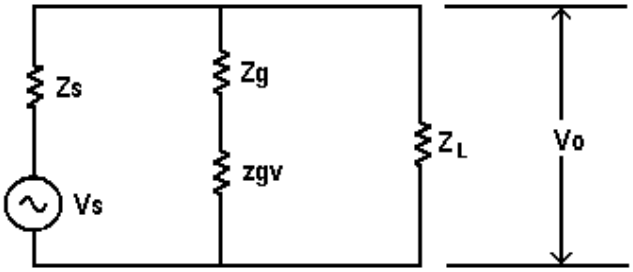


Fig 2.12 Equivalent circuit of Rothenberg's impedance detection method.

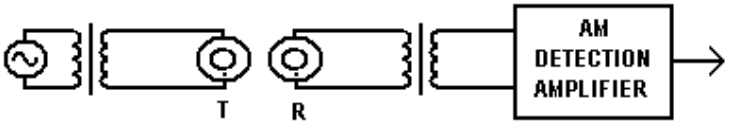


Fig 2.13 Impedance detection method used in instrument patented by Fourcin. Adapted from [25].

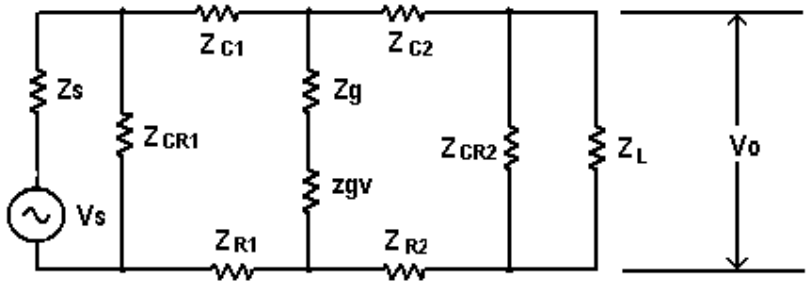


Fig 2.14 Equivalent circuit of Fourcin's impedance detection method.

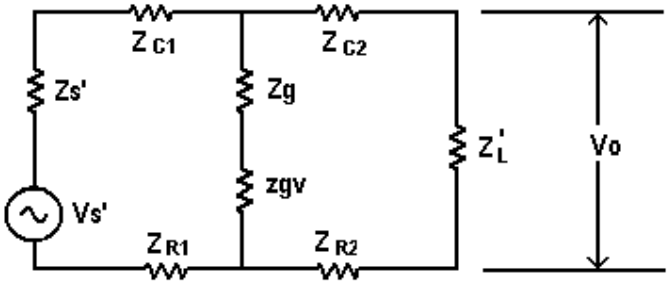


Fig 2.15 Reduced equivalent circuit of impedance detection method of Fourcin's.

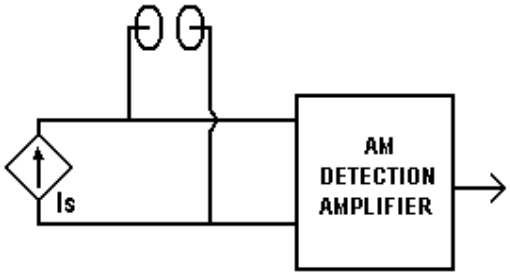


Fig 2.16 Impedance detection method used in the instrument developed at IIT Bombay [16][17].

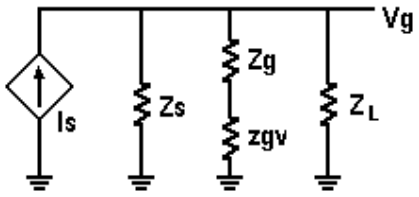


Fig 2.17 Equivalent circuit of impedance detection circuit based on current excitation, as used in the instrument developed at IIT Bombay [16][17].

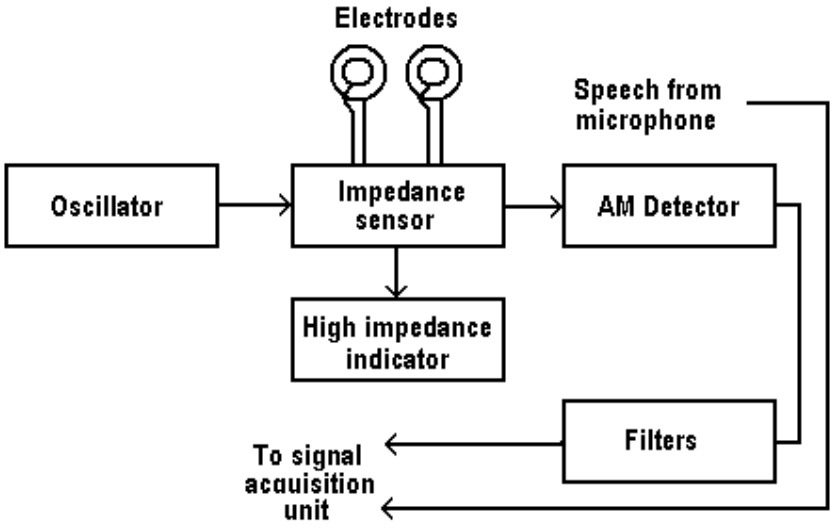


Fig 2.18 Block diagram of Impedance glottogram developed by Patil in 2000. Adapted from [17].

Chapter 3

HARDWARE DESIGN

3.1 Introduction

A block diagram of impedance glottograph, to measure the glottal impedance variation due to vibration of vocal folds, is shown in Fig 3.1. Excitation is generated by an oscillator, which provides a 400 kHz sinusoidal input to the impedance sensor, voltage to current converter. A pair of electrodes is held in contact with the skin on both sides of the thyroid cartilage. The electrodes are connected across the feedback path of the V/I converter circuit. Hence, a constant current flows through the two electrodes. When the vocal folds vibrate, the impedance between the vocal folds change. Since the current is constant, the voltage across the two electrodes changes in accordance with the impedance between the vocal folds and we get an amplitude modulated voltage waveform, across the electrodes. This signal is passed through the instrumentation amplifier, and given to the demodulator circuit. The demodulator circuit consists of a precision full-wave rectifier and low pass filter that suppresses the high frequency carrier. The demodulated output is given to filter module which consists of a notch filter to remove 50 Hz pick up, passive first order high pass filter to remove dc offset and fifth order elliptic low pass filter. The voltage output represents the impedance variation, and is known as impedance glottogram.

3.2 Oscillator module

For excitation, we need a sinusoidal oscillator to produce a frequency of approximately 400 kHz. Wien bridge oscillator was selected, for this purpose, considering several features like circuit simplicity, frequency selection and amplitude stability. The circuit diagram of the oscillator is shown in Fig 3.2 [27][28]. It is formed by op amp IC1 and associated circuitry. The frequency of oscillation is given by

$$f_o = 1/ 2\pi(R_{16}C_{16}\, R_{17}C_{17})^{0.5} \quad (3.1)$$

Selecting $R_{16} = R_{17} = 3.9$ kΩ, $C_{16} = C_{17} = 100$ pF we get $f_o = 408$ kHz.

The amplitude of the output is maintained at 6Vp-p using the JFET Q1 and associated circuitry by sensing the amplitude of the negative half cycle of the output. Increase in output voltage will result in increasing the gate bias and, thus, an increase in channel resistance and hence a decrease in the gain and a decrease in the output, and vice versa.

The oscillator circuit was assembled with three op amps namely LF 356 (JFET input, internally compensated op amp with f_T = 5 MHz), CA 3100 (uncompensated op amp with f_T = 38 MHz), TL 084 (JFET input, internally compensated quad op amp with f_T = 3 MHz). It was decided to use LF 356 for the oscillator, V/I converter, instrumentation amplifier and the demodulator module as the results with LF 356 and CA 3100 are comparable but LF 356 being JFET input (low bias currents), internally compensated and relatively less expensive is preferred. For the filter module, TL 084 is preferred as the results of all the three are comparable and TL 084 being a quad op amp effectively reduces the chip count and is thus preferred.

3.3 V/I converter module

The electrodes are driven by a sinusoidal current obtained using the V/I converter circuit, shown in Fig 3.3. The output of oscillator of Fig 3.2 is fed to the attenuator buffer formed using IC2, resistors R1, R41 and P2. The output of buffer, V2, can be varied by varying the resistor P2, to 0-2.5 V_{p-p} range to set the current level to be injected. The current source for current injection is the V/I converter formed using the op-amp IC3 and resistor R4. The injection current is given as

$$I_e = V_2 / R_4 \quad (3.2)$$

With R_4 = 2.2 kΩ, the current can be set in the range of 0 -1 mA peak to peak. The electrodes placed across the thyroid cartilage are put across terminals E1 - E2. i.e., in the feedback path of the op amp IC3. Capacitors C2 and C7 are used to prevent the passage of any dc current through the electrodes. These capacitors should have negligible leakage and their ac impedance at excitation frequency should be negligible as compared to 500 Ω

(maximum estimated impedance in the electrode path, when in contact with the skin). Resistor R2 has been put in the feedback path, to limit the dc gain of the circuit and to restrict the dc errors. Its resistance should be large compared to the impedance between the electrodes. One of the electrodes, E1, is at virtual ground. Therefore, the voltage across the electrodes is given by the output voltage

$$V_3 = I_e Z_g \tag{3.3}$$

where Z_g is impedance in the electrode path. Due to variation in Z_g , V_3 gets amplitude modulated.

3.4 High impedance indicator module

The instrument should indicate an inappropriately connected electrode. The high impedance indicator circuit is used to verify proper contact at the skin-electrode interface. The circuit diagram of the high impedance circuit used for the purpose is shown in Fig 3.4. Two indicator circuits are used, one for each half cycle of the sinusoidal waveform. A comparator compares the average of +ve half cycle of input V_2 to that of the output V_3. Another comparator is similarly used for –ve half cycle. If the electrodes are not properly connected, both the LED's are turned on. If both LED's remain off, it implies electrodes are properly connected. If one of the LED is turned on, it indicates corresponding half cycle is in saturation. It was decided to compare rectified average values of the waveform, rather than the instantaneous values in order to reduce switching currents.

3.5 Instrumentation Amplifier

The signal obtained from the electrodes may have common mode noise pick up. To remove this common mode noise pick up, an instrumentation amplifier is used. The instrumentation amplifier is built around IC 6, IC 11, IC 12. The output at the two electrodes is given as an input to the instrumentation amplifier. The input signal is high pass filtered using a passive first order high pass filter comprising C21, R13 and C22, R20 with a cut off frequency of 185 kHz. The gain of the amplifier is given by

$$A_v = V_4 / (V_{3a} - V_{2a}) = (1+ 2R_{f1} / R_{i1})\, R_{f2} / R_{i2} \quad (3.4)$$

where $R_{f1} = R_{52} = R_{53}$, $R_{i1} = R_{54}$, $R_{f2} = R_{55} = R_{56} + P_3$, $R_{i2} = R_{57} = R_{58}$

With the values shown in Fig 3.5, the gain of first and second stages are 1.3 and 6.8 respectively, with an overall gain of 8.84. A low pass filter is formed by R55, C24 and R56, P3, C25. The cut off frequency of low pass filter is 709.2 kHz ≈ 700 kHz.

3.6 Demodulator module

The output V4 from the instrumentation amplifier is capacitively coupled to the demodulator circuit, shown in Fig 3.6. It consists of a full-wave precision rectifier and first order low pass filter. The precision half-wave rectifier is formed using op amp IC7, R21, R22, R23, D7, D8. The output of the precision half wave rectifier and the input are summed by the adder to give precision full-wave rectification [28], by selecting $R_{73} + P_4 = R_{24} / 2$. Capacitor C28 in parallel with R26 provides low pass filtering with the cut off frequency of

$$f_c = 1/ 2\pi R_{26} C_{28} \quad (3.5)$$

Selecting R_{26} = 22 kΩ, C_{28} = 560 pF, we get f_c =13 kHz

The low pass filter removes harmonics of the high frequency carrier and the demodulated signal is obtained at the output of the low pass filter.

3.7 Filter module

The impedance glottogram waveform needs a 80 Hz – 5 kHz bandwidth for its faithful acquisition. In order to have reasonably good phase response in the pass band, filters with sharp transition cannot be used. Hence it was decided to treat 5 – 10 kHz as transition band, with an attenuation of > 40 dB for frequencies > 10 kHz. This has been realized as fifth order elliptic state variable low pass filter with a cut off frequency of 5 kHz. To eliminate the dc offset and attenuate the 50 Hz pick up, we have used a filter comprising of cascade of notch filter to attenuate 50 Hz pick up and a first order passive high pass filter with a cut off of 2 Hz.

The notch filter is formed using IC 13 and associated circuitry as shown in Fig 3.7. By summing the high pass and low pass outputs of the state variable filter, the notch filter is made [28].

The circuit diagram of cascade of first order passive high pass filter and fifth order elliptic low pass filter is shown in Fig 3.9. The low pass filter is realized by cascading second order and third order state variable sections as shown in Fig 3.8. The transfer function of this section shown in Fig. 3.8 is

$$T(s) = -\left(\frac{R_6}{R}\right)\left(\frac{s^2 + \left(\frac{1}{R_2 R_3 C^2}\right)\left(1 + \left(\frac{R_3 R}{R_4 R_5}\right)\right)}{s^2 + s\frac{1}{R_1 C} + \frac{1}{R_2 R_3 C^2}}\right) \tag{3.6}$$

where $R_1 = R_4$ and $R_2 = R_3$. The numerator roots result in a pair of imaginary zeros, and the denominator roots determine a pair of complex poles.

For this design, the specifications are : maximum ripple in pass band < 1 dB, minimum attenuation in stop band > 45 dB. The transition band is taken as 5 – 10 kHz. For a steepness factor, $A_s = f_s/f_c = 2$, $\theta = 30^o$, $\rho = 25$ % ($R_{dB} = 0.28$ dB) [29]. The normalized complex poles and zeros are

First section : $\alpha = 0.3386$, $\beta = 0.6879$, $\omega_\alpha = 3.0480$

Second section : $\alpha = 0.1091$, $\beta = 1.0310$, $\omega_\alpha = 1.9690$

Real Pole : $\alpha_0 = 0.4673$

The denormalized poles and zeros are as follows

First section : $\alpha' = 10637$, $\beta' = 21611$, $\omega_\alpha' = 95756$

Second section : $\alpha' = 3427$, $\beta' = 32390$, $\omega_\alpha' = 61858$

Real Pole : $\alpha_0 = 14681$

First section component values are calculated as follows :

Let $R = 100$ kΩ, $C = 1$ nF, $A = 158$ to 316

$\omega_0' = 24087$

$$R_1 = R_4 = 1/2\ \alpha' C \tag{3.7}$$

With $\alpha' = 10637$, $C = 1$ nF, $R_1 = R_4 = 47$ kΩ

$$R_2 = R_3 = 1/\ \omega_0' C \tag{3.8}$$

With $\omega_0' = 24087$, $C = 1$ nF, $R_2 = R_3 = 41.5$ kΩ ≈ 41.2 kΩ (39 kΩ in series with 2.2 kΩ)

$$R_5 = 2\,\alpha'\omega_0' R/((\omega_\alpha')^2-(\omega_0')^2) \quad (3.9)$$

With α'=10637, ω_0' = 24087, R = 100 kΩ, ω_α'= 95756 we get R_5= 5.9 kΩ ≈ 6 kΩ (3.3 kΩ in series with 2.7 kΩ)

$$R_6 = AR(\omega_0'/\omega_\alpha')^2 \quad (3.10)$$

With A = 158 to 316, R = 100 kΩ, ω_0' = 24087, ω_α'= 95756 we get R_6 = 1 MΩ to 2 MΩ.

Second section component values are calculated as follows :

Let R = 100 kΩ, C = 1 nF, A = 1

ω_0' = 32570

Using (3.8), with α'= 3427, C = 1 nF,

$R_1 = R_4$ = 145.9 kΩ ≈ 147 kΩ (120 kΩ in series with 27 kΩ)

Using (3.9), with ω_0'= 32570, C = 1 nF,

$R_2 = R_3$ = 30.7 kΩ ≈ 30.3 kΩ (27 kΩ in series with 3.3 kΩ)

Using (3.10), with α'= 3427, ω_0' = 32570, R = 100 kΩ, ω_α'= 61858 we get

R_5 = 8 kΩ ≈ 8.2 kΩ

Using (3.11), with A = 1, R = 100 kΩ, ω_0' = 32570, ω_α'= 61858 we get

R_6 = 27.7 kΩ ≈ 27 kΩ

Last section component values are calculated as follows :

$$C_6 = 1/\alpha_0' R_6 \quad (3.11)$$

With α_0'= 3427, R_6 = 27 kΩ we get

C_6 = 2.45 nF ≈ 2.42 nF (2.2 nF in parallel with 220 pF)

3.8 Electrodes used

In impedance glottography, 2-or 4-electrode arrangement can be used. In 2-electrode arrangement, as shown in Fig 3.10, both the electrodes are used for current injection and voltage sensing. In the 4-electrode arrangement, as shown in Fig 3.11 various configurations can be used for current injection and votage sensing, as shown in Fig 3.12a-c.

In the configuration shown in Fig 3.12a, both current injection and voltage sensing is done by center electrodes. And the ring electrodes are kept floating. This has no advantage as compared to the 2-electrode arrangement. In the configuration shown in Fig 3.12b, both current injection and voltage sensing is done by center electrodes. And the ring electrodes are grounded. In this arrangement, common mode pick up is reduced

but the disadvantage of this configuration is that some current passes superficially from the center electrodes to the ring electrodes, thus reducing the sensitivity of the instrument. In the configuration shown in Fig 3.12c, both current injection and voltage sensing is done by center electrodes. And the ring electrodes are actively driven by buffers driven by the respective center electrodes. This can reduce the common mode pickup, provides better directivity to the excitation current and should increase the sensitivity. But some additional current flows through the larynx, between the two ring electrodes.

The choice of electrode configuration is very important as the sensitivity of the circuit is affected by the electrode configuration. If the electrodes are used without guard rings, that is, in the 2-electrode arrangement, current finds an alternate path superficially across the skin, thus reducing the sensitivity of the circuit. In our design, the 4-electrode arrangement of Fig 3.12c is modified, as shown in Fig 3.13. The electrodes are connected across the feedback path of the voltage to current converter. So CE1 is at virtual ground, and thus, RE1 need not be actively driven and can be connected to the ground. CE2 is connected to the output of the op amp, thus, RE2 need not be actively driven and can be shorted to the CE2. Drawback of this method is that cable capacitance and the impedance between CE1 and RE1 will come in parallel with the input impedance of IC3 and hence may degrade the performance of V/I converter. Therefore, a careful experimental evaluation of the arrangements is necessary.

