A GLOTTAL PITCH EXTRACTOR

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

by

Mandar S. Chitnis
96307602

Guide

Dr. P.C. Pandey

Department of Electrical Engineering

Indian Institute of Technology, Bombay

June 1998

Mandar S. Chitnis : A Glottal Pitch Extractor, M.Tech. Dissertation, Department of Electrical Engineering, I.I.T. Bombay, June 1998. Guide : Dr. P. C. Pandey

ABSTRACT

Detection and estimation of voice pitch, the fundamental frequency of vocal fold vibrations, finds wide use in speech analysis and recognition, diagnosis of voice disorders, and speech training aids for the hearing impaired. The aim of this project is to develop a low cost battery operated instrument using electroglottography, a non-invasive technique for measuring impedance variation across the thyroid cartilage of the larynx. This impedance variation provides information about the dynamics of the closure of vocal folds, and can be used for obtaining the voice pitch.

A high frequency (300 kHz), low intensity (~3 mA) current is passed through the central discs (15 mm dia.) of a pair of plate electrodes held in contact with the skin on both sides of the thyroid cartilage. A guard ring around each of the discs is actively driven to the same potential as the central disc, in order to reduce the superficial component of the sensing current. The impedance variations, caused by varying contact area between the vocal folds, result in amplitude modulated voltage waveform across the central disc electrodes. This waveform is demodulated to get the impedance variation. A low cost microcontroller based signal acquisition, analysis, and LCD graphics display unit has been developed as a part of this instrument for displaying the impedance variation waveforms for diagnosis of speech disorders. A serial interface has been provided for interfacing the instrument to a computer for downloading the measurement results and to take a plot of the stored waveform on a plotter. A signal acquisition and analysis software has been developed on a notebook PC with a data acquisition card, for capturing the impedance variation waveform and plotting the pitch histogram.

ACKNOWLEDGMENT

I wish to express my sincere gratitude to my guide, Dr. P.C. Pandey for his valuable guidance and continuous encouragement during this project work. I take this opportunity to thank ENT specialists Dr. Pramod Kothari and Dr. Sanjay Helale for all the support given by them.

I thank Shri Apte and Shri Muntode for their help during my project work. I would also like to thank all of my friends who encouraged me and gave me lot of support for the completion of this project.

IIT Bombay
June 1998

Mandar S. Chitnis
(96307602)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
INTRODUCTION

1.1 OVERVIEW

Speech communication is the transfer of information from one person to another via speech, which is composed of a sequence of sounds produced by the vocal apparatus. The speech organs can be divided into three main groups: lungs, larynx, and the vocal tract. The lungs are the source of air flow that passes through the larynx and the vocal tract before exiting the mouth as pressure variations constituting the speech signal. The vocal tract, which is the region above the larynx, is effectively an acoustic enclosure with several resonance cavities and acts as an acoustic filter shaping the spectrum of the generated sound. The source of most speech occurs in the larynx where the vocal folds (also known as vocal cords), can partially or completely obstruct air flow from the lungs. The vocal cords are elastic structures of muscles, and mucous membrane. Due to the pressure variations, these vocal folds are set into vibrations and the fundamental frequency of these vocal fold vibrations is known as the pitch. The resonance frequencies of the vocal tract are called formants and depend upon the shape and dimensions of the vocal tract cavities which can be changed by the movement of the articulators (velum, tongue, teeth, lips, and jaw). Different speech sounds are produced due to the changes in vocal tract configuration and its mode of excitation [25], [28].

Pitch detection and estimation finds wide use in speech analysis and recognition, for diagnosis of voice disorders, in speech training aids for the hearing impaired, and in sensory aids as compliment to lip-reading. The profoundly deaf persons, unable to make use of hearing aid, make use of lip-reading for everyday communication. Lip-reading can be made much more efficient by providing the pitch of the sound externally. This is because, words such as 'bet' and 'pet' which have an identical lip movements, and as such are indistinguishable to a deaf lip-reader, are differentiated by the vocal fold activity. There are various ways of presenting the pitch information [27].

Pitch can be estimated using speech processing methods, where the speech is captured using a microphone and then it is analyzed to find out the pitch. However, this method involves lot of computations because the speech captured using this technique includes the effect of resonance due to nasal, paranasal, and oral cavities.

During phonation, the vocal folds are set into vibration. The movement of the vocal folds consists of three phases; contact, separation, and open phase. The variation in electrical impedance between the vocal folds gives information about the different phases between them. This information about the actual contact is very useful, since the nature of contact depends upon the physical condition of the vocal folds.

Electroglottography, uses the measurement of impedance variations in the vicinity of the glottis, for sensing the vocal fold contact. The pitch can be obtained by measuring the time period of the impedance variation waveform [5]. The glottal pitch extractor measures the electrical impedance variations of the larynx using a pair of electrodes held in contact with the skin on both sides of the thyroid cartilage. The base impedance across the thyroid cartilage is approximately 500 Ω and the change in the impedance variations is less than 1 Ω [5]. The usable frequency range of the current carrier is 100 kHz - 5 MHz [5], [9], [10], [17]. The voltage developed across the electrodes gets amplitude modulated due to the variations in the impedance in the current path, caused by the vibrations of the vocal folds. The modulation depends upon the change in the tissue impedance in the current path. The impedance is minimum when the vocal folds are in full contact, the effective tissue length for the RF current, is smallest between the electrodes in this vocal fold configuration. The impedance increases as the folds separate and it is at its maximum when the folds are completely separated. The impedance level does not show any considerable change, even if the folds are wider apart [5].

The principle of impedance variation for pitch measurement was used by the group of Dr. A. J. Fourcin of University College of London in 1974 in an instrument called 'laryngograph'. This instrument has been used as a benchmark for the fundamental frequency measurements. This instrument is currently marketed by Kay-Electronics [17]. A number of glottal pitch extractors are available, such as the Electroglottograph (type EG830) and Portable Electroglottograph (type EG80) from F-J Electronics A/S [9], [10]. The method of impedance variation detection gives information about the contact phase of vocal fold vibration cycle and hence has a lot of potential for clinical and pathological use [5].

1.2 PROJECT OBJECTIVES

The impedance variations to be sensed in electroglottography are generally less than 1Ω over a base impedance of 500Ω [5]. The aim of this project was to design and develop a low-cost, battery operated glottal pitch extractor to detect these impedance variations across the human larynx. A low cost microcontroller based signal acquisition, analysis, and LCD graphics display unit is also to be developed as a part of this instrument for displaying the impedance variation waveforms. The instrument is to be designed with a RS-232 serial interface for downloading the measurement results on to a computer and for taking a hardcopy of the stored waveforms on a plotter. A signal acquisition and analysis software is to be developed on a notebook PC with an A/D card, for capturing the impedance variation waveforms and carrying out the analysis.

This project is a continuation of development work done earlier at IIT Bombay by Bhagwat (1990) [3], Sriram (1991) [31], Thajudin (1994) [33], and Mahajan (1995) [18]. The circuits developed earlier had noise problems and the impedance change of less than 1Ω could not be easily detected. In this project, the aim is to reduce the noise problems and to increase the sensitivity of the circuit for proper extraction of the impedance variation waveform. Also, it is intended to develop a stand alone data acquisition and display unit which would be useful for capturing, displaying and storing the impedance variation waveforms obtained from the impedance variation detector unit.