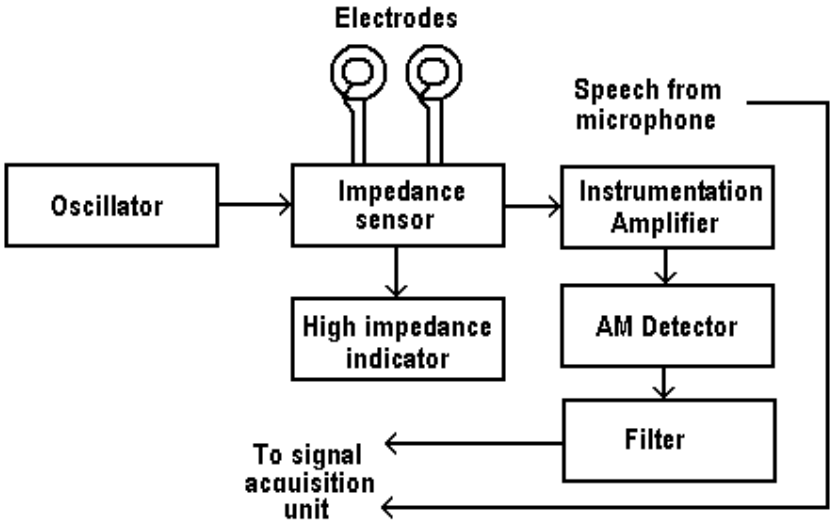


Fig 3.1. Block diagram of the impedance glottograph.

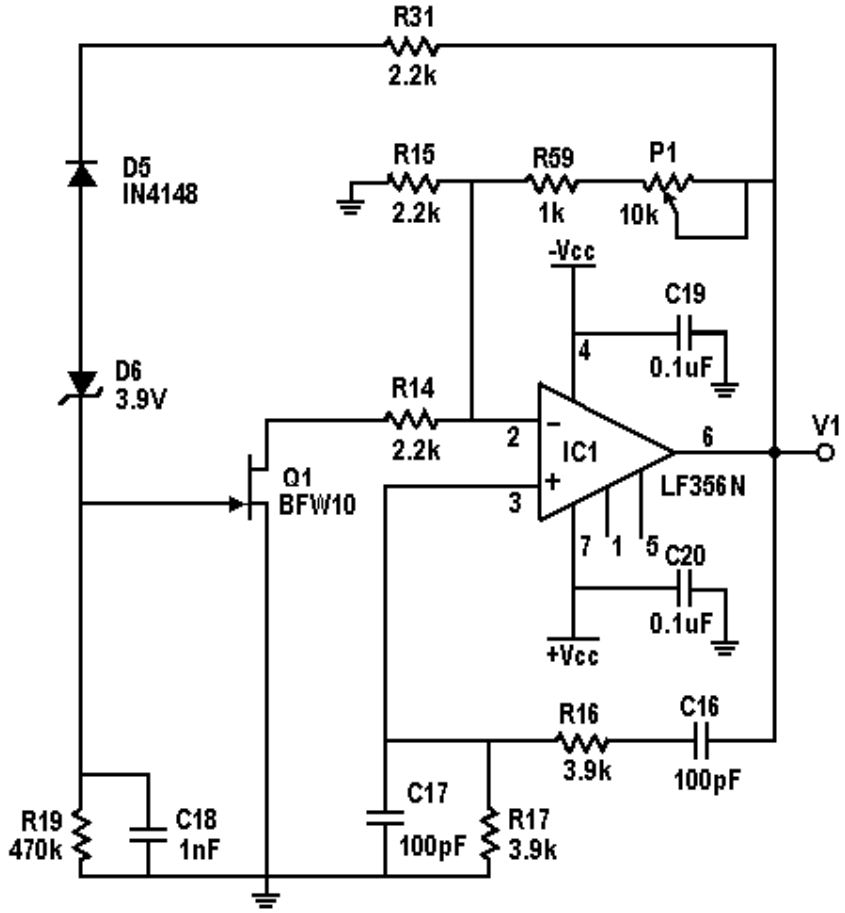


Fig 3.2. Wien bridge oscillator.

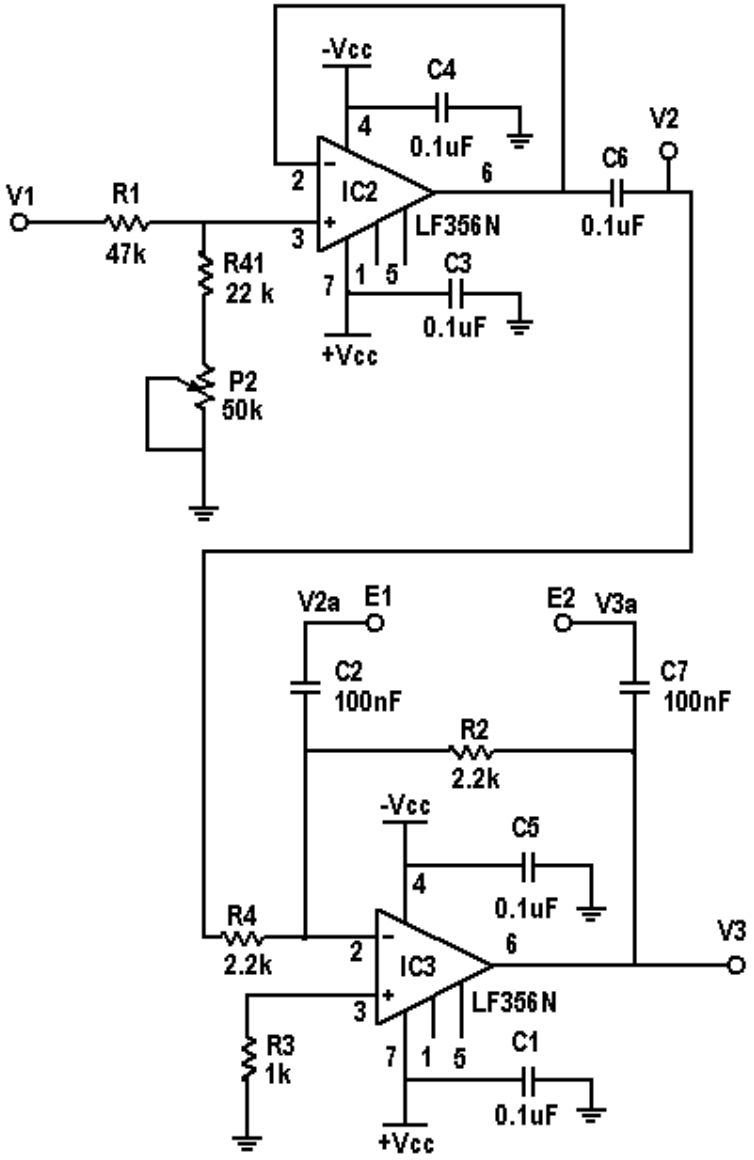


Fig 3.3. V/I converter module.

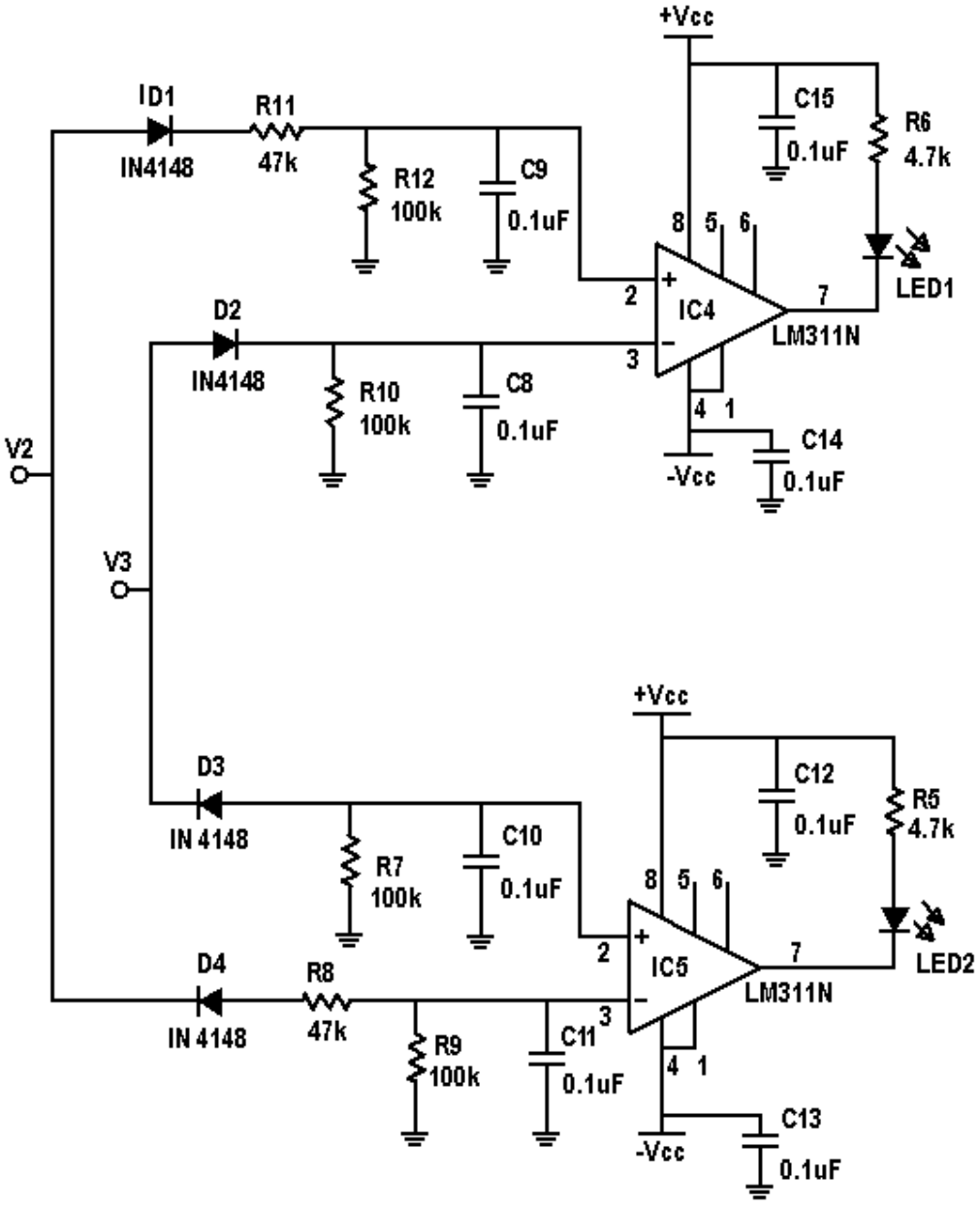


Fig 3.4. High impedance indicator module.

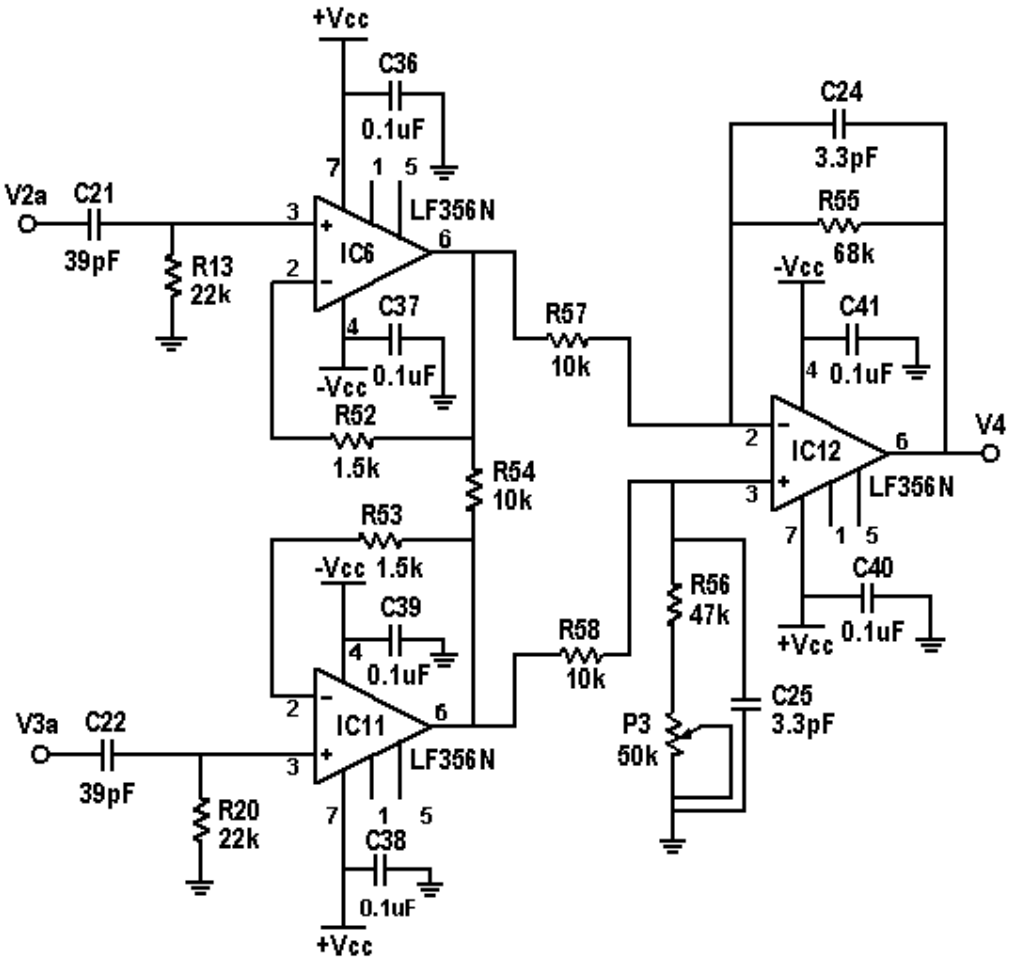


Fig 3.5. Instrumentation amplifier.

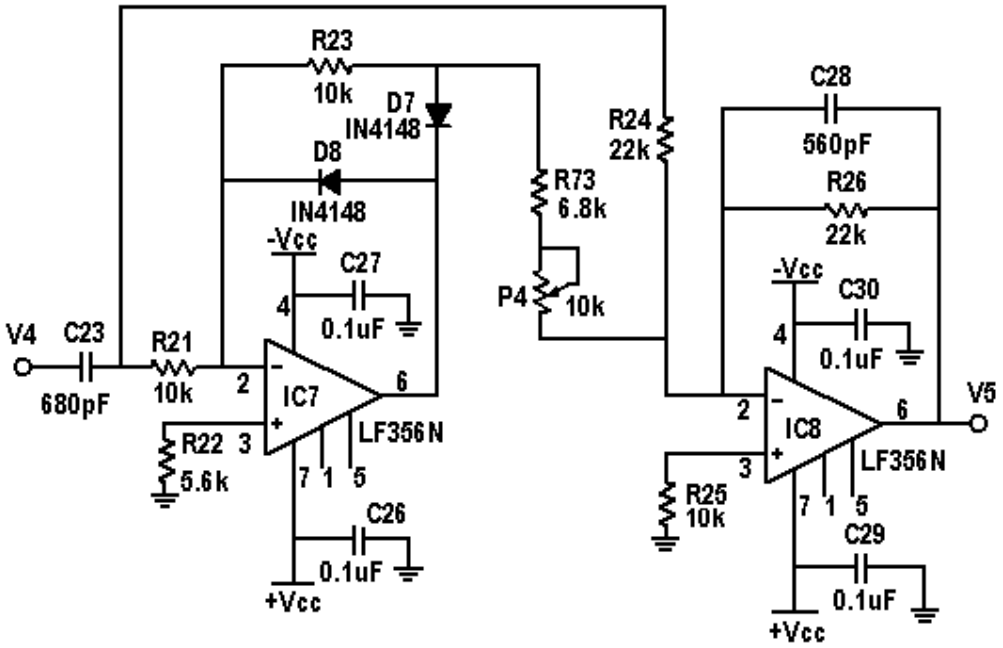


Fig 3.6. Demodulator circuit.

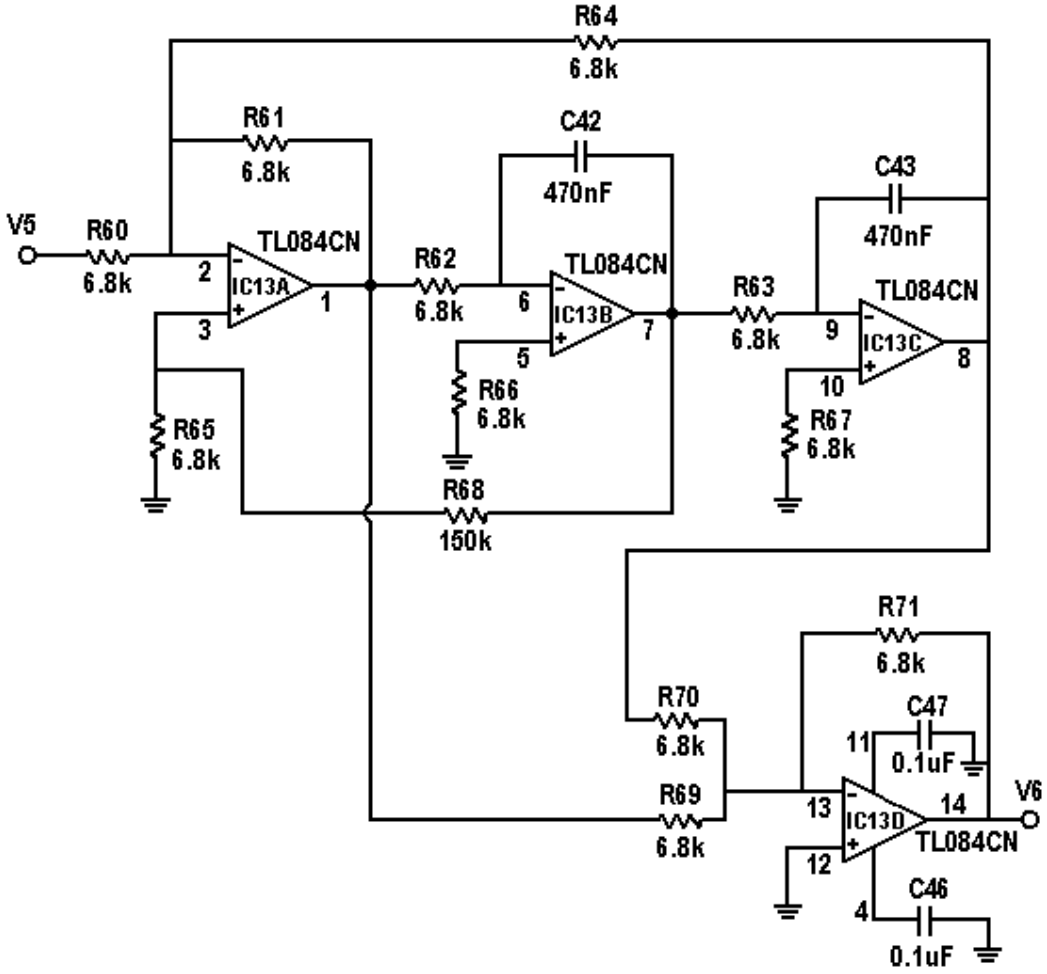


Fig 3.7. Notch filter

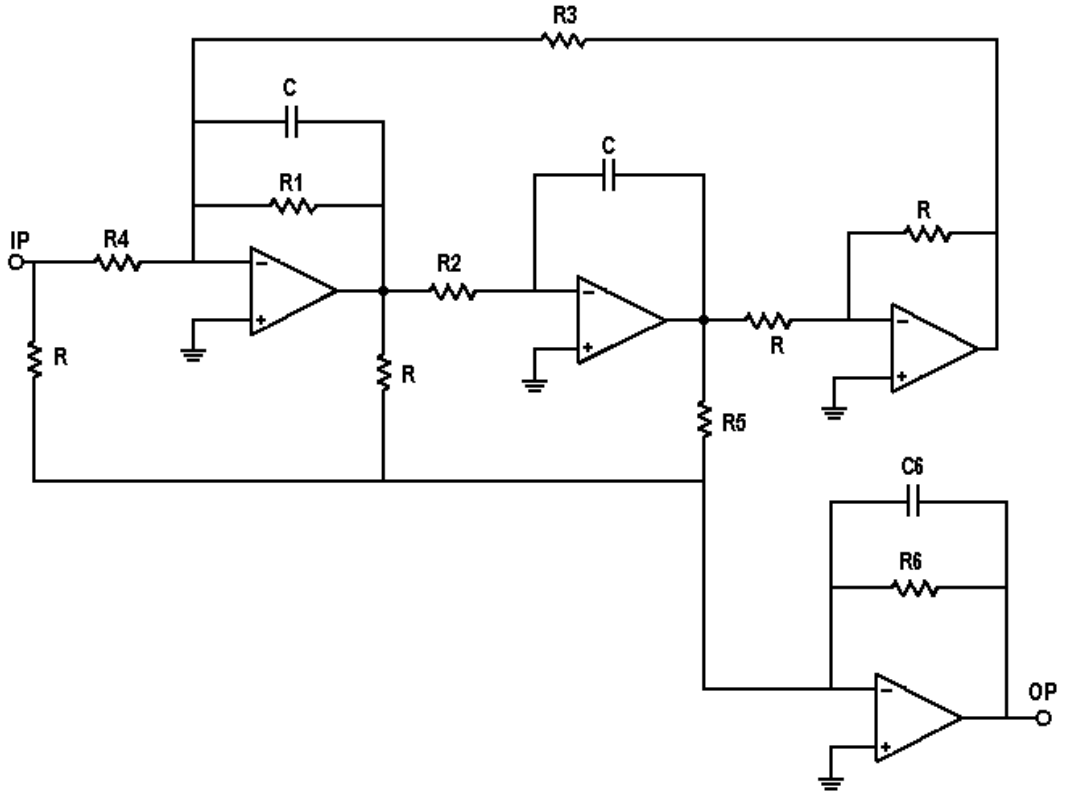


Fig 3.8. Schematic of a state variable low pass filter

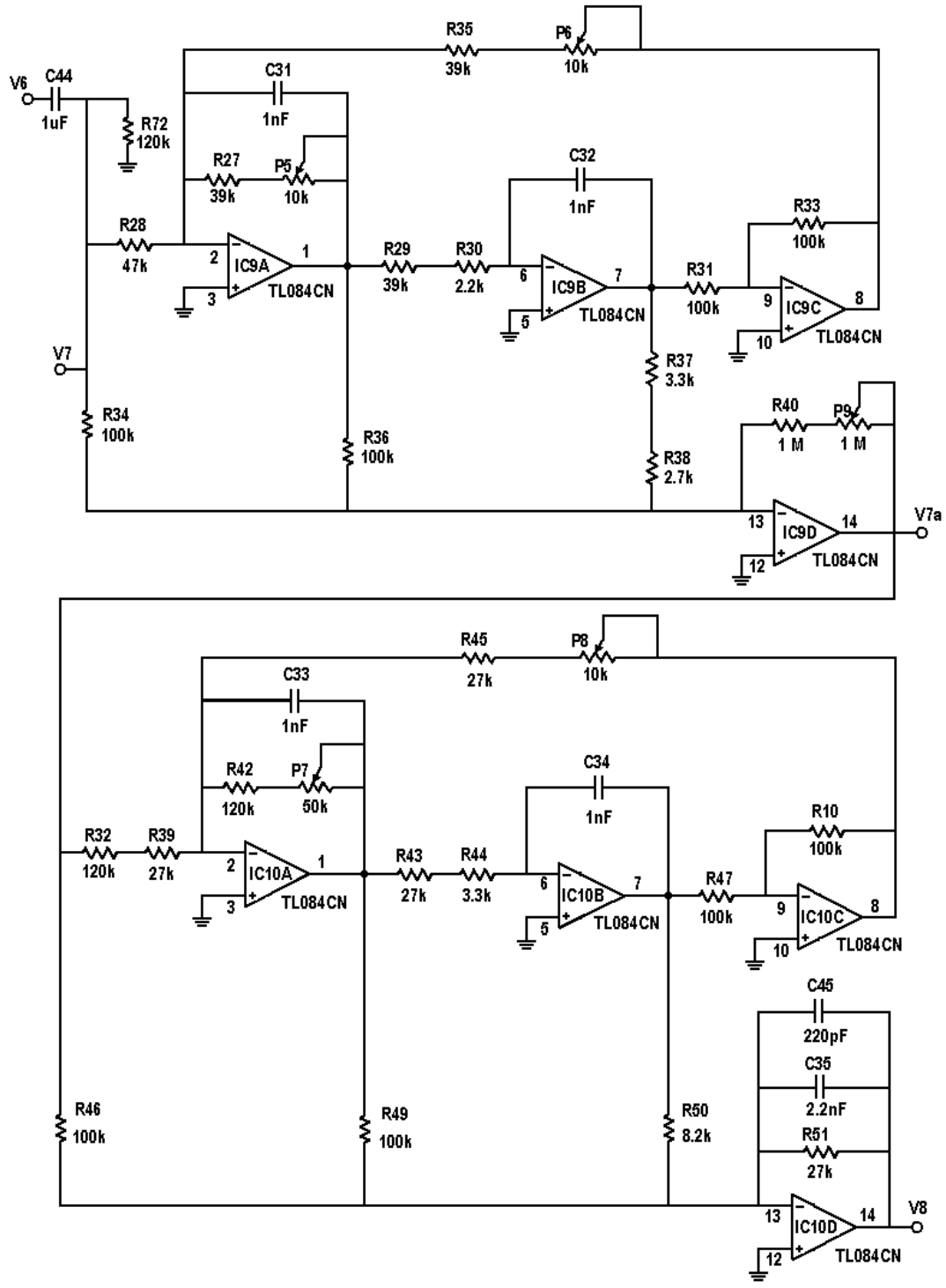


Fig 3.9. Cascade of passive first order high pass filter and fifth order elliptic state variable low pass filter.

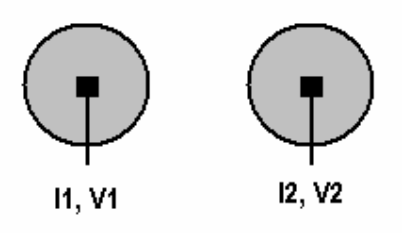


Fig 3.10. 2 electrode arrangement

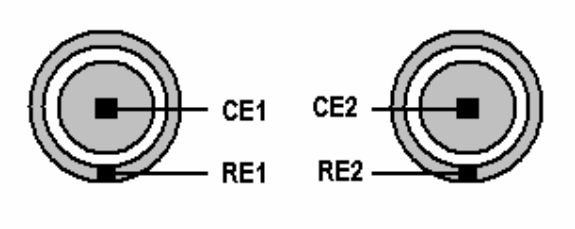
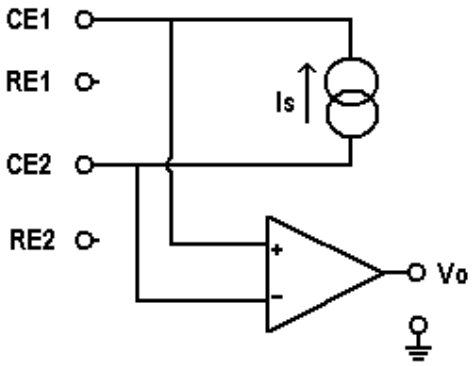
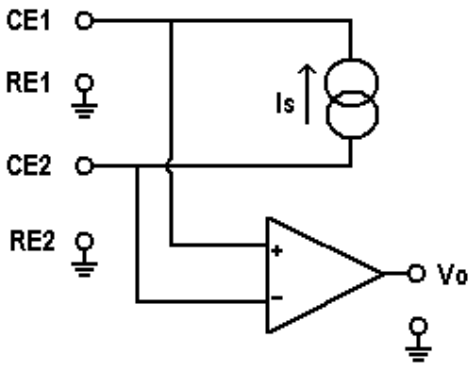


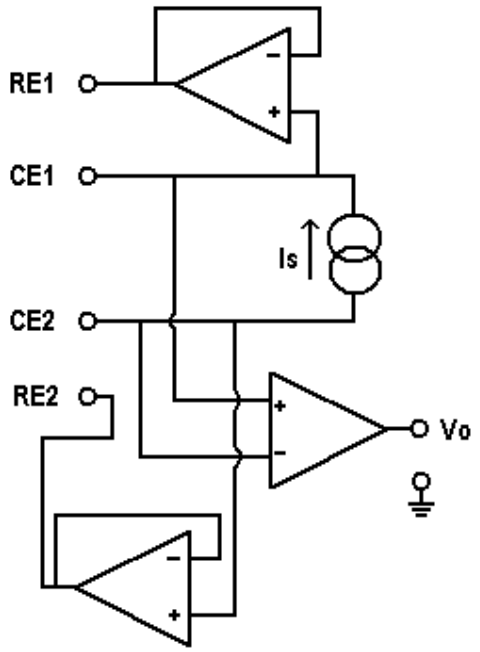
Fig 3.11. 4 electrode arrangement



(a)



(b)



(c)

Fig 3.12a-c. Various 4 electrode configurations

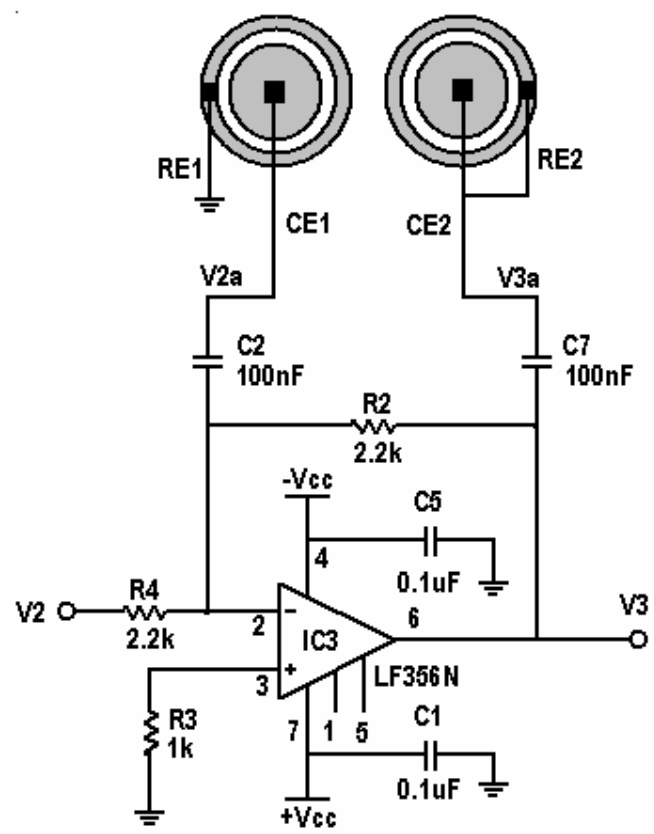


Fig 3.13. Connections of electrodes to the V/I converter

Chapter 4

SIGNAL ACQUISITION AND ANALYSIS

4.1 Introduction

Signal acquisition, display and analysis system has been developed around a PC. Signal acquisition is done using the two channel inputs of the sound card. From the acquired waveform, a segment can be selected and displayed, and it can be analyzed for glottal pulse, pitch frequency Fx plot, single period and multi period time and frequency histograms. This software, “siggada” is Windows based, developed in Visual Basic. Basics of this software implementation and operation are presented in this chapter. Detailed description of the software is given in Appendix 'E'.

4.2 Signal acquisition

Stereo line input of the PC sound card is used for simultaneously acquiring speech waveform on one channel and IGG waveform on the other channel. Speech waveform is obtained using a microphone and amplifier. The IGG waveform is obtained from the hardware developed. The two channels are connected to the line in of the sound card through a stereo jack. The acquired signal is saved as 16-bit signed integer valued samples. The sampling rates can be selected from 8, 10, 11.025, 16, 20, 22.050, 32 or 44.1 K Sa/s. For stereo channel recording at 44.1 K Sa/s, 162 MB of RAM memory is required for a recording of 15 minutes and 59 seconds. Recording duration has, thus, been provided from minimum of 1 second to maximum of 15 minutes and 59 seconds, in increments of 1 second. For longer recordings, "Goldwave" or some other audio recording software may be used and the signals stored as 16-bit signed integer ".wav" format for analysis by “siggada”. Signal acquired using “siggada” or other audio recording software can be in the stereo mode for speech as well as impedance glottogram waveform, or mono for impedance glottogram waveform only.

The speech waveform and IGG waveform can be played to be heard individually or together using stereo output. It has also been provided to play the selected speech and IGG segments independently and also as a stereo segment.

4.3 Pitch Calculation

For calculation of glottal pitch period, the IGG waveform is converted to glottal pulses by passing it through a band pass filter and a hysteresis comparator using a technique developed earlier in our lab [4]. Analysis is carried out assuming a sampling rate of 22 kSa/s. A block diagram of signal processing is given at Fig 4.1. The band pass filter is a 257- coefficient FIR filter with 10 - 600 Hz bandwidth. The average $a(n)$ is calculated by using a first order low pass filter as the input IGG waveform $x(n)$

$$a(n) = \alpha a(n\text{-}1) + (1\text{-}\alpha)x(n) \qquad (4.1)$$

The peak $p(n)$ and valley $v(n)$ are calculated with a fast attack response (towards the peak / valley) and slow release response (towards the average)

$$\begin{aligned} p(n) &= \beta p(n\text{-}1) + (1\text{-}\beta)\, x(n) && \text{for } x(n) \geq p(n\text{-}1) \\ &= \gamma p(n\text{-}1) + (1\text{-}\gamma)\, a(n) && \text{otherwise} \qquad (4.2) \\ v(n) &= \beta v(n\text{-}1) + (1\text{-}\beta)\, x(n) && \text{for } x(n) \leq v(n\text{-}1) \\ &= \gamma v(n\text{-}1) + (1\text{-}\gamma)\, a(n) && \text{otherwise} \qquad (4.3) \end{aligned}$$

Value $\alpha = 0.997$, $\beta = 0.5$ and $\gamma = 0.995$ as suggested earlier in [4] for satisfactory performance have been used. Provision is made in the software to change these values in the run time. So one can see the effect of varying the values on the same segment in another window. Thresholds for the hysterisis comparator, $x_L(n)$ and $x_H(n)$ are calculated from the average $a(n)$, peak $p(n)$, valley $v(n)$.

$$x_L(n) = a(n) - 0.1(p(n)\text{-}v(n)) \qquad (4.4)$$

$$x_H(n) = a(n) + 0.1(p(n)\text{-}v(n)) \qquad (4.5)$$

The hysteresis comparator is used to convert the input into a rectangular waveform. The output of this hysteresis comparator is given as

$$\begin{aligned} y(n) &= 1 && \text{if } x(n) \geq x_H(n) \\ &= -1 && \text{if } x(n) \leq x_L(n) \\ &= y(n-1) && \text{otherwise} \end{aligned} \qquad (4.6)$$

The rectangular glottal pulses obtained can be displayed. Pich periods are calculated by counting the number of samples between alternate zero crossings.

$$\text{Pitch value} = \text{Sampling rate} / \text{No. of samples in one cycle} \qquad (4.7)$$

The instantaneous pitch values thus obtained are plotted as Fx-plot.

4.4 Histogram

Histogram of the pitch values can be used for studying voice disorders [21][24]. These are obtained from the pitch values calculated as in (4.7). Histogram can be plotted on the basis of the number of pitch periods or on the basis of time, and the two can be called as frequency and time histogram respectively. The frequency range from 0 to 1000 Hz is divided in a fixed number of bins (from 10 to 50). Same segment can be analyzed by plotting various histograms with different number of bins. Histograms are obtained by updating the bin counts after each glottal pulse. For frequency histograms, count in a frequency bin is incremented by one when the pitch frequency falls within that bin. For time histogram, the corresponding bin count is incremented by the number of samples in the glottal pulse. At the end of the segment, histogram is obtained by dividing the bin counts by the sum of all the bin counts. The histograms calculated by the above method are single period histograms. For *m* period histograms, value of each bin is incremented when m successive pitch values fall in the same bin. Normalizing is done by dividing the bin counts by the total number of pitch periods for the frequency histogram and the total number of samples in the segment for the time histogram. Multi period histograms show the stability of pitch frequency.