1.3 REPORT OUTLINE

Chapter two explains the significance of pitch and briefly gives the anatomy and physiology of the larynx and explains the vocal fold vibrations.

The principle of impedance variation and its measurement along with the applications of electroglottography are explained in the third chapter.

The hardware design of the impedance detector module and its testing is explained in the fourth chapter.

The fifth chapter explains the implementation of the signal acquisition and display unit.

The sixth chapter gives the details of the software developed on the notebook PC with DAQ-700 data acquisition card for signal acquisition and analysis.

Finally, the test results and discussion is given in the seventh chapter.

The last chapter provides summary and conclusions.

Chapter 2
VOICE PITCH

This chapter explains the structure of larynx and some of the associated organs, which are involved in the process of the generation of sound. It explains the cause of vocal fold vibrations which determine the pitch of the sound [25].

2.1 ANATOMY AND PHYSIOLOGY OF LARYNX

The speech organs are divided into three groups namely lungs, larynx, and the vocal tract. Fig. 2.1 shows the overall structure of these organs. The lungs are the source of airflow that passes through the larynx and the vocal tract before exiting the mouth as pressure variations constituting the speech signal. The voicing source occurs in the larynx at the base of the vocal tract, where airflow can be interrupted periodically by the vocal folds.

Larynx is an important organ which helps in the generation of sound waves. It lies in the frontal part of the neck, provides a passageway for air between the pharynx and the trachea. During swallowing the larynx moves upwards, occluding the opening into it from the pharynx. This ensures that food passes into the esophagus and not into the lower respiratory passages [25], [35].

The larynx is composed of several irregularly shaped cartilages attached to each other by ligaments and membranes. The main cartilages are thyroid cartilage, cricoid cartilage, two arytenoid cartilages, and an epiglottis. The structure of the larynx as viewed from behind, and front is given in Fig. 2.2 [35].

The thyroid cartilage is the most prominent and consists of two flat pieces of cartilages fused anteriorly (in the front), forming the laryngeal prominence (Adam's apple). The inlet to the larynx lies in the anterior wall of the pharynx. Continuing from the inlet, the cavity of larynx expands into a wide vestibule. Two sets of thick, membranous ridges protrude from the laryngeal wall at the lower end of the vestibule. The upper pair, called vestibular folds or false vocal folds, are not involved in the production of sound. The lower pair, the vocal folds or true vocal cords, contain the vocal ligaments and the intervening space together form the sound producing apparatus known as glottis. Each vocal cord is stretched between the thyroid cartilage anteriorly and an arytenoid cartilage posteriorly (in the back) [35].

2.2 VOCAL FOLD VIBRATIONS

During voiced segments of speech, the vocal cords are set into vibrations. The vibration cycle consists of three main phases: the contact phase, the separation phase, and the open phase. Vibration of the vocal folds, or phonation, occurs when (a) the vocal folds are sufficiently elastic and close together and (b) there is a sufficient difference between the pressure below the glottis (subglottal pressure) and the pressure above the glottis (supraglottal pressure). As the air is breathed in, the vocal cords adduct (come together). When the air is exhaled from the lungs, the vocal cords abduct (move apart). As the velocity increases in the narrow glottis, the local pressure drops. When sufficient pressure difference exists across the glottis to cause a large airflow, a negative pressure develops in the glottis that forces the vocal folds to close. Glottal closure interrupts the airflow, and a pressure gradient develops across the glottis, eventually building to a point where the vocal folds open again [25]. Thus the vocal folds are set into vibrations. Typical speech uses a pitch range of about an octave 80 - 160 Hz for males; while singers often use a two octave range. Average pitch values are 132 Hz and 223 Hz for males and females, respectively [25], [34].

Although vocal fold closures are not in response to individual muscle contractions, muscle variations can change characteristics of the airflow. Variations in the subglottal pressure and vocal fold elasticity due to the muscles of chest and larynx, respectively, cause changes in the waveshape of glottal air pulses, including their duration and amplitude, which lead to changes in pitch and loudness [25]. Fig. 2.3 shows the position of the vocal cords when they are opened and closed [35].

The pitch of the voice, i.e., the number of vocal fold vibrations per second, depends upon the length and tightness of the cords. In adults, the vocal cords are longer in the male than in the female, typically 15 mm long in male, and about 13 mm long in female, thus the male voice has the lower pitch [25], [34]. The loudness of the voice depends upon the force with which the cords vibrate. The greater the force of expired air, the more the cords vibrate and the louder the sound. The quality and resonance of the voice depends upon the shape of the mouth, the position of the tongue and the lips, the facial muscles, and the air sinuses in the bones of the face and skull [35].

Fig. 2.4 shows the representation of the sequence in a vocal fold vibration cycle [17]. The two folds come into contact due to a wavelike motion starting at the lower end as shown in Fig. 2.4 (e) and (f). Fig. 2.4 (a) shows the condition of maximum contact between the two folds. Fig. 2.4 (b) and (c) represents the separation phase, which is also gradual from the lower surface upwards. Fig. 2.4 (d) shows the open phase where the two folds are completely apart.

2.3 SIGNIFICANCE OF PITCH

An important aspect of speech perception concerns prosody, whose domain of variation extends beyond the phoneme boundary into units of syllables, words, phrases, and sentences. The perception of rhythm, and stress patterns helps the listener understand the speech messages by pointing out important words. The basic function of prosody are to segment and to highlight [25].

In addition to the speaker identity and emotional state of the speaker, voice pitch variations convey many other kinds of information. The patterns of pitch variations that occur during a phrase, known as intonation, mark the boundaries of syntactic units and sometimes affect their meaning. It has been determined that pitch and amplitude patterns vary with emotions, that emotions often raise pitch and amplitude levels and their variability [25].

In English, statements have a falling pitch; while sentences meant as questions have a rising pitch. In many other languages, pitch variations known as "tone", also contribute to the meaning of the word [25].

Having given the above introduction and background material, we will examine, in the following chapter, how the different phases of phonation can be determined using the principle of impedance variation.

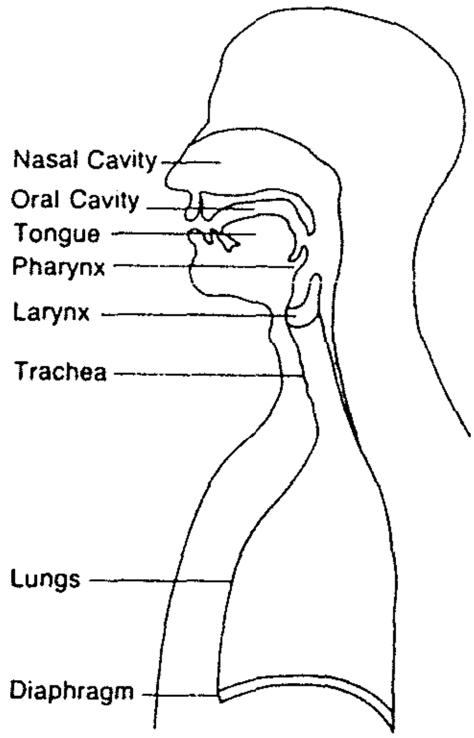
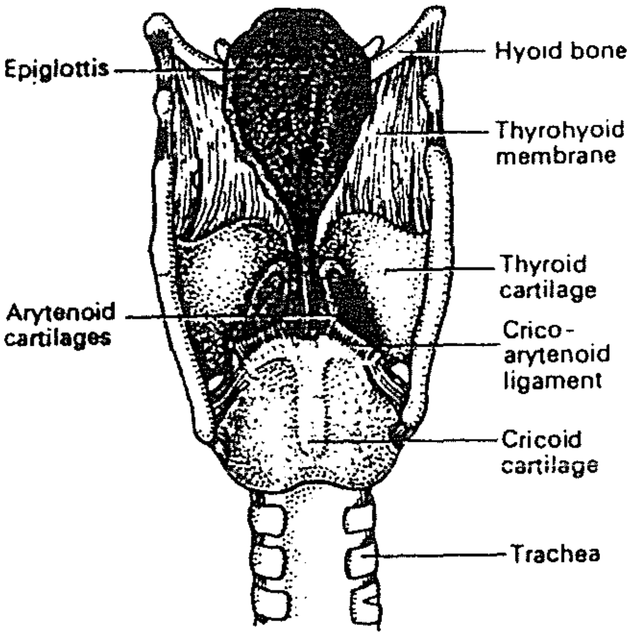
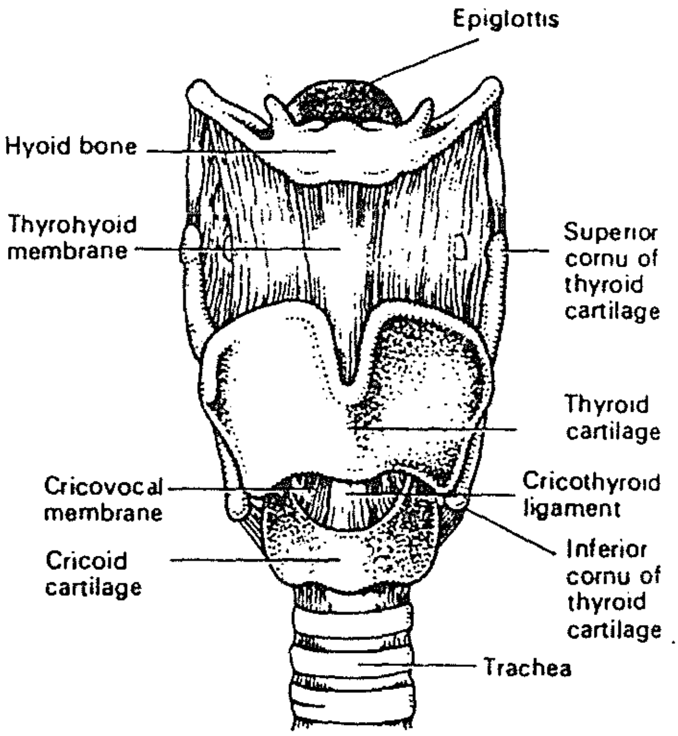


Fig. 2.1 The speech organs [25].



(a)



(b)

Fig 2.2 Structure of larynx (a) viewed from behind, (b) viewed from front [35].

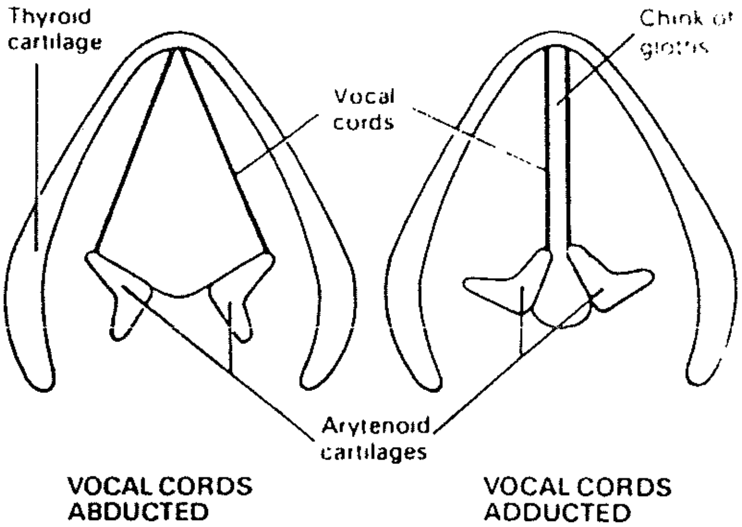


Fig. 2. 3 Positions of the vocal cords [35].

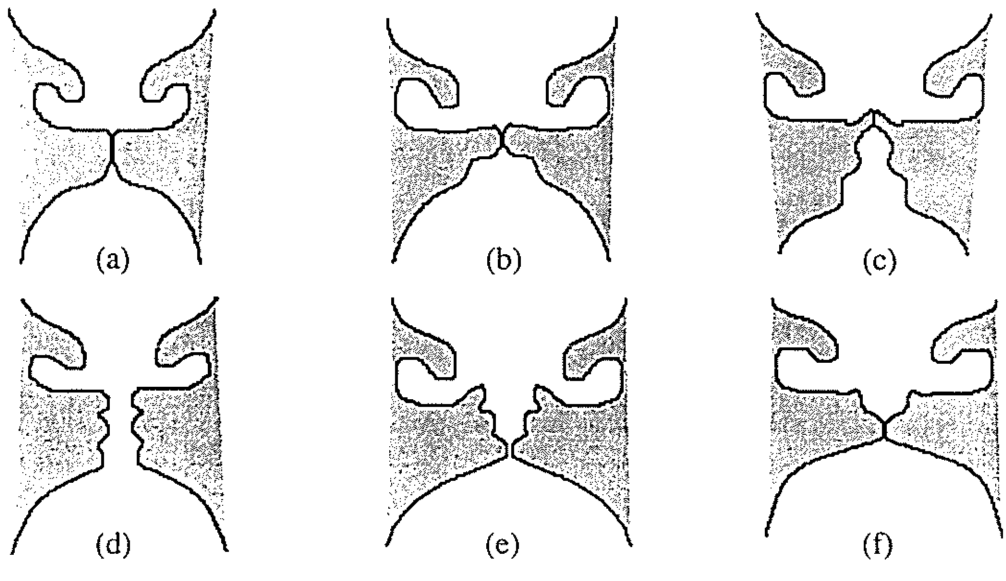


Fig. 2.4 The vocal fold vibration sequence [25].

Chapter 3

PITCH MEASUREMENT BY IMPEDANCE VARIATION TECHNIQUE

The pitch of the sound can be found using the technique of impedance variation across the larynx. The electrical impedance variation provides information about the contact phase of the vocal fold vibration cycle. This chapter explains the technique and its application.

3.1 IMPEDANCE VARIATION ACROSS THE LARYNX

As said before, phonation consists of three phases; the contact phase, the separation phase, and the open phase.