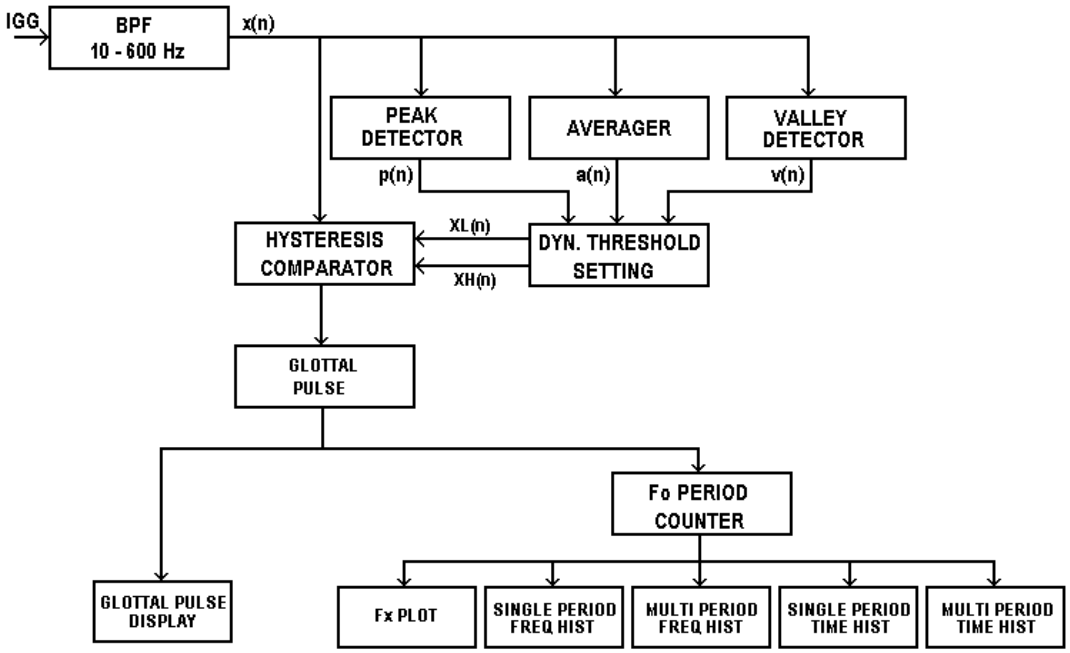


Fig 4.1 Signal processing for glottal pulse, Fx plot and histograms

Chapter 5

CIRCUIT ASSEMBLY AND SYSTEM TESTING

5.1 PCB Design.

The IGG circuit has been assembled on a double sided PCB with plated through holes (PTH). It consists of Wien bridge oscillator, impedance detector, high impedance indicator, instrumentation amplifier, demodulator and filter circuits. Size of PCB is 105 mm by 140 mm. The PCB layout has been designed with special consideration for reducing the power supply noise by providing tight coupling between the supply and ground, stabilizing the ground by providing a large ground plane, and reducing the noise pick-up by shielding the signal lines with ground plane. Ground plane has been provided on the component side, and supply and signal lines are kept on the solder side to reduce environmental pick up. Care is taken to minimize the length of the supply path for ICs. The entry points of the dc supply on the PCB are decoupled by 10 uF/63 V electrolytic capacitors in parallel with 0.1 uF ceramic capacitor. Positive and negative supply for each IC is decoupled by 0.1 uF surface mounted capacitors placed as close as possible to the IC. Surface mounted capacitors of 0.1 uF are also provided where ever the supply lines change side and this has been kept to the minimum. The electrodes are connected to the PCB by shielded coaxial cable. The instrument is powered by ± 9 V dual supply. Two 9 V batteries may be used for this. The PCB layouts are shown in Appendix C.

5.2 Assembly

The IGG instrument has been assembled in a cabinet. The cabinet is made of 3 mm acrylic sheet. All the controls, indicators, input and output connectors have been provided on the front panel, as shown in Appendix F. Input and output connectors are PCB mounted RCA connectors. The box has been designed, fabricated and assembled to provide easy access to all the connectors, controls and the PCB when the box is opened for testing. The cabinet design and dimensions are shown in Appendix D. The list of components with approximate cost is shown in Appendix B.

5.3 Testing of the circuit

The circuit comprises of six modules namely oscillator module, voltage to current converter module, high impedance indicator module, instrumentation amplifier, demodulator module, and filter module. Each module was first tested separately.

The oscillator module was tested for amplitude stability against power supply variations. For ± 9 V supply, the output voltage was 6.03 Vp-p. A plot of variation in the output voltage is shown in Fig. 5.1. The amplitude of the carrier signal is stable at 6.03 Vp-p, with p-p variation < 30 mV for power supply variations from ± 5.5 V to ± 10 V. As the supply is lowered from ± 5.5 V to ± 5 V, the p-p amplitude drops by 220 mV. The short time and long time amplitude stability of the oscillator was also monitored. In short time, the positive peak varied from + 2.99 V to + 3.03 V and, the negative peak varied from - 2.99 V to - 3.00 V. The RMS value was found to vary from 2.08 V to 2.1 V. Long time stability was checked by keeping the oscillator on for 10 hrs and the output was measured every one hour. It was found that oscillator output was stable at 6.03 V with p-p variation < 30 mV.

The voltage to current converter was tested for linear response by connecting different loads across CE1 and CE2. The response was found to be linear for a load resistance from 330 Ω to 1 kΩ. The plot of load resistance vs output voltage is as shown in Fig. 5.2.

High impedance indicator has been designed to indicate proper contact at the skin electrode interface. It has been tested by connecting different loads across CE1 and CE2. LEDs turn off for load resistance upto 1 kΩ indicating proper contact and are not turned off for loads above 1 kΩ indicating improper contact.

Instrumentation amplifier is designed to reject the common mode pick up. It was tested by using a 50 mV differential input and 2.19 V common mode input. Differential gain and common mode gain obtained for various frequencies are plotted in Fig 5.3. Differential gain and common mode gain obtained for 400 kHz was 13.1 dB and -18.7 dB respectively giving a CMRR of 31.8 dB.

Demodulator module together with filter module was tested for detection of modulating signal for 1% modulation index with a carrier of 400 kHz, for varying modulating sinusoidal frequencies from 100 Hz to 5 kHz. The output obtained was a sinusoidal wave of the respective frequency used for modulating the carrier. The

demodulator module together with filter module was also tested for detection of modulating signal for 1% modulation index with varying carrier frequencies from 60 kHz to 480 kHz. The amplitude of the output was found to vary from 50 mV to 200 mV for the same input for varying carrier frequencies.

Filter module is designed to obtain the signal between 80 Hz and 5 kHz and attenuate the 50 Hz pickup and remove dc offset and attenuate frequencies greater than 10 kHz by around 40 dB. Thus, a notch filter is used to attenuate the 50 Hz pick up, first order passive high pass filter is used to remove the dc offset and fifth order elliptic state variable filter is used to attenuate the signal greater than 10 kHz. The cutoff frequencies of high pass filter is 2 Hz. The cut off frequency of low pass filter is 5 kHz and the stop band frequency for low pass filter is 10 kHz. Maximum ripple in pass band is less than 1 dB. Attenuation of low pass filter is ~ 55 dB. The magnitude and phase responses of the cascade of notch filter, high pass and low pass filter are shown in Fig. 5.4 and Fig. 5.5 respectively. The filter exhibits a near flat magnitude response in 75 Hz – 5 kHz band, which is needed for the glottal pulse. Also the phase response is seen to be nearly linear in this band. Band width limitation and deviation in linear phase response cause distortion in the IGG waveform. As an indication of overall response of the filter for glottal pulses, filter response was observed for square waves of different frequencies. A typical output is shown in Fig. 5.6, and the characteristic parameters of the output pulse are given in Table 5.1. The output was also seen by varying the input voltage. It was seen that t1, t2, f1, f2 don't vary with input variation. However, Vo, Vo1, Vo2, d1, d2, d3, d4 vary linearly with variation in input voltage.

Table 5.1. Square wave response of the demodulation filter

f (Hz)	**Vin (mV)**	**Vo (V)**	**Vo1 (V)**	**Vo2 (V)**	**d1 (V)**	**d2 (V)**	**d3 (V)**	**d4 (V)**	**t1 (μs)**	**t2 (μs)**	**f1 (kHz)**	**f2 (kHz)**
100	43.2	7.19	3.69	3.50	1.50	1.50	1.81	1.81	216	216	5	5
200	43.2	6.5	3.31	3.19	1.50	1.38	1.25	1.19	216	216	5	5
300	43.2	6.5	3.31	3.19	1.50	1.50	1.19	1.13	216	216	5	5
400	43.2	6.44	3.25	3.19	1.56	1.56	1.13	1.13	216	216	5	5
500	43.2	6.5	3.31	3.19	1.56	1.56	1.13	1.13	212	212	5	5

A comparison of the instrument developed is made with available instruments at Appendix G.

5.4 Testing of the hardware with impedance simulator

The instrument developed was tested with an impedance simulator. The design of the simulator is discussed at Appx F. The impedance simulator was developed by Patil [17]. It simulates the vocal fold opening and closing as step changes. The variation in impedance is ~ 1 Ω over a base impedance of ~ 590 Ω. The output of the instrument at various frequencies is shown in Fig. 5.7.

5.5 Testing of the signal acquisition, display and analysis software

Signal acquisition, display and analysis system has been developed around a PC. Signal acquisition is done using the two channel inputs of the sound card. From the acquired waveform, a segment can be selected and displayed, and it can be analyzed for glottal pulse, pitch frequency Fx plot, single period and multi period time and frequency histograms.

The analysis software has been tested by generating data file for sine wave, whose frequency is swept linearly from 100 Hz to 400 Hz over a period of 3 seconds. Fig 5.8 shows the Fx plot and single period frequency and time histogram for the swept sine wave. For swept sine wave, Fx plot rises linearly, single period frequency histogram rises linearly whereas single period time histogram remains more or less constant

5.6 Acquiring IGG from subjects

The glottal pitch extraction was done by recording IGG waveform from impedance glottograph hardware and acquiring and analyzing signal using the software developed. Recordings are done for subjects with normal vocal folds. The recordings obtained are given in Fig 5.9 to 5.12.

The speech segment, IGG segment, filtered IGG segment are shown for sustained vowel /*a*/, /i/, /u/ for a male and a female speaker in Fig 5.9 and 5.10 respectively. The recordings are displayed for 100 ms for male speaker and 50 ms for female speaker. The pitch lies in the range 100 to 120 Hz for male speaker AL and 200 to 220 Hz for female

speaker AC. It may be seen that the IGG waveform is almost the same for the same speaker uttering different vowels.

The single period and multi period, time and frequency histograms are plotted for reading a continuous text

"*It is raining cats and dogs. The streets are flooding. Tide is high.*"

in Fig 5.11 and 5.12 for male speaker AL and female speaker AC respectively. It was again found that the pitch lies in the range 100 to 120 Hz for male speaker AL and 200 to 220 Hz for female speaker AC.

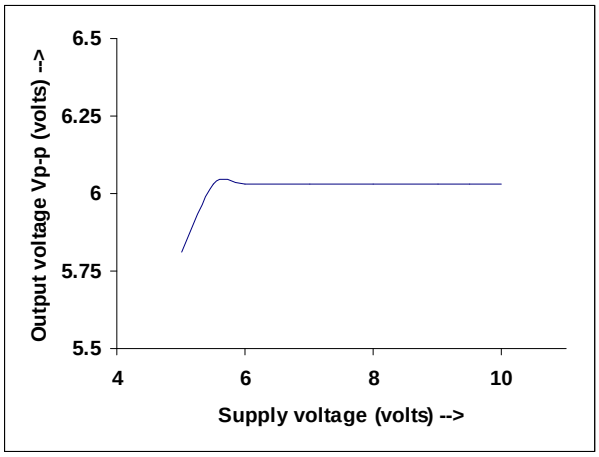


Fig 5.1. Oscillator output vs supply voltage

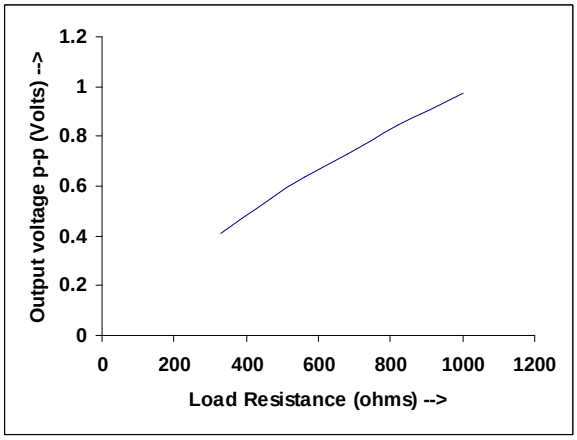


Fig 5.2. Output of the voltage-to-current converter vs load resistance

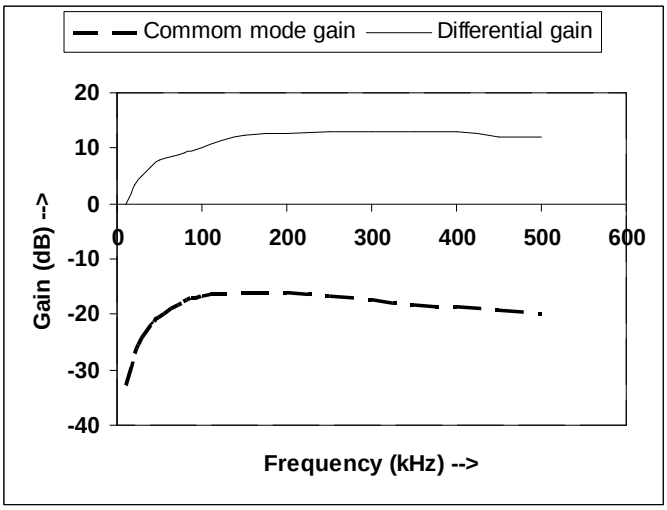


Fig 5.3. Common mode and differential mode gains of the instrumentation amplifier