As the two vocal cords come into contact, the impedance across them reduces, it is the least when they are in complete contact. As the cords separate, the impedance across them gradually increases and it is at its maximum when the cords are completely separated i.e., the open phase. The impedance level does not show any considerable change even if the cords are wider apart. The impedance level is sensitive to the contact area and thus gives information about the contact phase. This is the only parameter giving information about the actual contact and is very useful, since the nature of contact depends upon the physical condition of the vocal cords. Thus, electroglottography turns out to be very important for the laryngeal function assessment [5], [30].

Fig. 3.1 shows a model of the impedance variation waveform. This figure labels critical points on the model waveform which is a straight line approximation for the waveform. At all the key points, the positions of the vocal cords are shown starting from the closed phase to the open phase [5].

3.2 MEASUREMENT OF IMPEDANCE VARIATION

A high frequency current is passed through the electrodes in contact with the skin placed on the side of thyroid cartilage. The voltage developed across the electrodes gets amplitude modulated due to the variations in the impedance in the current path, caused by the vibrations of the vocal folds. The percentage of amplitude modulation depends upon the change in the tissue impedance in the current path. The impedance is minimum when the vocal folds are in full contact since the tissue path, i.e., the effective tissue length for the RF current, is smallest between the electrodes for this vocal fold configuration. Since the total impedance is a function of the tissue path length (as well as the tissue cross-sectional area), the impedance increases as the folds open because the current path becomes less direct [5]. This modulated waveform is then demodulated to obtain the electroglottogram. This method gives one peak for each vocal fold contact.

The base impedance across the thyroid cartilage is approximately 500 Ω and the change in the impedance due to vocal fold vibrations is less than 1Ω [5], [9], [10]. The usable frequency range of the current carrier is in the range of 100 kHz - 5 MHz [5], [9], [10]. The current used is generally in the milliampere range. The voltage level between the electrodes is dependent on the tissue impedance and the current, but is typically of the order of 0.5 V [9], [10].

The glottal pitch extractor is an instrument which measures the electrical impedance variations of the larynx using a pair of electrodes held in contact with the skin on both sides of the thyroid cartilage. It registers the contact between the vocal folds as a time varying signal. The amplitude variations of this signal are representative of the amount of contact between the vocal folds.

Electroglottogram, defined as the waveform used to describe laryngeal behavior in the vicinity of the glottis, has number of applications such as the physical assessment of the vocal folds, lip-reading, speech analysis & recognition. Out of these, the laryngeal function assessment and the improvement of lip-reading are explained in the following sections.

3.3 LARYNGEAL FUNCTION ASSESSMENT

One of the major fields where electroglottography plays an important role is the assessment of the laryngeal function. Since the electroglottograph provides the information regarding the actual contact phase of the vocal folds, it is possible to detect the physical condition of these folds [5], [30]. The demodulated output from the electroglottograph, proportional to impedance variation, is known as L_x (laryngogram) waveform.

Fig. 3.2 (a) shows the L_x waveform for normal voice. During normal voicing, the cover of the vocal cords is deformed by a wavelike motion from the beginning of the open phase till its end. The onset of contact is due to the bridge formed by the mucous membrane on the two sides of the vocal cords. When the two covers come together and when the area of contact is maximum, a peak is obtained. Separation of the vocal cords is also initiated at the lower surfaces. The two sides gradually pull apart and the contact area becomes minimum which leads to the open phase [8], [17], [25].

For breathy voicing, the air flow required is larger than that needed normally. Here, the vocal cord contact does not necessarily occur during vibration. Fig. 3.2 (b) shows the L_x waveform for breathy voicing which shows small and well defined closure peaks which are more symmetrical than for normal voicing. The open phase is quite long between the end of separation and beginning of the contact phase [8], [17], [25].

Creaky voice is characterized by its low and irregular pitch and sharply defined vocal tract resonance. The L_x waveform shown in Fig. 3.2 (c) shows a pair of vocal cord contact-separation sequences. A small peak precedes a larger peak, both occurring with considerable irregularity. Very long closure durations are indicated by the width of the larger peak [8], [17], [25].

Analysis of normal L_x waveform was carried out by Fourcin [8] and the following reference features were established :

1. Uniform L_x waveform peaks are likely to be associated with uniform acoustic output.
2. Long closure duration (separation and closure) is likely to be associated with well defined formants.
3. Sharply defined contact implies good acoustic excitation of the vocal cords.

4. Regular, sharply defined contact periodically gives well defined pitch.
5. Progressive change in sharply defined L_x period lengths are associated with smoothly changing voice pitch.

In pathology, an indication of the nature and degree of the disorder can be made by interpreting the L_x waveform with the reference to the above points.

A plot of instantaneous pitch frequency (obtained from measurements of instantaneous pitch periods) as a function of time is known as F_x waveform. This is considered useful for observing perturbation in the pitch, and also for detecting unvoiced segments.

The F_x histogram is a probability distribution of vocal fold vibration frequencies observed over a representative time segment of speech signal. The frequency range of interest is divided into a number of equal bins. Each bin is incremented when the pitch value falls within its frequency range. The content of each bin divided by the total number of pitch periods gives the measure of the probability distribution. A histogram obtained by this method is known as a single period histogram. In a triple period histogram the content of each bin is incremented only when three successive pitch periods fall within its range. The relative height and width of this distribution provides a measure of vibrational regularity [8].

The healthy larynx characteristically has sharp edges to its frequency distribution and well defined frequencies of vibrations. Fig. 3.3 shows the single period F_x histogram for normal voice [8]. Pathological conditions show characteristic F_x histograms. Laryngitis shows limited frequency range and a higher pitched voice. For older people, lower frequencies have a more gradual distribution. For a younger person, the low frequency edge distribution is abrupt and higher frequency periods are gradually reduced in probability. Smoking increases the irregularity of vocal fold vibrations. F_x histograms thus prove to be very useful research tool and also has applications in clinical therapy [8].

3.4 IMPROVING LIP-READING

Lip-reading is one of the alternatives for profoundly deaf person, to understand a conversation. Lip-reading, sometimes becomes very difficult for differentiation of certain words which have identical lip movements. Some of the phoneme pairs which creates the discrimination problem are /g/ and /k/, /b/ and /p/, and /v/ and /f/. Words such as 'vat' and ‘fat’ which have identical lip movements are very difficult to distinguish. The only difference between ‘vat’ and ‘fat’ is that the vocal folds vibrate throughout the /v/ and not the /f/. It has been shown that when pitch is provided externally, lip-readers are able to comprehend speech much better [27]. It has been shown that if the voice pitch is provided externally, normal listeners are able to lip-read a person reading continuous text, at a rate of two and half times faster when they are lip-reading without the provision of pitch [27]. The provision of pitch greatly improves lip-reading ability in those who have acquired an auditory knowledge of speech before becoming deaf [27].