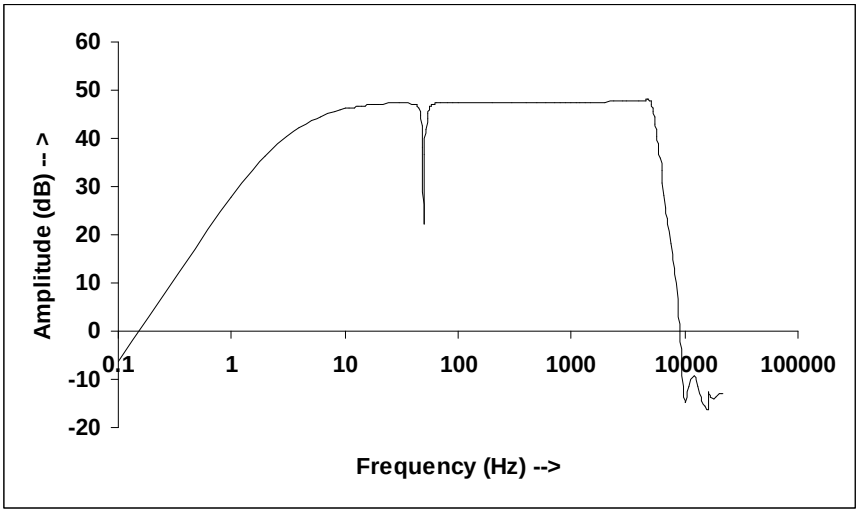


Fig 5.4. Magnitude response of the demodulation filter

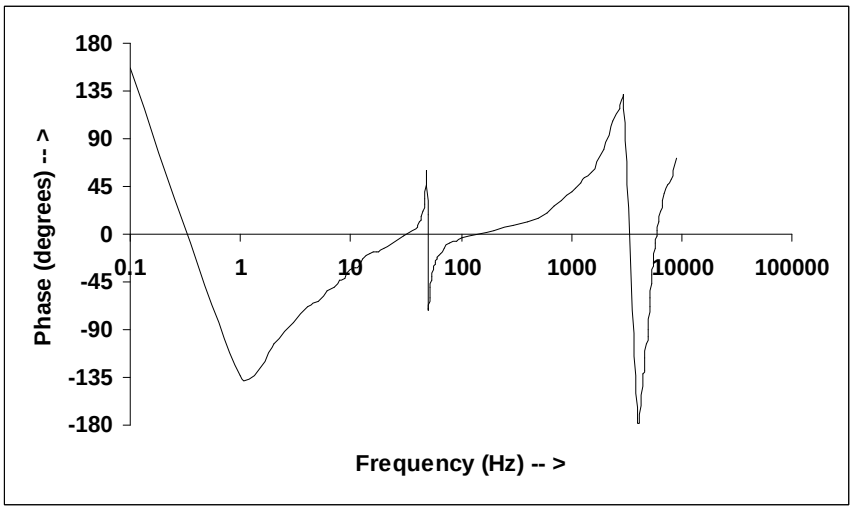


Fig 5.5. Phase response of the demodulation filter

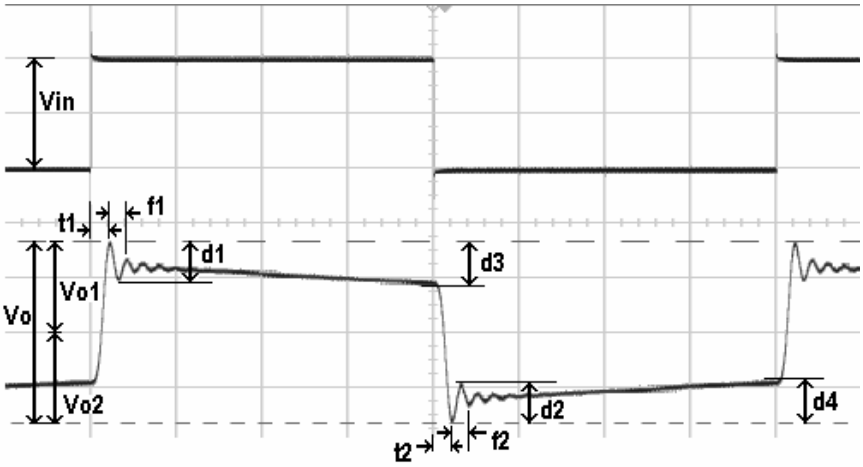


Fig 5.6. Square wave response of the demodulation filter

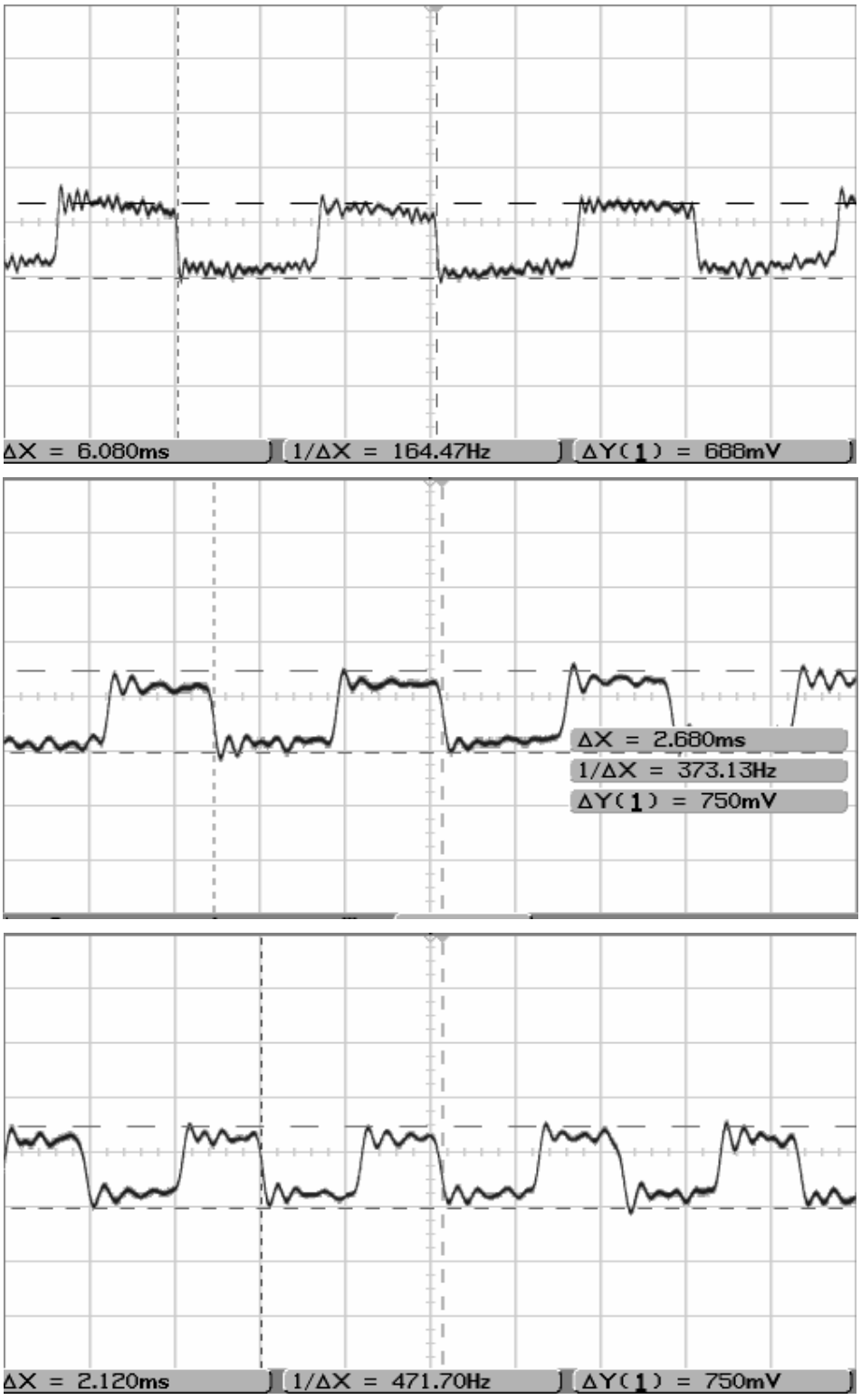


Fig 5.7. Simulator response of the instrument at 164 Hz, 373 Hz, and 471 Hz.

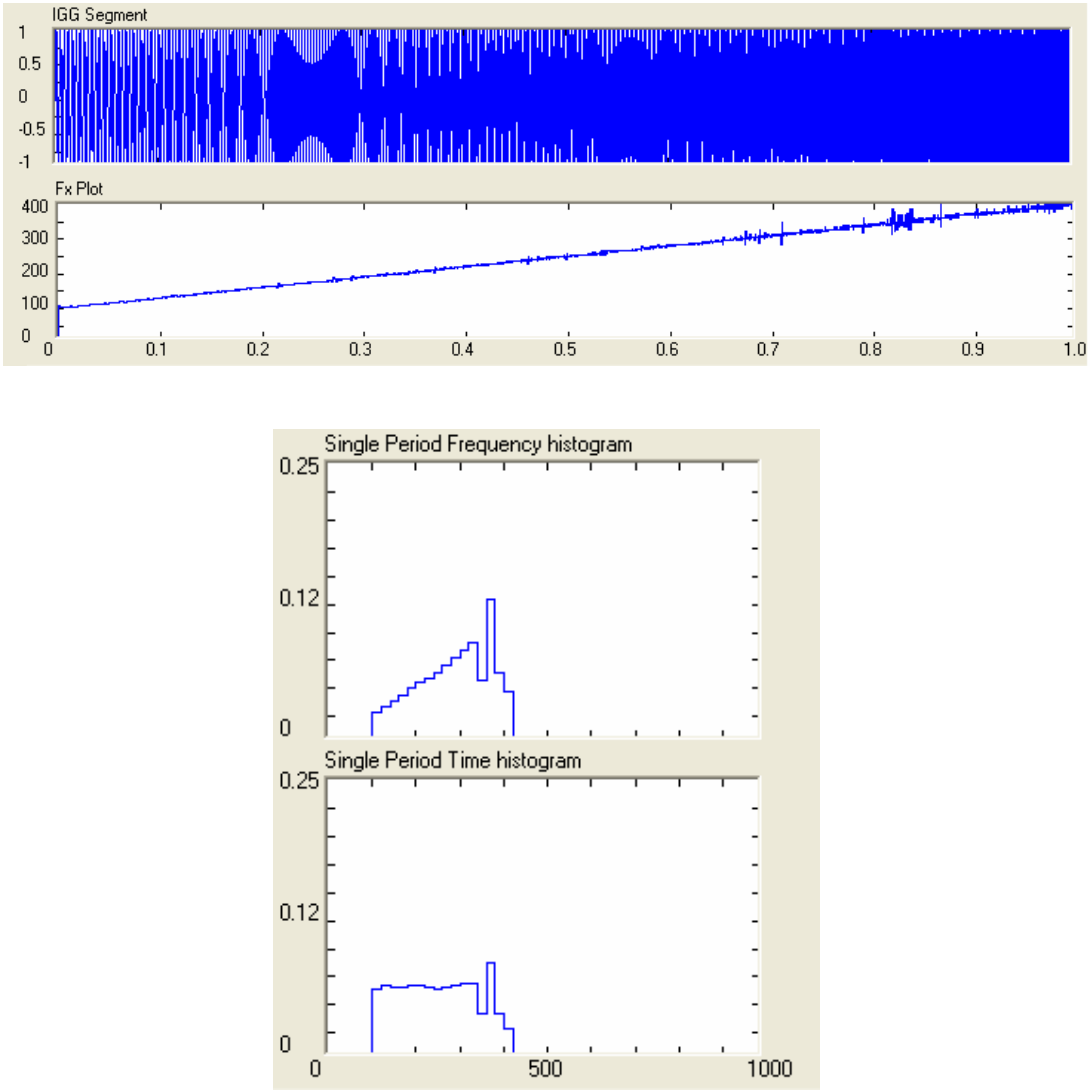


Fig 5.8. Plot of IGG, Fx plot and single period time and frequency histogram of sine wave swept linearly from 100 Hz to 400 Hz over 3 seconds

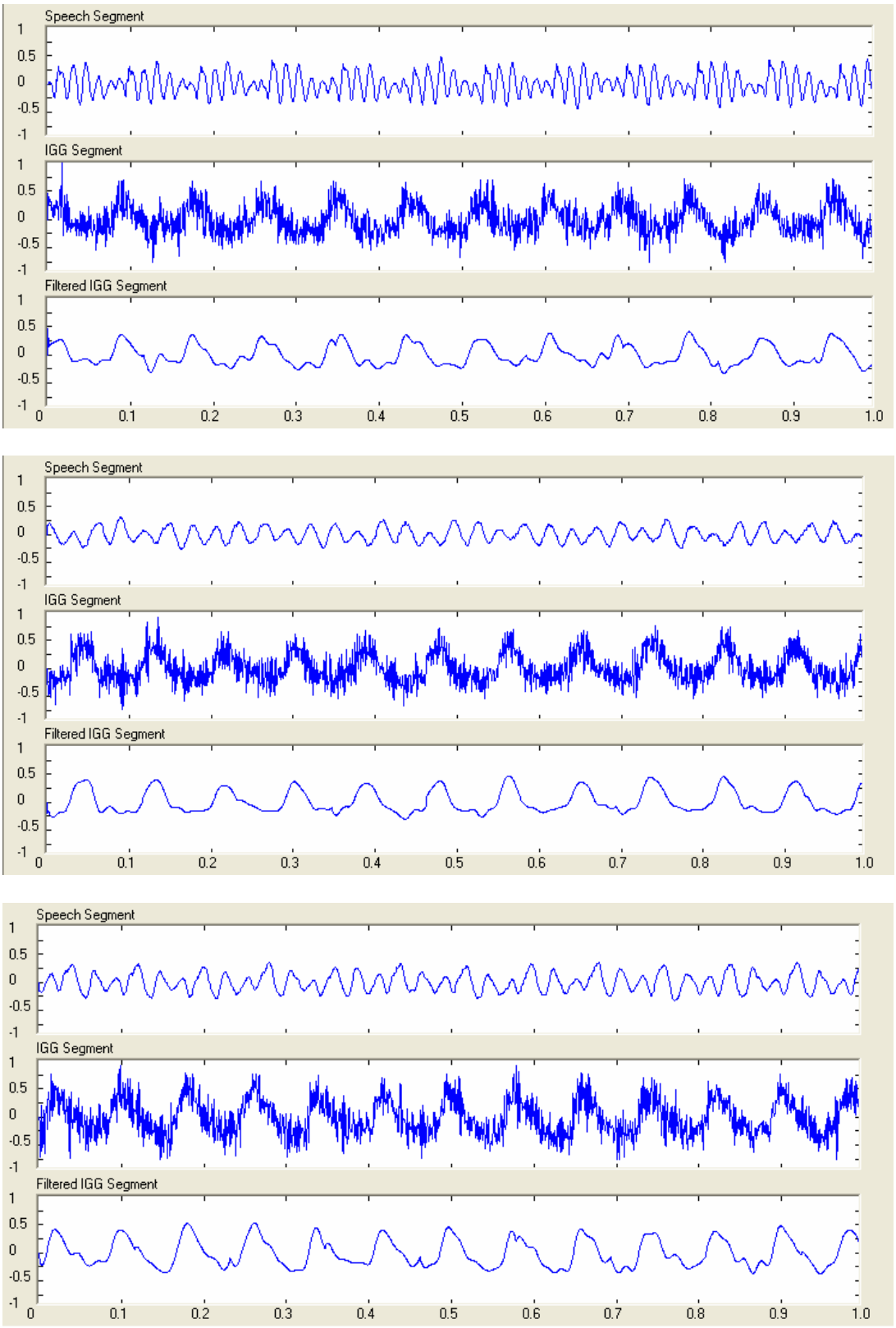


Fig 5.9. Plot of speech, IGG, filtered IGG segment of male speaker AL for sustained vowel /*a*/, /i/, /u/. Segment duration 100 ms. Vertical axis is normalized over ± 1.

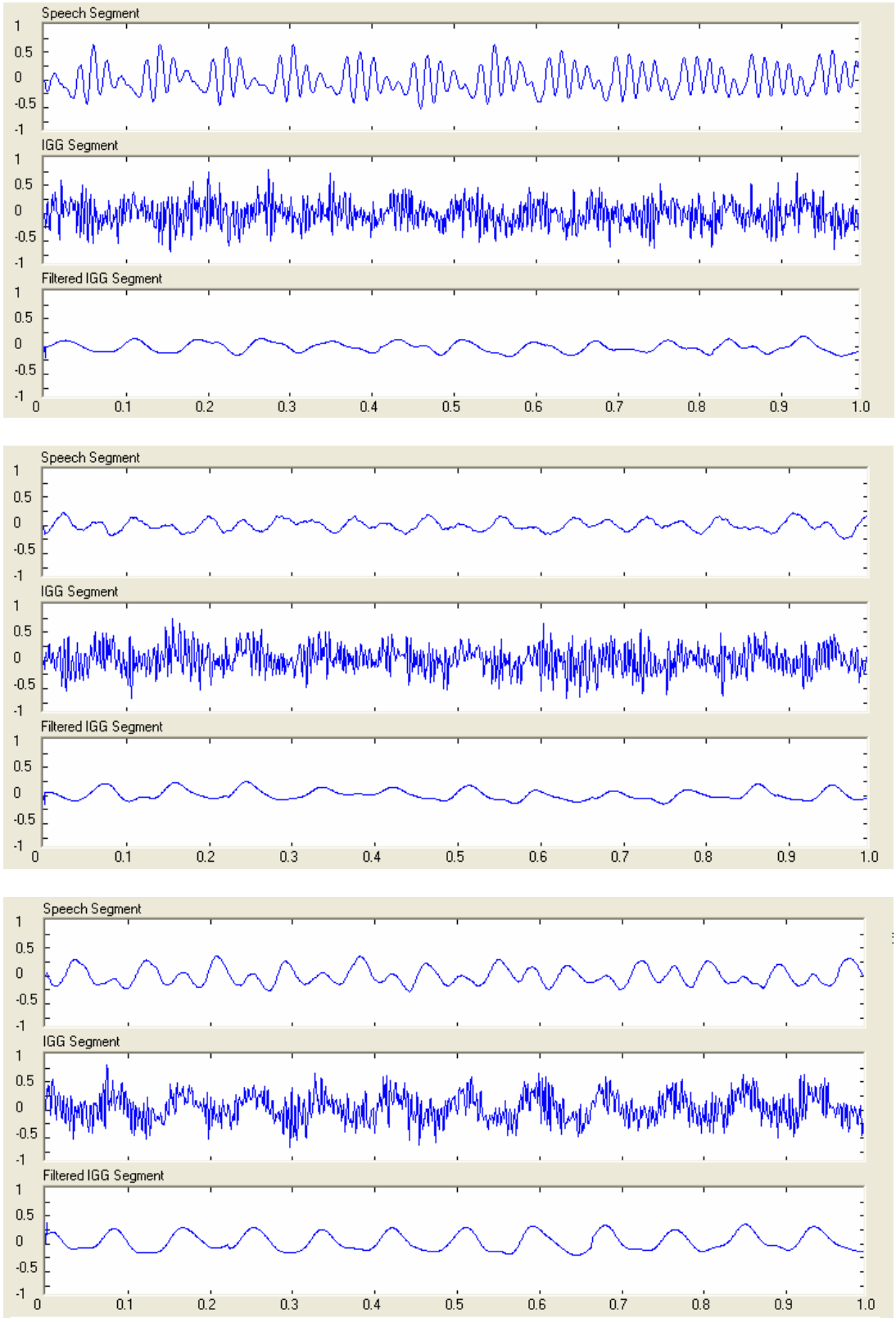


Fig 5.10. Plot of speech, IGG, filtered IGG segment of female speaker AC for sustained vowel /*a*/, /i/, /u/. Segment duration 50 ms. Vertical axis is normalized over ± 1.

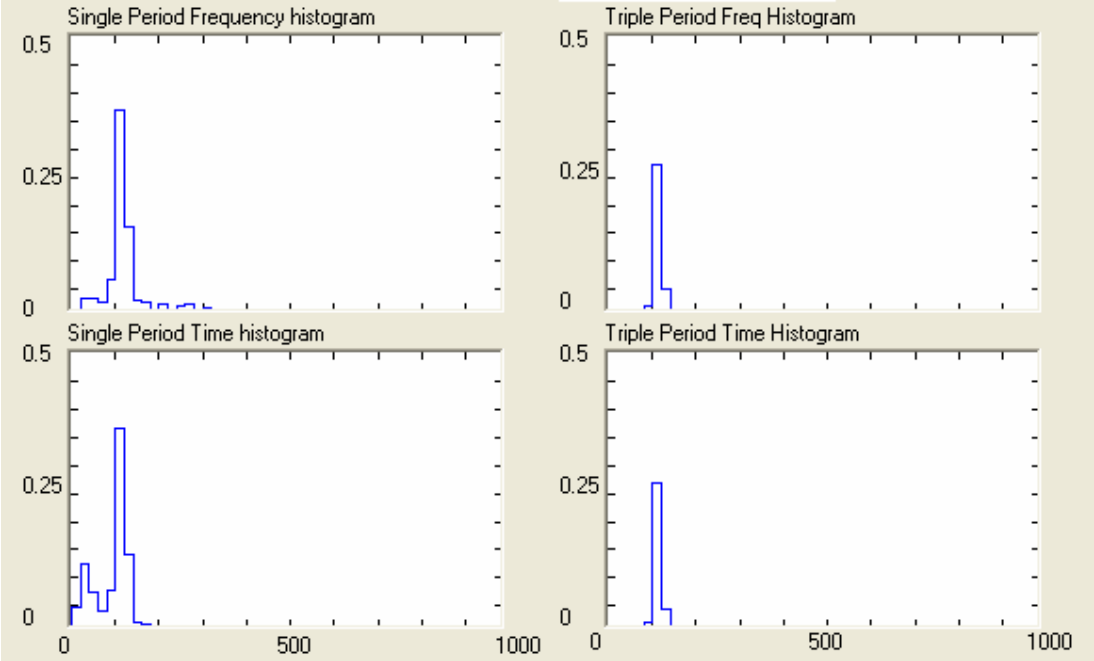


Fig 5.11. Plot of single period and triple period histograms
of male speaker AL reading continous text for 5 s

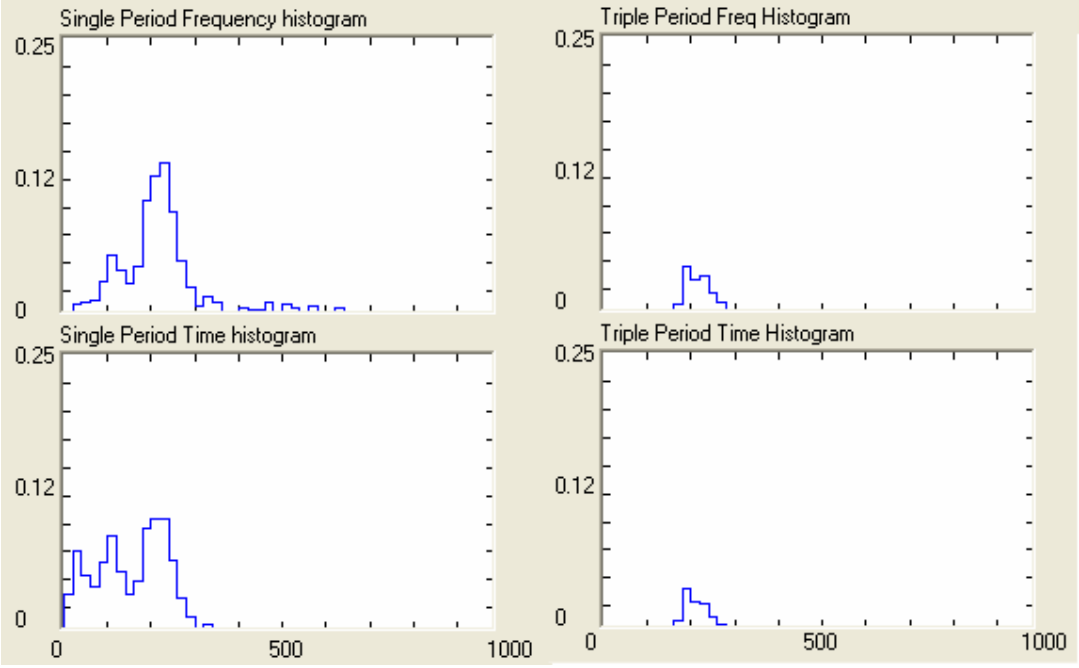


Fig 5.12. Plot of single period and triple period histograms
of female speaker AC reading continous text for 5 s

Chapter 6

SUMMARY AND CONCLUSIONS

The aim of this project was to (i) develop an impedance glottograph, using earlier developed circuits, by increasing the sensitivity of the instrument, keeping the noise level low, and increasing the bandwidth to 5 kHz for more faithful acquisition of the waveform, (ii) study the electrode configurations thoroughly to select the most appropriate one, and (iii) develop a Windows based software for processing and analyzing the output waveform.