3.5 SPEECH PROCESSING APPLICATIONS

EGG waveform finds wide application in the area of speech processing, mainly in the following

(a) Voiced-unvoiced classification

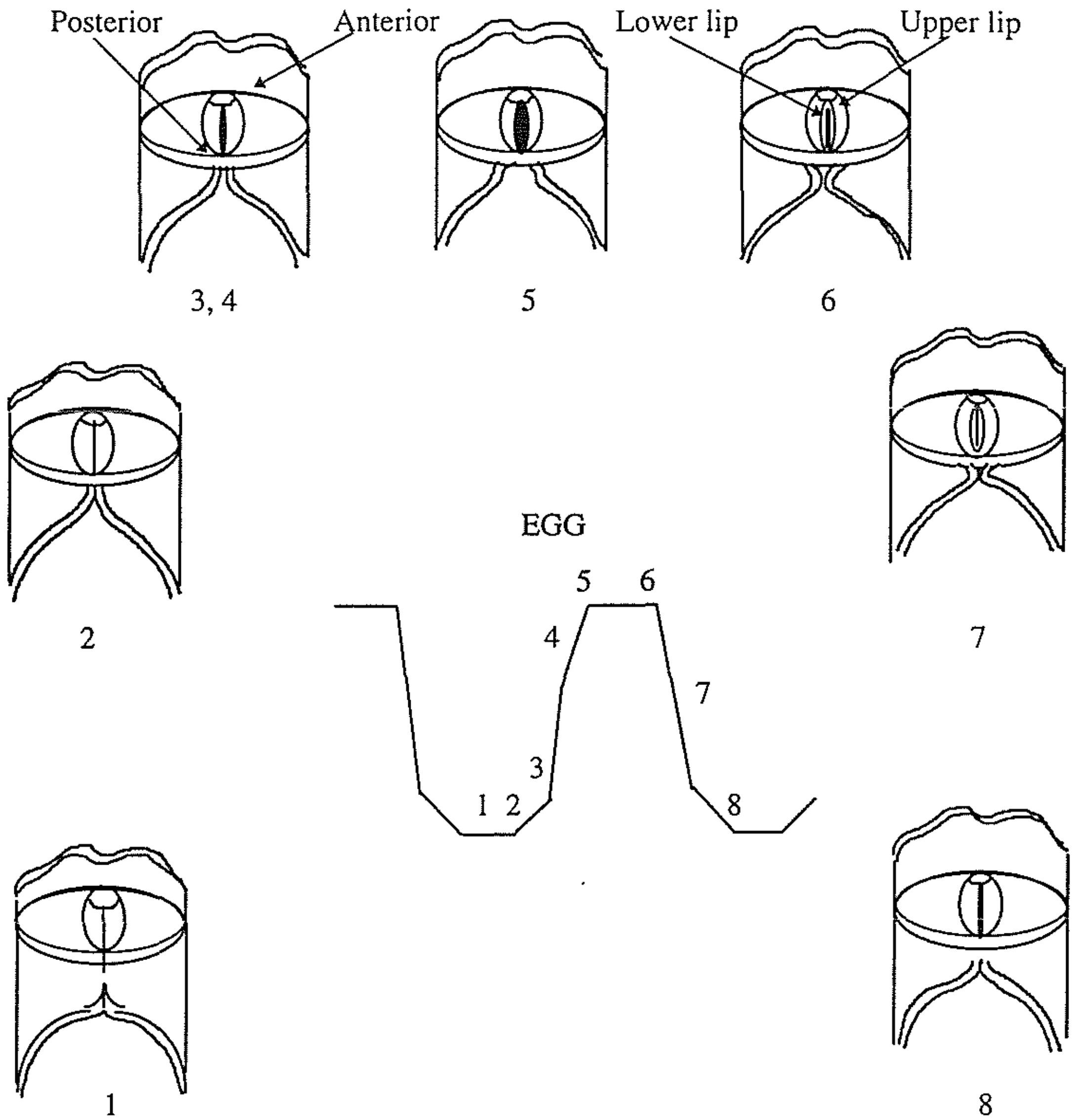
Algorithms for classification of a speech segment into voiced, unvoiced, mixed, and silent have taken on new levels of accuracy and simplicity when EGG signal is used as an aid in the decision making process. For just a voiced-unvoiced decision, thresholding the EGG signal alone suffices since the EGG is ideally zero during nonvoiced regions and periodic and nonzero during voiced regions [5].

(b) Fundamental frequency estimation

Using the EGG, the pitch calculations becomes simpler and accurate, since usually the pitch value is based on either zero crossings or the distance between the minima in the differentiated EGG [5].

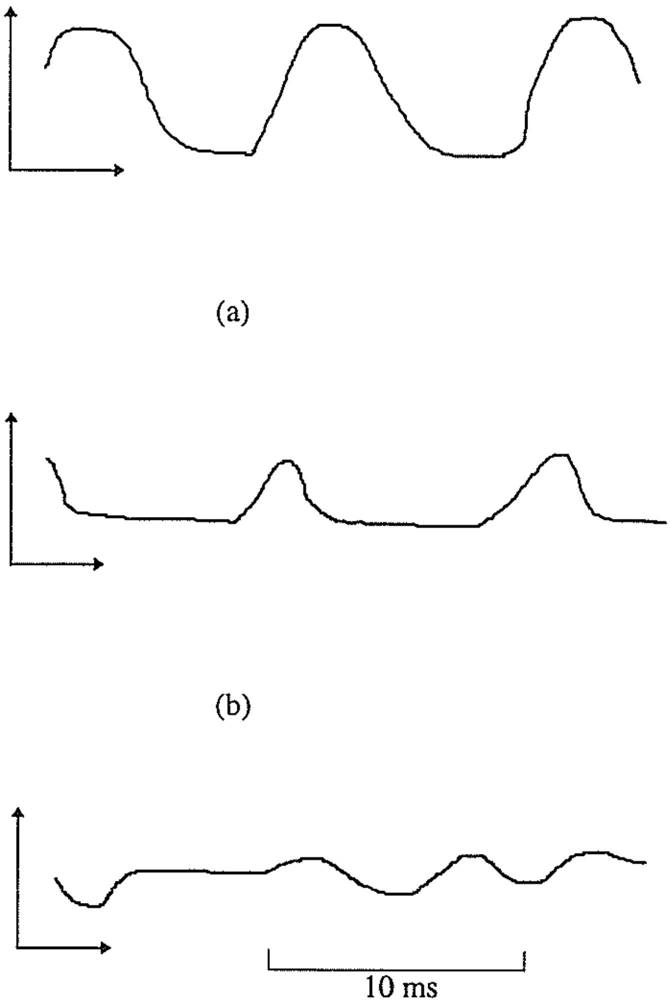
(c) Synthesis applications

The EGG-aided analysis play a key role in the synthesis of high quality speech. This is the case since naturalness and intelligibility of synthesized speech are influenced by factors such as accuracy in vocal tract modelling, voiced-unvoiced classifications, pitch detection, and the nature of excitation used [5].



1 Vocal folds maximally closed.
2 Folds separating along lower margin toward upper margin.
3 Folds separating along upper margin.
4 Folds continuing to open.
5 Folds apart, no lateral contact.
6 Folds starting to close, starts at the lower margin from posterior to anterior but still no contact.
7 Folds making first lateral contact along lower margin at anterior.
8 Folds completing closure with increasing lateral contact.

Fig. 3.1 Model of EGG waveform with corresponding drawings of vocal fold configurations [5].



(a)

(b)

(c)

Fig. 3.2 L_x Waveforms for characteristic voicing (a) L_x waveform for normal voice, (b) L_x waveform for breathy voice, (c) L_x waveform for creaky voice [8], [17]. Y-axis indicates the impedance variation.

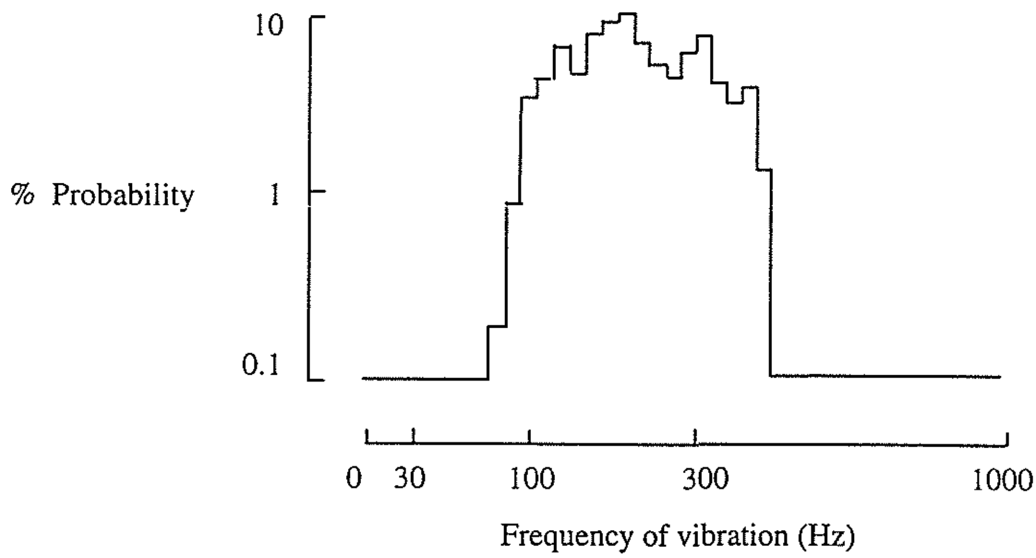


Fig. 3.3 F_x histogram for normal male voice [8].

Chapter 4

DESIGN OF THE GLOTTAL PITCH EXTRACTOR

This section describes the design of the glottal pitch extractor and also the design of the electrodes used for acquiring the impedance variation waveform across the larynx.

4.1 DESIGN OF THE IMPEDANCE VARIATION DETECTOR

A schematic of the earlier instrument design is shown in Fig. 4.1 [5]. It consists of a RF source for generating the high frequency carrier signal which is given to the electrodes through a transformer and an AM detector for demodulating the received amplitude modulated waveform, across the load R_l. The instrument can be designed by using a scheme as shown in Fig. 4.2. One of the main advantages of this scheme is that it avoids the use of a transformer as used in the earlier designs.

Fig. 4.3 shows the detailed block diagram of the impedance variation detector. The overall scheme consists of an oscillator which generates a carrier frequency with a stabilized amplitude. It is then applied to the glottal impedance sensor (GIS). The amplitude modulated voltage waveform obtained from the GIS is then demodulated. The design of each block of the circuitry is explained in the following subsections.

4.1.1 Oscillator

This block generates a sinusoidal voltage waveform of 300 kHz with stabilized 5 V p-p amplitude. Fig. 4.4 shows the circuit diagram of this oscillator [4], [11], [29].

A Wien bridge oscillator is designed, using op-amp U1, for generating the carrier signal and the amplitude of this signal is stabilized using the FET Q_1 as a voltage variable resistor. The condition for oscillation, i.e., unity loop gain is satisfied when

$$(R_3 + R_d) \parallel R_6 = R_7 / 2$$

where R_d is the dynamic resistance of Q_1, and is determined by the negative bias V_{GS}, which is a function of oscillation amplitude. The potentiometer (R_7) can be adjusted to set the oscillation amplitude. The frequency of oscillations is given as

$$f_0 = \frac{1}{2\Pi RC}$$

where $R = R_1 = R_2$, and $C = C_1 = C_2$.

The values of the resistors and capacitors which decides the frequency of oscillations has been obtained for f_o = 300 kHz. High frequency op-amp, CA3140 [29] has been used as op-amp U1.

4.1.2 Glottal Impedance Sensor

The glottal impedance sensor (GIS) corresponds to the voltage-to-current converter, electrodes for applying the current and the amplifier for sensing the voltage waveform across the electrodes as shown earlier in Fig. 4.2. The construction of the electrodes is explained in the next section.

The V-I converter, as shown in Fig. 4.5, produces a sinusoidal current source [4], [11], [29]. The output current is determined by the resistors R_8 and R_9, and is given as

$$I = \frac{V_C}{R8 + R9}$$

R_8 is made variable so that the current injected can be adjusted. A fixed resistor R_9 is connected is series with R_8, so that the current always remains within the safe limit even when the potentiometer is shorted accidentally.

The value of R_{11} is chosen in such a way that the voltage drop across it is not too large to cause the output saturation of the voltage to current converter. R_{11} should preferably be much larger than the load impedance, i.e. impedance in the path of the electrodes.

This excitation current is applied to the electrodes through the capacitors C_6 and C_7, in order to prevent the passage of any DC current. Since one end of the load is virtual ground, this voltage to current converter circuit, also provides the sensed output voltage V_m, and there is no need for a difference amplifier.

4.1.3 Demodulator

The sensed output V_m is high pass filtered by C_7 & R_{12}, with a cutoff frequency of 16 kHz to obtain the amplitude modulated waveform, from which the impedance variation waveform can be extracted by the demodulator.

The demodulator consists of an amplifier, a precision full wave rectifier and a low pass filter as shown in Fig. 4.6. This signal is first amplified using a non-inverting amplifier. The gain of this amplifier can be adjusted by varying the resistance R_{14} from 1 to 10. The capacitor C_8 connected in series with R_{13} acts as a high pass filter which suppresses the power line (50 Hz) hum. The capacitor connected at the output of this amplifier acts as a high pass filter along with R_{15} and R_{18} with a cutoff frequency of 16 kHz, so that the power line interference (50 Hz) gets further suppressed [4], [11], [29].

This amplified signal is then given to a precision full wave rectifier circuit, with gain = 1. Both halves of the waveform are amplified equally by satisfying the following condition

$$R15 = R17 = R18 = R,$$
$$R19 = R/2,$$
$$R21 = R$$

The summer stage of the precision rectifier also acts as a low pass filter with a cutoff frequency of 13 kHz, which eliminates the high frequency noise and the carrier frequencies.

4.1.4 Amplifier

The output of the demodulator is passed through two stages of inverting amplifiers as shown in Fig. 4.7 [4], [11], [29]. The gain of each stage is fixed at 70. The gain of the first stage is

$$A1 = -\frac{R23}{R22} \cdot \frac{1}{1 + j\omega\ C12R23} \cdot \frac{j\omega\ C11R22}{1 + j\omega\ C11R22}$$

C11 and R22 provide a high pass filtering with cutoff of 7 Hz for filtering out the mean DC. C12 and R23 provide low pass filtering with cutoff frequency of 10 kHz for eliminating the harmonics of the carrier.

The second stage has a gain

$$A2 = -\frac{R27}{R25} \cdot \frac{1}{1 + j\omega\ C13R27}$$

R27 and C13 provide a low pass filtering with cutoff frequency of 10 kHz.