As part of the circuit development, the high impedance indicator module, required to indicate proper contact at the skin electrode interface, has been redesigned. The demodulator filter module has also been redesigned replacing the earlier second order Butterworth low-pass and high-pass filters by cascade of notch filter, single order passive high-pass filter, and fifth order elliptic state variable low-pass filter. The cut off frequency of high-pass filter is 2 Hz and that of low-pass filter is 5 kHz. The bandwidth of the circuit has, thus, been increased from 50 Hz – 2.4 kHz to 2 Hz – 5 kHz, thus improving the acquisition of waveform. The measured roll off of low-pass filter is ~ 45 dB / octave. Notch filter is used to reduce the 50 Hz pickup. The CMRR of the instrumentation amplifier (to reduce common mode pickup) is ~ 32 dB.

The modules designed and developed have been integrated on a PCB, designed with special consideration to reduce the noise pick-up from signal lines and from power supply lines by providing tight coupling between the supply lines and the ground, and shielding of the signal lines with ground.

The impedance glottogram output from the instrument can be digitally acquired through the line input of the PC sound card. For acquisition of the waveform and its analysis, a Windows based menu driven signal acquisition, display and analysis software, “siggada” has been developed. This software can be used for acquisition and analysis of speech and IGG signals using stereo line in of the sound card. The operation provided by the software include (i) selection, display, and playback of speech and IGG waveforms,

(ii) band-pass filtering and display of IGG segment, (iii) glottal pulse extraction and display, (iv) Fx plot, (v) histogram plots.

IGG recordings were taken from two normal subjects, a male and a female, using vowel /*a*/, /i/, /u/ and continuous text. The waveform and analysis results conformed to the expected wave shapes and values.

To make the instrument compact, a single battery based power pack needs to be used. Short time stability of the oscillator needs to be improved to improve SNR in the demodulator output. Study of electrode configurations also needs to be undertaken to select the most appropriate one.

It is suggested that tests should be carried out on more number of subjects and also tests are carried out on some subjects with various voice disorders.

Appendix A

INSTRUMENT SPECIFICATIONS

A1 Electrodes

Electrodes	: Neck Electrodes
Electrode area	: Circular gold plated glass epoxy PCB, two sizes - 33 mm in diameter and 38 mm in diameter
Electrode distance	: Adjustable

A2 Electrical characteristics

Carrier frequency	: 400 kHz
Current injected	: ~ 1 mA p-p
Voltage across electrodes	: ~ 1V
Bandwidth for impedance detection	: 80 Hz to 5 kHz
Output voltage	: 1 Vp-p
Out of band rejection by the filter	: 55 dB
Impedance variation sensitivity	: 1V/ohm over basal impedance of ~ 500 ohms
Signal to noise ratio in sensed voltage	: ~ 4 dB
Signal to noise ratio in processed voltage	: ~ 7 dB
Supply voltage	: ± 6 V to ± 10 V
Current drain	: ~ 80 mA from each supply
Power source	: Two Nickel-Hydride batteries. 9V/ 120 mAh
Dimensions of the PCB	: 105 mm × 140 mm

Appendix B

COMPONENT LIST

B.1 Electronic components :

Reference / Designator	Part No. / Value	Part description	Qty.	Appx. rate (Rs)	Appx. Cost (Rs)
C1, C3, C4, C5, C12, C13, C14, C15, C19, C20, C26, C27, C29, C30, C36, C37, C38, C39, C40, C41, C46, C47, C48, C49, C50, C51, C54, C55	0.1 uF	Capacitor (SMD)	28	1.00	28.00
C2, C6, C7, C8, C9, C10, C11,	0.1 uF	Capacitor (Ceramic)	7	0.40	2.80
C16, C17	100 pF	Capacitor (Ceramic)	2	0.40	0.80
C18, C31, C32, C33, C34	1 nF	Capacitor (Ceramic)	5	0.40	2.00
C21, C22	39 pF	Capacitor (Ceramic)	2	0.40	0.80
C23	680 pF	Capacitor (Ceramic)	1	0.40	0.40
C24, C25	3.3 pF	Capacitor (Ceramic)	2	0.40	0.80
C28	560 pF	Capacitor (Ceramic)	1	0.40	0.40
C35	2.2 nF	Capacitor (Ceramic)	1	0.40	0.40
C42, C43	470 nF	Capacitor (Ceramic)	2	0.40	0.80
C45	220 pF	Capacitor (Ceramic)	1	0.40	0.40
C44	1 uF, 63 V	Capacitor (Electrolytic)	1	1.00	1.00
C52, C53	10 uF, 63 V	Capacitor (Electrolytic)	2	1.00	2.00
D1, D2, D3, D4, D5, D7, D8	IN4148	Diode	7	1.00	7.00
D6	3.9 V	Zener diode	1	1.00	1.00
IC1, IC2, IC3, IC6, IC7, IC8, IC11, IC12	LF356N	Op amp	8	18.00	144.00
IC4, IC5	LM311N	Comparator	2	7.00	14.00
IC9, IC10, IC13	TL084P	Quad Op amp	3	10.00	30.00
P1, P4, P5, P6, P8	10k	Trim Potmeters	5	3.00	15.00
P2, P3, P7	50k	Trim Potmeters	3	3.00	9.00
P9	1M	Trim Potmeters	1	3.00	3.00
Q1	BFW10	Transistor	1	27.00	27.00
LED1, LED2	LED5mm	LED	2	1.50	3.00
R1, R8, R11, R28	47k	Resistor	4	0.40	1.60

R2, R4, R14, R15, R18, R30	2.2k	Resistor	6	0.40	2.40
R3, R59	1k	Resistor	2	0.40	0.80
R5, R6	4.7k	Resistor	2	0.40	0.80
R7, R9, R10, R12, R31, R33, R34, R36, R46, R47, R48, R49	100k	Resistor	12	0.40	4.80
R13, R20, R24, R26	22k	Resistor	4	0.40	1.60
R16, R17	3.9k	Resistor	2	0.40	0.80
R19	470k	Resistor	1	0.40	0.40
R21, R23, R25, R54, R57, R58, R73	10k	Resistor	7	0.40	2.80
R22	5.6k	Resistor	1	0.40	0.40
R27, R29, R35	39k	Resistor	3	0.40	1.20
R32, R42, R72	120k	Resistor	3	0.40	1.20
R37, R44	3.3k	Resistor	2	0.40	0.80
R38	2.7k	Resistor	1	0.40	0.40
R39, R43, R45, R51	27k	Resistor	4	0.40	1.60
R40	1M	Resistor	1	0.40	0.40
R41	33k	Resistor	1	0.40	0.40
R50	8.2k	Resistor	1	0.40	0.40
R52, R53	1.5k	Resistor	2	0.40	0.80
R55	68k	Resistor	1	0.40	0.40
R56	47k	Resistor	1	0.40	0.40
R60, R61, R62, R63, R64, R65, R66, R67, R69, R70, R71	6.8k	Resistor	11	0.40	4.40
R68	150k	Resistor	1	0.40	0.40
FB1, FB2, FB3, FB4,	4.3 uH	Ferrite bead	4	6.00	24.00
		PCB	1	250.00	250.00
		Connectors	12	3.00	36.00
		Total			**633.00**

Note : All the resistors are of 5% tolerance.

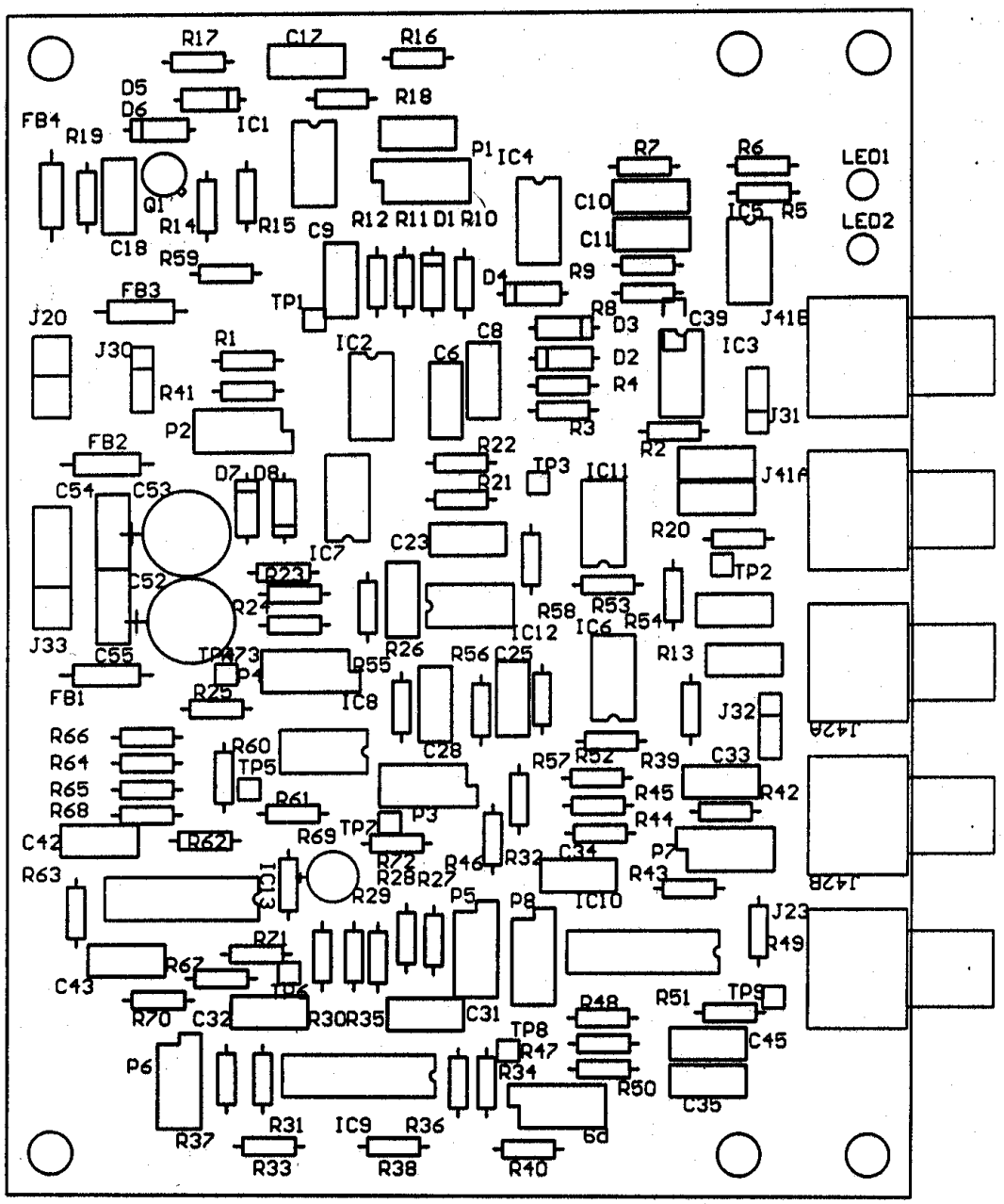
B.2 Other components

Part description	**Qty.**	**Appx. rate (Rs)**	**Appx. Cost (Rs)**
Electrodes	1 pair	300.00	300.00
9V Nickel Hydride battery	2	150.00	300.00
Cabinet	1	200.00	200.00
Total			**800.00**

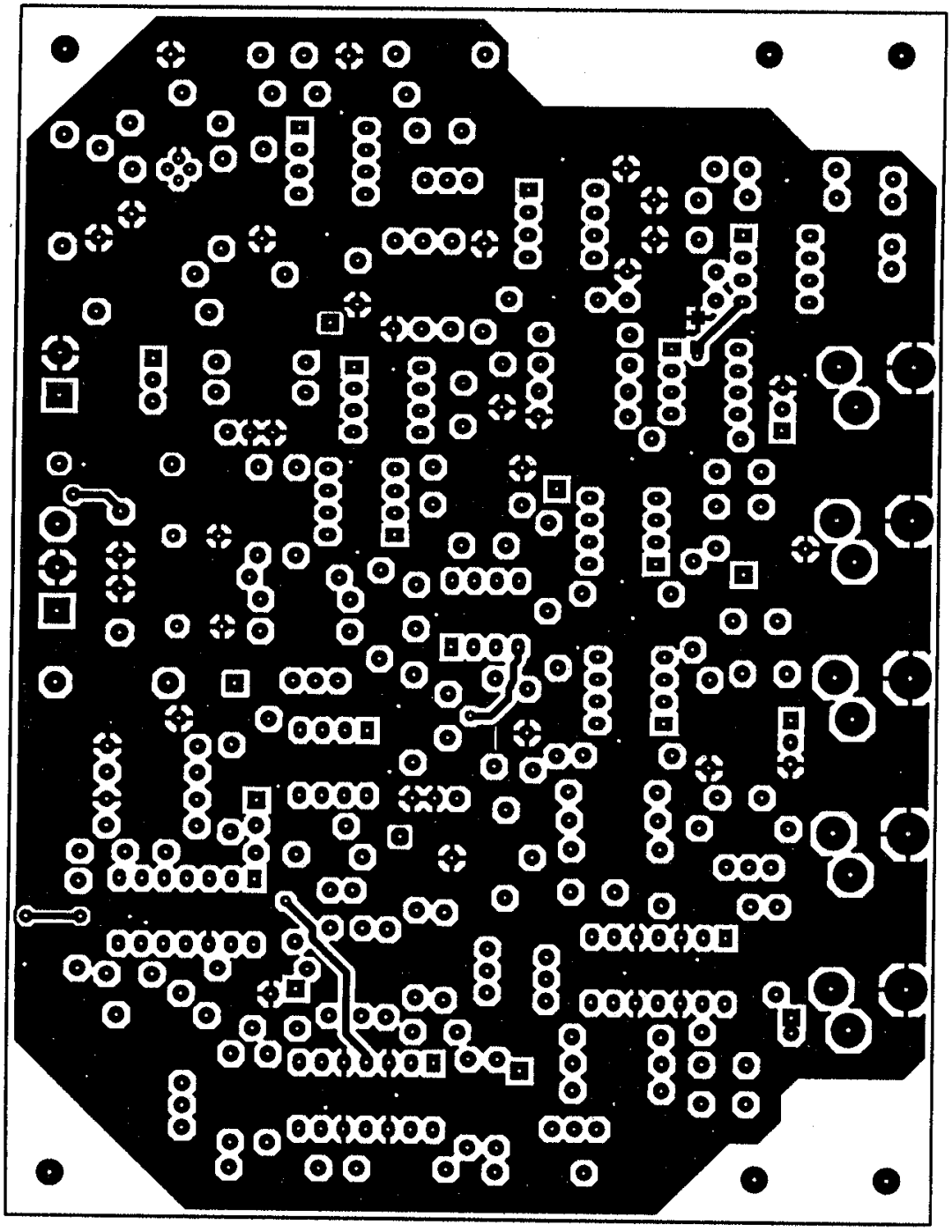
Appendix C

PCB LAYOUTS

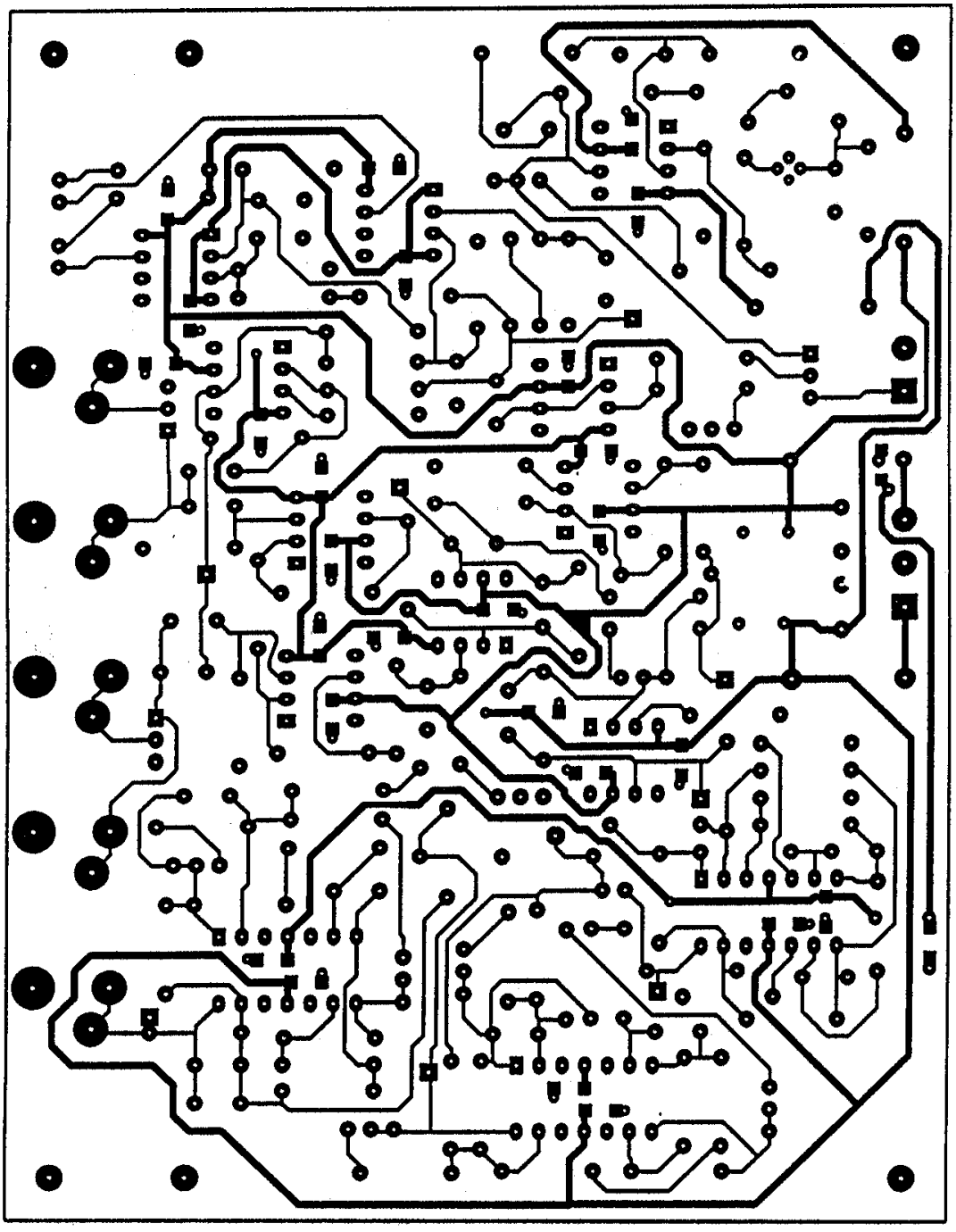
C.1 Component placement layout on PCB



C.2 Component side PCB layout



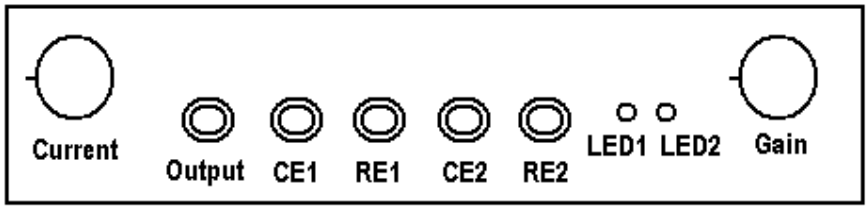
C.3 Solder side PCB layout



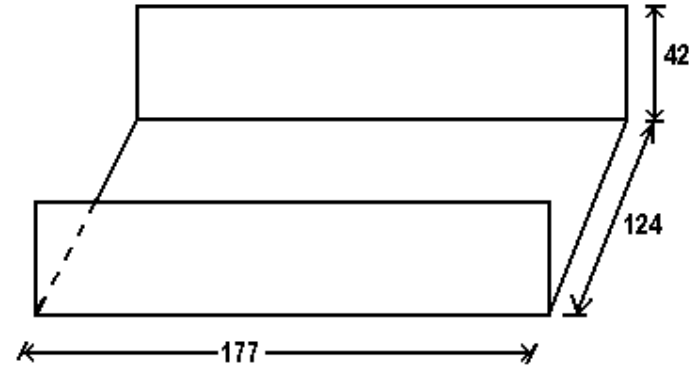
Appendix D

INSTRUMENT CABINET

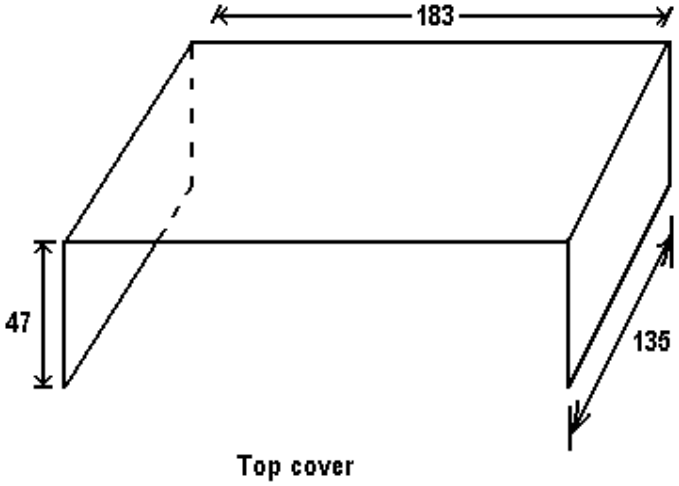
D.1 Front panel of the instrument



D.2 Cabinet dimensions



Bottom cover



Top cover

Appendix E

OPERATION MANUAL FOR SOFTWARE "SIGGADA"

E.1 Introduction

Signal acquisition and analysis program, "siggada", is a Windows based program developed to acquire the signal on two channels simultaneously, one channel for speech and other for impedance glottogram. The signal is acquired on stereo line input of the PC sound card. The acquired signal is saved as a ".wav" file. Signal acquired using some other package, like "Goldwave", in 16-bit signed integer values and saved as ".wav" file, can also be analysed using this program. The display form used to display the acquired signal and its analysis results is shown as Fig E.1.

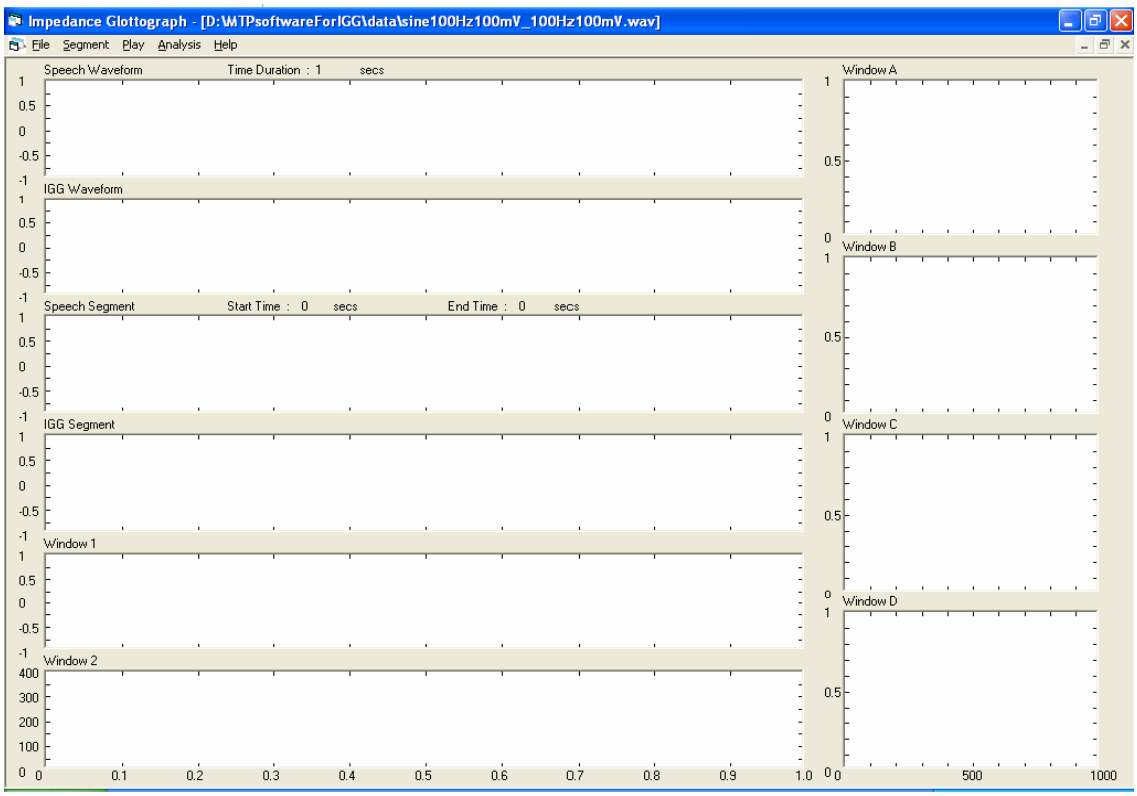


Fig E.1 Display form

The display form has ten windows, arranged in two columns. In the column 1, first two windows are used to display the speech and IGG waveform, next two windows are used to display the selected speech and IGG segment respectively. Last two windows in column 1 are used to display analysis results: filtered IGG waveform, detected glottal

pulses, and Fx plot. Any of the three functions can be displayed in either window. Also same function can be displayed in both the windows for two different segments. The four windows in column 2 are used to display the various histograms. Either of the histograms can be displayed in either of the window. Also four different segments can be compared in four different windows for the same function.

The program is structured as a menu driven application. Fig E.2 lists out all the functions provided and these are discussed in the following sub-sections.

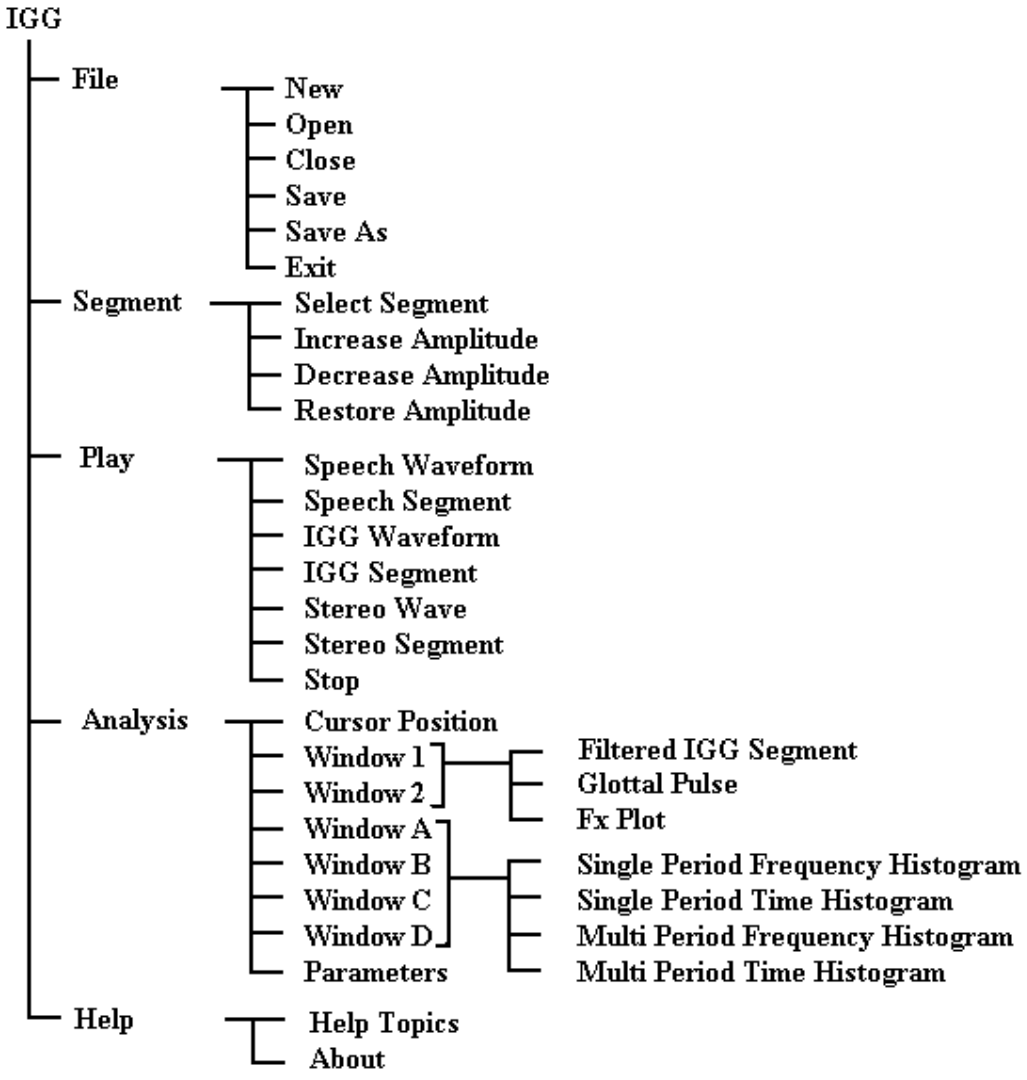


Fig E.2 Menu Tree

E.2 Starting the program

Double clicking the program icon executes the program and the main form of the program appears. The main menus are file, segment, play, analysis and help as can be seen on the main form displayed in Fig. E.3.

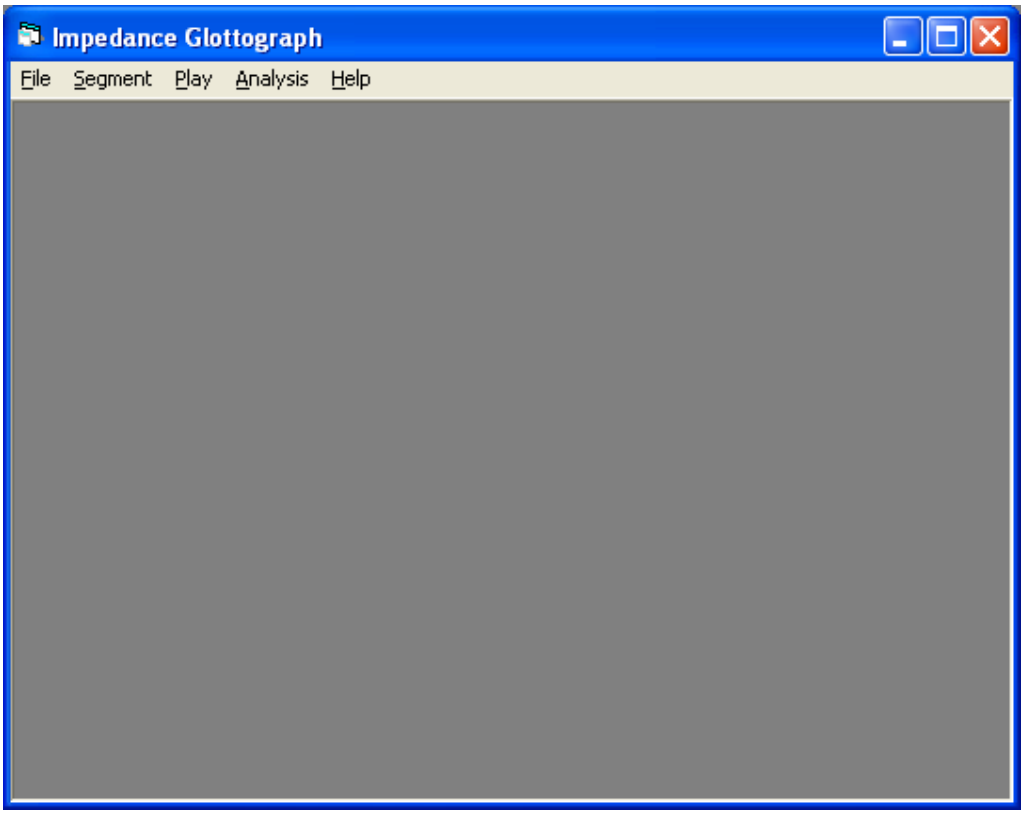


Fig E.3 Main form

E.3 "File" menu

The "File" menu as shown in Fig E.4, allows a number of functions incorporated as sub menus. At the onset, only relevant functions are enabled. Other functions are enabled at the appropriate time.

Fig E.4 "File" menu

E.3.1 "New" sub menu

On clicking the "New" sub menu, another dialog box will be displayed as shown in Fig E.5 This dialog box presents the options to select number of channels, sampling rate, recording duration.

Number of Channels : Mono / Stereo
Sampling Rate (Sa/s) : 8000, 10000, 11025, 16000, 20000, 22050, 32000, 44100.
Recording Duration : 1 s — 15 mins 59 s, selectable in 1 s steps.

On pressing the "OK" button, recording starts. A progress bar is displayed for the duration of recording. When the duration of recording is over, the "Save As" sub menu, under the file menu, is enabled to enable saving the recording.

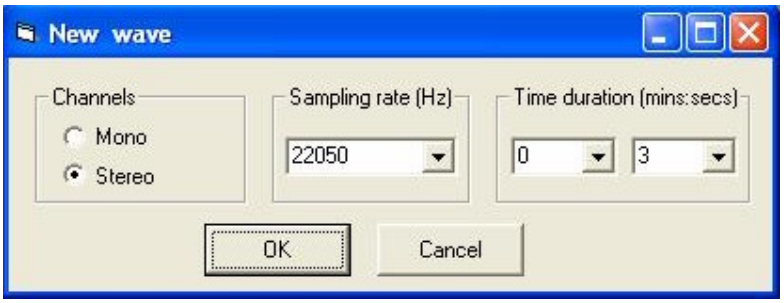


Fig E.5 "New wave" form

E.3.2 "Open" sub menu

This is used to open an already saved file. Clicking this sub menu, an open dialog box appears to enable us to choose a file to open. Upon choosing a file, the signals in the file are displayed in the "speech" and "IGG" waveform windows of the display form, as shown earlier in Fig E.1. The name of the file appears as the caption of the form. Only ".wav" files can be opened by this package.

E.3.3 "Close" sub menu

"Close" sub menu is used to close an already open file. The form displaying the waveform is closed and the main form is displayed with relevant functions enabled.

E.3.4 "Save" / "Save As" sub menu

As already brought out in sub section E.3.1, after a new file is recorded, "Save As" sub menu is used to save the file. On clicking this sub menu, a "Save As" dialog box appears to choose the directory, drive, and the file name. "Save" sub menu also functions the same way.

E.3.5 "Exit" sub menu

This is used to exit the signal acquisition and analysis program.

E.4 "Segment" menu

This menu has four sub menus as shown in Fig E.6. When a file is opened, "Select Segment" sub menu is enabled. "Increase Amplitude" and "Decrease Amplitude" sub menus are enabled once the segment is selected. "Restore Amplitude" sub menu is enabled when the amplitude of any segment is increased / decreased and is disabled when both the segments are of original amplitude.

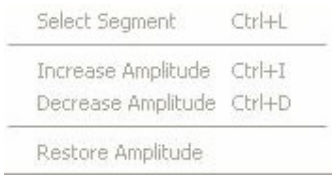


Fig E.6 "Segment" Menu

E.4.1 "Select Segment" sub menu

This sub menu is used to select and display a portion of the waveform. The segment can be selected by left clicking in the speech waveform window for placing the start and end markers. As the mouse pointer is moved in the speech waveform window, the X and Y axis positions are displayed. First left click will position the start marker displayed green in colour and second left click will position the end marker displayed red in colour. Portion of the waveform between the two markers will be displayed as speech segment and IGG segment in two different windows. Start time and end time of the segment selected is also displayed.

E.4.2 "Increase Amplitude" / "Decrease Amplitude" / "Restore Amplitude" sub menu

These sub menus can be used to appropriately scale the amplitude of the speech segment and IGG segments independently. On clicking "Increase Amplitude" sub menu, the sub menu will be checked and the mouse pointer will be changed when mouse is moved over the window displaying speech segment or IGG segment. On each left clicking in the window, the respective segment amplitude will be increased by 10% and waveform will be replotted. Also restore amplitude sub menu will be enabled. "Decrease Amplitude" sub menu can be used for decreasing the amplitude in a similar fashion.