4.2 ELECTRODES USED

The electrodes can be constructed in two ways; a two electrode system with a guard ring and a four electrode system without the guard ring. These constructions are explained below.

In the two electrode configuration (Fig. 4.8 a), the excitation electrodes are also used as sensing electrodes. In the simplest four electrode configuration (Fig. 4.8 b), the current is passed through the inner electrodes and pick up of the modulated voltage signal is done from the central ring. In both these methods, the current flow path is very diffuse, most of the current flow not being through the glottis area, and therefore the method results in very low sensitivity. Further, generally, there will be large common mode interference. In another method, the current is passed through the central electrodes and the modulated voltage is also collected across the same electrodes, and the guard rings are shorted together and thus these act as the current director rings. Another possible configuration (Fig. 4.8 d) consists of a buffer for driving the individual guard rings. The potential developed at the central rings is used for driving the guard ring of each probe separately through a buffer. This ensures the flow of the excitation current into the larynx and avoids the leakage of the current from the source across the skin of the neck, resulting in enhanced sensitivity. However, it is to be noted that there will be additional current flow between the outer rings.

The choice of a particular configuration becomes very important since the sensitivity of the circuit depends upon the configuration. If the electrodes without guard rings are used, then it is found that, the current flowing in the electrodes finds an easy path across the skin of the neck and thus reaches the second electrode. Thus very small or negligible amount of current passes through the larynx and hence the sensitivity of the arrangement reduces.

The electrodes are constructed on a glass epoxy printed circuit board as shown in Fig. 4.9. The PCB is double sided with plated-through-holes for establishing electrical connection between the two sides. The side of the electrode which is to be placed in contact with the skin has a complete connection free surface so that electrode application does not cause discomfort to the subject. The lead wires are soldered on the other side of the electrode. The electrodes are gold plated in order to avoid corrosion due to electrolysis.

Different configurations of the electrodes were tried out for obtaining the maximum sensitivity. Various configurations are listed below

(a) Guard rings of electrode 1 and electrode 2 shorted together and kept floating.
(b) Guard rings of electrode 1 and electrode 2 shorted together and connected to ground.
(c) Guard rings of electrode 1 and electrode 2 not connected to any point.
(d) Guard ring of electrode 1 connected to ground and that of electrode 2 shorted with its central ring, as shown in Fig. 4.10.

In our experiments, it was found that, the sensitivity is maximum when the voltages at the two sensing electrodes are buffered and used for actively driving the respective guard rings. In the V-I converter arrangement, this has been achieved in case of (d), when the guard ring of the electrode 1, that is connected on the input side of the V-I converter, is connected to ground and the guard ring of the electrode 2 is shorted with its central ring as shown in Fig. 4.10. This is because, the potential on the two rings of electrode 1 is now same and thus the current flowing in the central ring of that electrode passes into the larynx instead of flowing across the skin of the neck. However, it is to be noted that additional current will be flowing through the larynx between the guard rings. The two rings of electrode 2 in Fig. 4.10 are shorted together and hence represents a single ring. The electrode construction was kept similar for both the electrodes in order to make the design cost effective and also for trying out various configurations of the electrode connections.

It is to be noted here that, the electrodes should be positioned properly across the neck without applying an electrode gel. Since the carrier frequency is very high it is not affected by the electrode contact potential. The application of the gel will short the two rings of the electrodes. Thus the current, instead of passing into the larynx, will find an easier path between the two rings and the sensitivity of the circuit will reduce.

4.3 TESTING OF THE CIRCUIT

The amplitude stability of the oscillator was tested against the power supply variations. The amplitude of the carrier signal is stable at 4.5 V_{pp}, over supply variation from $\pm$ 8 V to $\pm$ 9 V and the frequency also remains quite stable at 305 kHz. The results obtained are given in Table 4.1.

The voltage to current converter circuit was tested by connecting different loads across pins 2 and 6 of U2. Table 4.2 gives the output voltages obtained for the different values.

The demodulator circuit was tested using an amplitude modulated signal, generated from a function/arbitrary waveform generator (HP 33120A). The carrier frequency was kept at 300 kHz and an external modulating signal of frequency varying from 80 Hz to 1000 Hz was applied to the waveform generator. The amplitude of the modulating signal was adjusted to give different percentage of amplitude modulation. The amplitude of the carrier signal was kept at 0.5 V peak-to-peak, which corresponds to the voltage developed across a base impedance of 500 Ω with a current of 1 mA. The modulation index "m" will correspond to impedance variation ΔZ

$$m = \frac{\Delta Z}{Z_b}$$

The output voltage was recorded as a function of modulation index, for different modulating frequencies in the range of 80 - 1000 Hz. The results are given in Table 4.3. From these simulated impedance variation results, we can say that for a base impedance of 500 Ω, impedance variation down to 0.6 Ω can be extracted. It was observed that as the percentage of modulation becomes less than 0.5%, the 50 Hz noise problem becomes more severe. This could be because of the signal source itself, since the modulating signal was generated using a separate signal generator which gives sufficient 50 Hz noise for very low amplitudes such as 20 mV.

Finally, the electrodes were connected to the circuit and the waveforms were observed. The oscillator was replaced by sinusoidal carrier generated by a function generator. The range of carrier frequencies at which the response is maximum, varies from 250 kHz to 400 kHz with an amplitude in the range of 4 to 5 V_{pp}. Therefore, the choice of the carrier frequency of 300 kHz can be considered to be appropriate. The amplitude of the carrier is kept at 4.5 V_{pp} for getting a stable output. The current passed through the larynx can be varied for different subjects. It was found that a current of 2 - 3 mA is enough for getting the impedance variation waveform across the larynx for normal persons.

Table 4.1 Output stability of the oscillator against power supply variations

Supply voltage $\pm V_s$ (volt)	Output voltage of the oscillator V_{pp} (volt)
7.0	4.00
7.5	4.20
8.0	4.50
8.5	4.50
9.0	4.50

Table 4.2 Output of the voltage to current converter for different loads with output current of 1mA.

Load resistance (Ohm)	Output voltage of the voltage to current converter V_{pp} (Volt)
680	0.68
820	0.82
1 k	1.10
1.5 k	1.51
1.8 k	1.82
2.2 k	2.23
2.7 k	2.61
4.7 k	9.00
10 k	9.00