"Restore Amplitude" sub menu restores the amplitude of the segment to the original value. On clicking restore amplitude sub menu, the sub menu will be checked. On left clicking in the speech segment or IGG segment window, the respective segment amplitude will be restored to original amplitude and the respective segment will be replotted with the original amplitude.

E.5 "Play" menu

"Play" menu as shown in Fig E.7 allows to play the speech waveform, speech segment, IGG waveform, IGG segment, stereo waveform, stereo segment. A stop sub menu is provided to terminate an ongoing play operation.

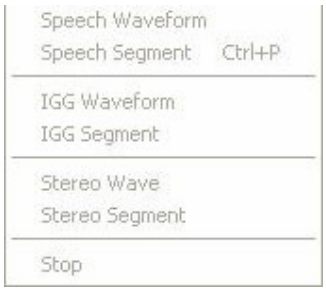


Fig E.7 "Play" Menu

E.6 "Analysis" menu

“Analysis” menu provides a number of functions like glottal pulse, Fx plot, single period time and frequency histograms, multi period time and frequency histograms, cursor position, dialog box to change analysis parameters at run time. Of these, “Glottal pulse” and “Fx plot” are placed under “Window 1” and “Window 2” sub menus as shown in Fig E.9. So either function can be plotted in either window, or same function can be

compared in both windows for two different segments. Similarly, single period and multi period time and frequency histograms are placed under “Window A”, “Window B”, “Window C” and “Window D” sub menus as shown in Fig E.10 and either function can be plotted in either window, or same function can be compared in all four windows for four different segments. “Analysis” menu is shown in Fig E.8.

Fig E.8 "Analysis" menu

E.6.1 "Cursor Position" sub menu

This sub menu provides the cursor measurement to facilitate measurements in any of the displayed windows.

E.6.2 "Window 1", "Window 2" sub menu

Under these sub menus, three sub sub menus, namely “Filtered IGG Segment”, "Glottal Pulse", and "Fx Plot" are provided as shown in Fig E.9. By clicking on either of the sub sub menus, the particular function is plotted in that particular window. For any segment two functions can be plotted in two different windows. Alternatively, for two different segments, same function can be plotted in two different windows for easy comparison.

Fig E.9 "Window1", "Window 2" sub menu

E.6.3 "Window A",… sub menu

Under "Window A", "Window B", "Window C", "Window D" sub menus, four sub sub menus, namely single period frequency histogram, single period time histogram, multi period frequency histogram and multi period time histogram are provided as shown in Fig E.10. By clicking on either of the sub sub menus, the particular function is plotted in that particular window. For any segment, four functions can be plotted in four different windows. Alternatively, for four different segments, same function can be plotted in four different windows for easy comparison.

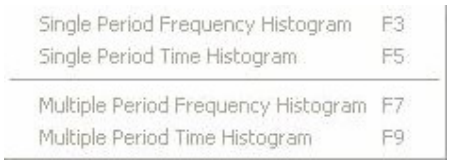


Fig E.10 "Window A", … sub menu

E.6.4 "Filtered IGG Segment", "Glottal Pulse", and "Fx Plot"

Input IGG waveform is digitally filtered to remove offset and noise, and the filtered waveform can be seen as "Filtered IGG Segment". Glottal pulse is the plot of the output of rectangular pulse obtained by processing the filtered IGG segment with a dynamic hysteresis comparator. Fx Plot is the plot of instantaneous pitch values obtained from glottal pulses. It can be considered as tracking of the pitch.

E.6.5 Histogram

Histograms are obtained from the pitch values calculated as in (4.7). Histogram can be plotted on the basis of the number of pitch periods or on the basis of time, and the two can be called as frequency and time histogram respectively. The frequency range from 0 to 1000 Hz is divided in a fixed number of bins (from 10 to 50). Same segment can be analyzed by plotting various histograms with different number of bins. Histograms are obtained by updating the bin counts after each glottal pulse. For frequency histograms, count in a frequency bin is incremented by one when the pitch frequency falls within that bin. For time histogram, the corresponding bin count is incremented by the number of samples in the glottal pulse. At the end of the segment, histogram is obtained

by dividing the bin counts by the sum of all the bin counts. The histograms calculated by the above method are single period histograms. For *m* period histograms, value of a bin is incremented when m successive pitch values fall in the same bin. Normalizing is done by dividing the bin counts by the total number of pitch periods for the frequency histogram and the total number of samples in the segment for the time histogram. This multi period histograms show the stability of pitch frequency.

E.6.6 “Parameters” sub menu

On clicking this sub menu, a dialog box appears as shown in Fig E.11, to set values at run time and compare the outputs. The parameters that are available are alpha, beta, gamma values for the dynamic hysteresis comparator used to obtain glottal pulses from the input IGG. The number of bins for histogram and number of periods for multi period histogram can also be set through this dialog box.

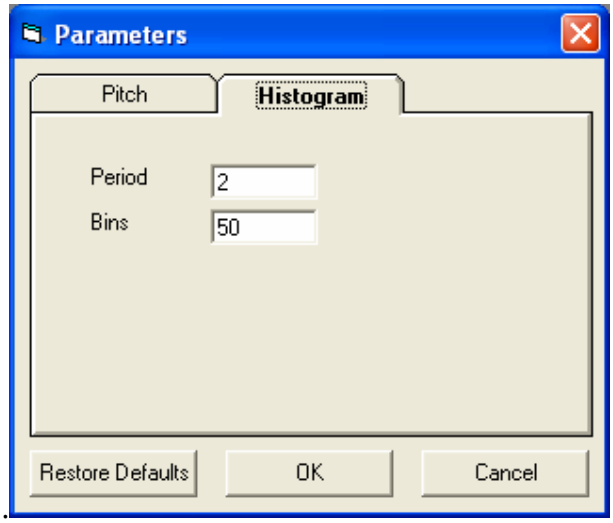


Fig E.11 Parameters form

E.7 “Help” menu

Under the help menu as shown in Fig E.12, “Help Topics” sub menu provides online help on this software, “About” sub menu gives basic information about the software as shown in Fig E.13.

Fig E.12 “Help” menu

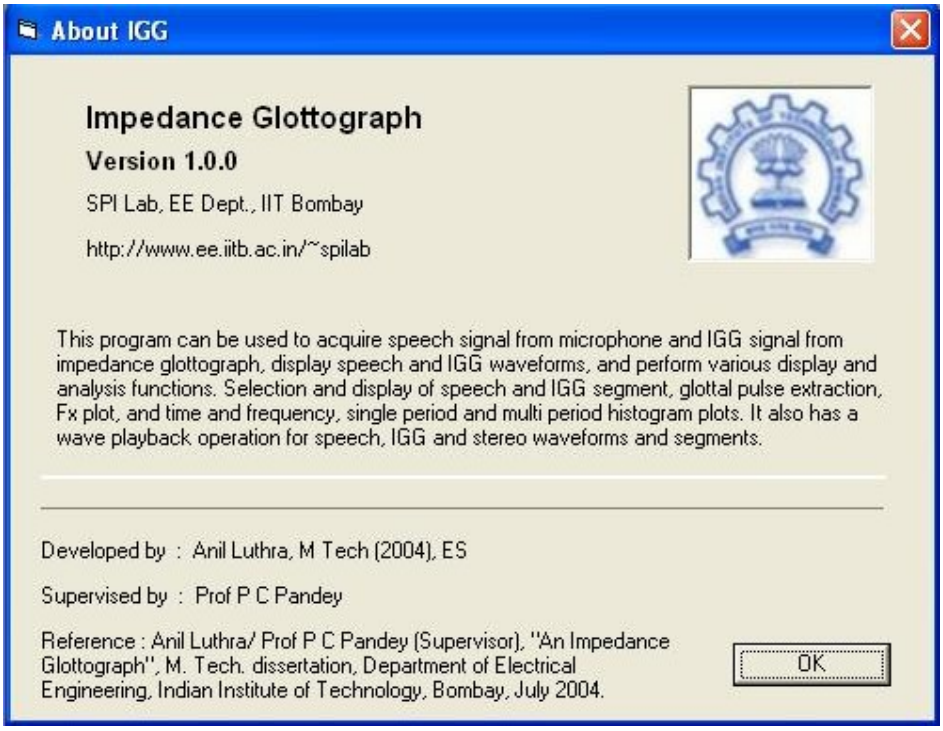


Fig E.13 “About IGG” form

E.8 System specifications

Processor : Minimum :
Intel Pentium Processor P1 133 MHz or equivalent
: Recommended :
Intel Pentium Processor P3 350 MHz or equivalent

Operating System **:** Minimum :
Windows 95
: Recommended :
Windows 98 / Windows 2000 / Windows XP

RAM : 64 MB (For a recording of upto 5 mins)
256 MB (For a recording of upto 15 mins)

Hard disk requirement : 300 MB of free hard disk space.

Sound Card specifications	: Analog stereo line-in with
	Coupling : AC
	Voltage range : ± 1 Vp-p
	Input resistance > 10 kΩ
	Quantization : 16-bit
	Sampling rate (Fs) : 11 — 44 kSa/s
	Bandwidth = 10 Hz to Fs/2.
	Analog stereo out : 4 W per channel for 4 Ω speakers (without external amplifier) or powered speakers (driven by external amplifier)

Appendix F

IMPEDANCE SIMULATOR

F.1 Introduction

For calibration of the impedance glottograph and for testing of various electrode configurations, an impedance simulator is required, that can simulate the impedance across the thyroid cartilage and impedance variation as step changes corresponding to the vocal chord being fully open and vocal chord being fully closed. The impedance across the thyroid cartilage is ~ 500 Ω and the change in impedance due to vocal fold vibration is ~ 1 Ω. The vocal fold vibration is in the range of 100 – 500 Ω for a human subject. An impedance simulator developed by Patil [17], was used for calibration of the impedance glottograph and study of various electrode configurations. The design of the impedance simulator is discussed in the next section.

F.2 Model of the impedance simulator

The impedance simulator should be able to simulate various fixed impedances and variations in the impedances due to vocal fold vibrations. At the frequency of current injection, all the impedances are predominantly resistive. The model of impedance simulator is given in Fig. F.1. R_Z represents the tissue impedance across the thyroid cartilage with vocal folds fully open. When the switch S is closed, R_Y is in parallel with Rz, denoting the condition of vocal folds being fully closed. R_{X1}, R_{X2}, R_{X3}, and R_{X4} represent the resistances across the outer surface of the neck between the electrodes. R_{Z1}, and R_{Z2} represent the resistances from the surface of the neck to the inner tissue. By controlling the switching frequency of switch S, the vocal folds can be simulated to be opening and closing at various frequencies. Switch S is an analog switch controlled by a astable multi vibrator.

When switch S is off,

$$R_{eq}(S_{off}) = R_X \,||\, (R_Z + R_{Z1} + R_{Z2}) \qquad \text{(F.1)}$$

When switch S is on,

$$R_{eq}(S_{on}) = \underline{R_X} \,||\, (R_{Z1} + R_{Z2} + R_Z || \, R_Y) \qquad \text{(F.2)}$$

where $R_X = R_{X1} + R_{X2} + R_{X3} + R_{X4}$

The values of resistances are so chosen that $R_{eq}(S_{off})$ = 591.1 Ω and $R_{eq}(S_{on})$ = 590.1 Ω. The frequency of switching is made variable from 100 Hz to 500 Hz.

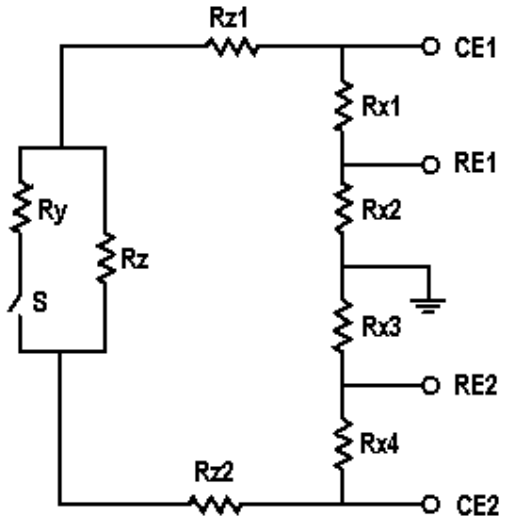


Fig F.1 Model of impedance simulator

Appendix G

COMPARISON OF VARIOUS PRODUCTS

		F J Electronics (EG 90)	Glottal Enterprises (EG 2 PC)	Instrument developed at IIT Bombay
1.	Electrodes	Strapped or hand held electrodes	Strapped electrodes	Strapped electrodes
2.	Electrode area	24 mm in diameter.	34 mm in diameter.	33 mm and 38 mm in diameter.
3.	Electrode material used	Bulging circular stainless steel	Gold plated.	Gold plated
4.	Electrode current	4.2 mA		1 mAp-p
5.	Material used to hold electrodes	Velcro strap.	Velcro strap.	Velcro strap.
6.	Power source	Alkaline battery 9 V, size PP3, 550 mAh.	Standard dual 6 Volt lead-acid gel cell batteries.	Two nickel hydride batteries. 9V / 120 mAh.
7.	BW for impedance detection (frequeny response)	3.5 Hz - 10 kHz.	5 Hz - 6 kHz.	2 Hz - 5 kHz.
8.	Carrier frequency	400 kHz	2 MHz	400 kHz
9.	Impedance variation sensitivity	2 V/ohm		1 V/ohm
10.	Total current drain	9.5 mA (?)	160 mA	158 mA
11.	Dimensions (in mm)	120 * 60 * 30	200 * 266 * 87	180 * 135 * 45

REFERENCES

[1] D. O. Shaughnessy, *Speech Communication Human and Machine*, New York : Addison-Wisley, 1987.

[2] A. J. Fourcin and E. Abberton, “First applications of a new laryngograph”, *Medical and Biological Illustration*, Vol 21, pp 172-182, July 1971.

[3] M Hirano, *Clinical Examination of Voice, Disorders of Human Communication*, Austria : Spring-Verlag Wien New York, 1981.

[4] P.K. Lehana and P.C. Pandey, "A low cost impedance glottograph and glottal pitch analyzer," *Proc. Bio Vision 2001 International Conference on Biomedical Engineering*, (Bangalore, India), pp 33-37, December 2001.

[5] T. Baer, A Lofqvist and N.S. McGarr, “Laryngeal vibrations : A comparison between high speed filming and glottographic techniques”, *J. Accoust. Soc. Am.*, Vol. 73, No. 4, pp 1304-1308, April 1983.

[6] “Electroglottograph”, http://www.drspeech.com/Electroglottograph.html, site of Dr. Speech Software Group, downloaded on 14 Mar 03.

[7] A.J. Fourcin, “Laryngographic examination of vocal fold vibrations”, *Ventillatory and Phonatory Control Systems, An International Symposium*, New York : Oxford University Press, 1974.

[8] K. Marasek, “Electroglottography”, http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee1.htm, site of Institute of Natural Language Processing, University of Stuttgart, Germany, Jul 2001, downloaded on 16 Mar 03.

[9] “Laryngograph Processors”, http://www.laryngograph.com/products/lxproc.htm, site of Laryngograph Ltd, downloaded on 11 Mar 03.

[10] “Portable Electroglottograph”, http://www.f-jelectronics.dk/htm/electroglotto.htm, site of F-J Electronics, downloaded on 11 Mar 03.

[11] “Electroglottographs”, http://www.glottal.com/electroglottograph.html, site of Glottal enterprises, downloaded on 11 Mar 03.

[12] R. Bhagwat, “Glottal pitch extractor”, *B Tech Project Report*, Supervisor: P.C. Pandey, IIT Bombay, 1990.

[13] A. Sriram, “Glottal pitch extractor”, *B Tech Project Report*, Supervisor: P.C. Pandey, IIT Bombay, 1991.

[14] A.V.I. Thajudin, “Glottal pitch extractor”, *M Tech Mini Project Report*, Supervisor: P.C. Pandey, IIT Bombay, 1994.

[15] K.K. Mahajan, “Glottal pitch extractor”, *M Tech Mini Project Report*, Supervisor: P.C. Pandey, IIT Bombay, 1995.

[16] M.S. Chitnis, "A glottal pitch extractor", *M Tech Dissertation*, Supervisor: P.C. Pandey, Dept. of Electrical Engineering, IIT Bombay, June 1998.

[17] A.V. Patil, "A glottal pitch extractor", *M Tech Dissertation*, Supervisor: P.C. Pandey, Dept. of Electrical Engineering, IIT Bombay, January 2000.

[18] A.C. Guyton, *Textbook of Medical Physiology*, Philadelphia : Saunders, 1986.

[19] A.K. Krishnamurthy and D.C. Childers, “Two channel speech analysis”, *IEEE Trans. Acoustic, Speech, Signal Processing*, vol. ASSP-34, pp. 730-743, August 1986.

[20] K. Marasek, “Description of the EGG waveform”, http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee2.htm, site of Institute of Natural Language Processing, University of Stuttgart, Germany, Jul 2001, downloaded on 16 Mar 03.

[21] D.G. Childers and J.N. Larar, “Electroglottography for laryngeal function assessment and speech analysis”, *IEEE Trans. Biomedical Engineering*, Vol BME-31, No. 12, pp 807-817, December 1984.

[22] D.G. Childers, D.M. Hicks, G.P. Moore, and Y.A. Alsaka, “A model for vocal fold vibratory motion, contact area, and the electroglottogram”, *J. Accoust. Soc. Am.*, Vol. 80, No. 5, pp 1309-1320, November 1986.

[23] K. Marasek, “The EGG waveform and voice qualities”, http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee7.htm, site of Institute of Natural Language Processing, University of Stuttgart, Germany, Jul 2001, downloaded on 17 Mar 03.

[24] A.J. Fourcin, Laryngographic Assessment of Phonatory Function, *ASHA reports*, No 11, pp 116-127, 1981.

[25] A.J. Fourcin, “Appratus for speech pattern derivation”, *US Patent*, No.4139732 dated February 13, 1979.

[26] M. Rothenberg, “Tracking multielectrode electroglottograph”, *US Patent*, No. 4909261 dated March 20, 1990.

[27] F.R. Dungan, *Op Amps and Linear Integrated Circuits for Technicians*, New York : Delmar, 1992.

[28] J.M. Fiore, *Op Amps and Linear Integrated Circuits*, New York : Delmar, 2001.

[29] A.B. Williams, *Electronic Filter Design Handbook*, New York : Mc Graw Hill, 1981.