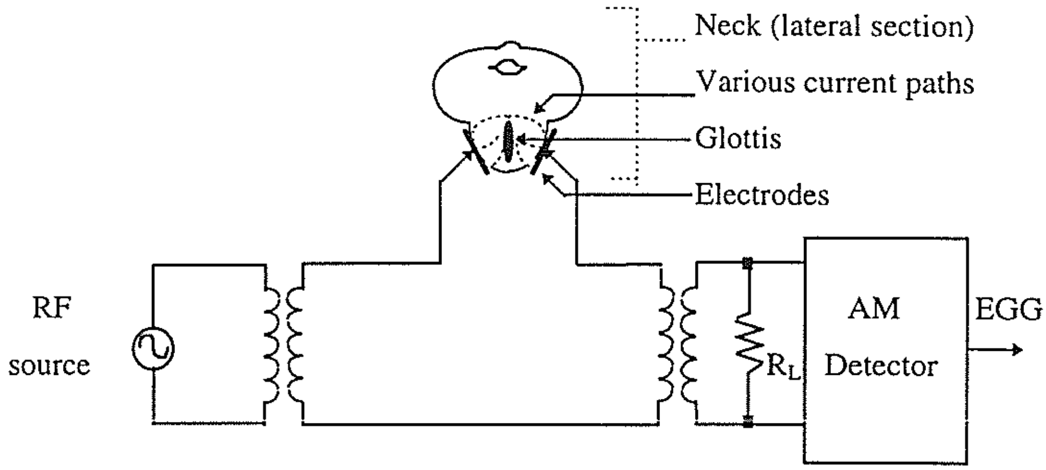


Fig. 4.1 A schematic of earlier design of electroglottographic instrument [5]

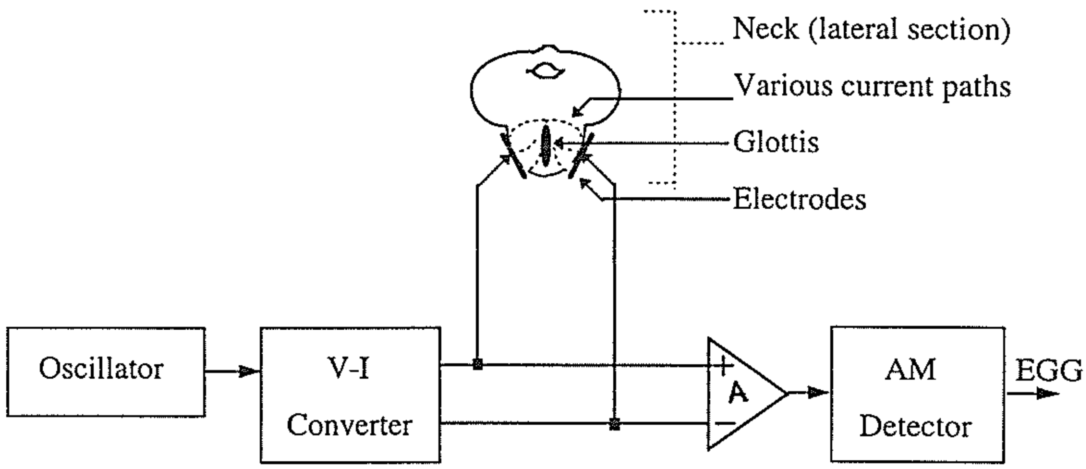


Fig. 4.2 Block diagram of the proposed system

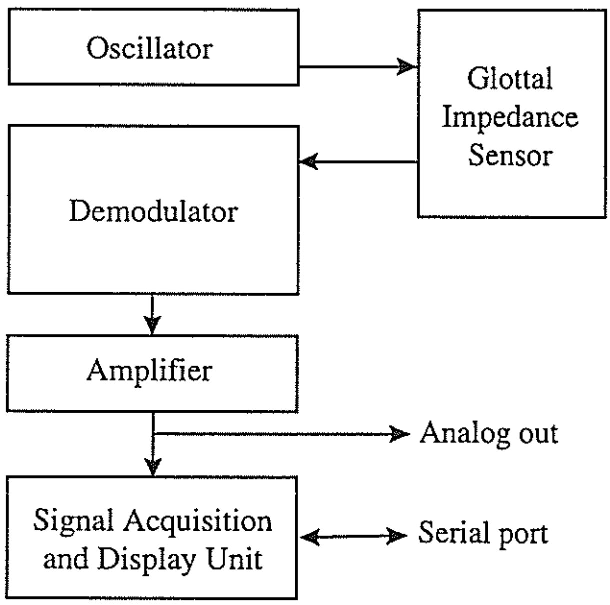


Fig. 4.3 Block diagram of the impedance variation detector

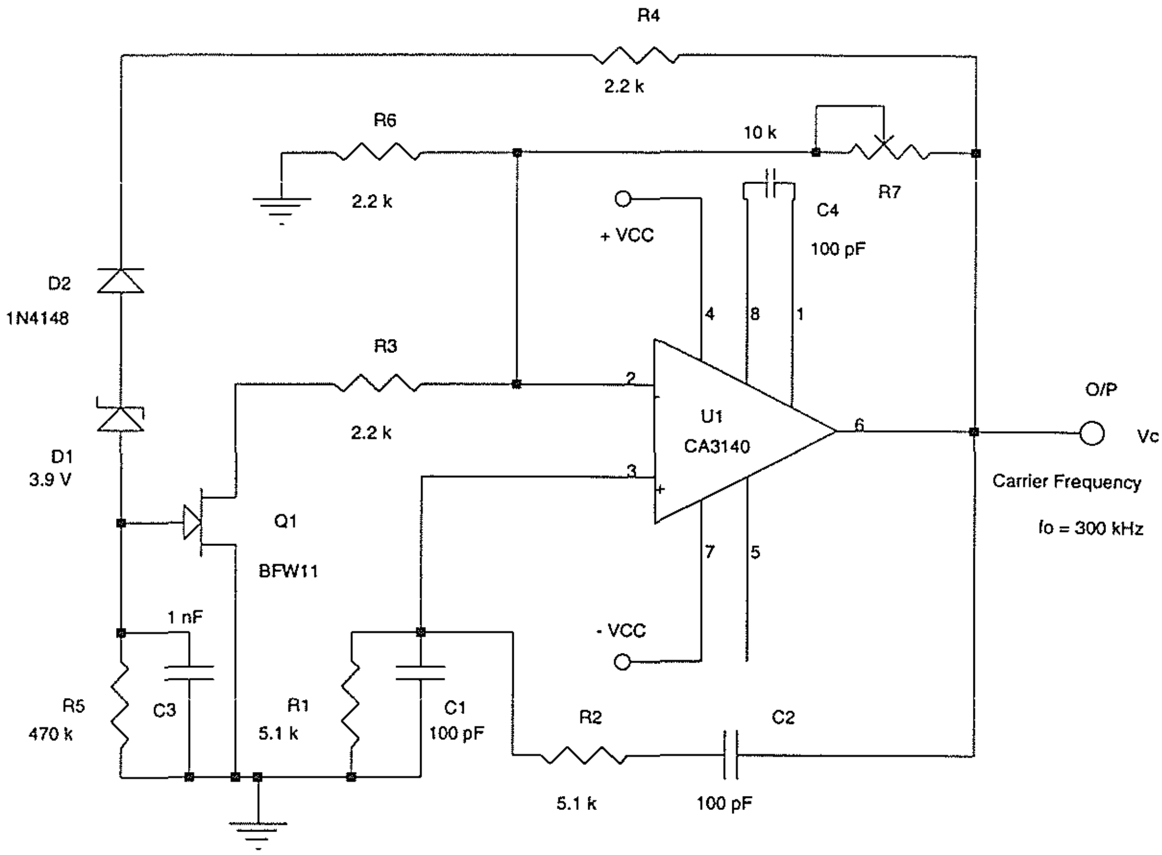


Fig. 4.4 Circuit diagram of the oscillator

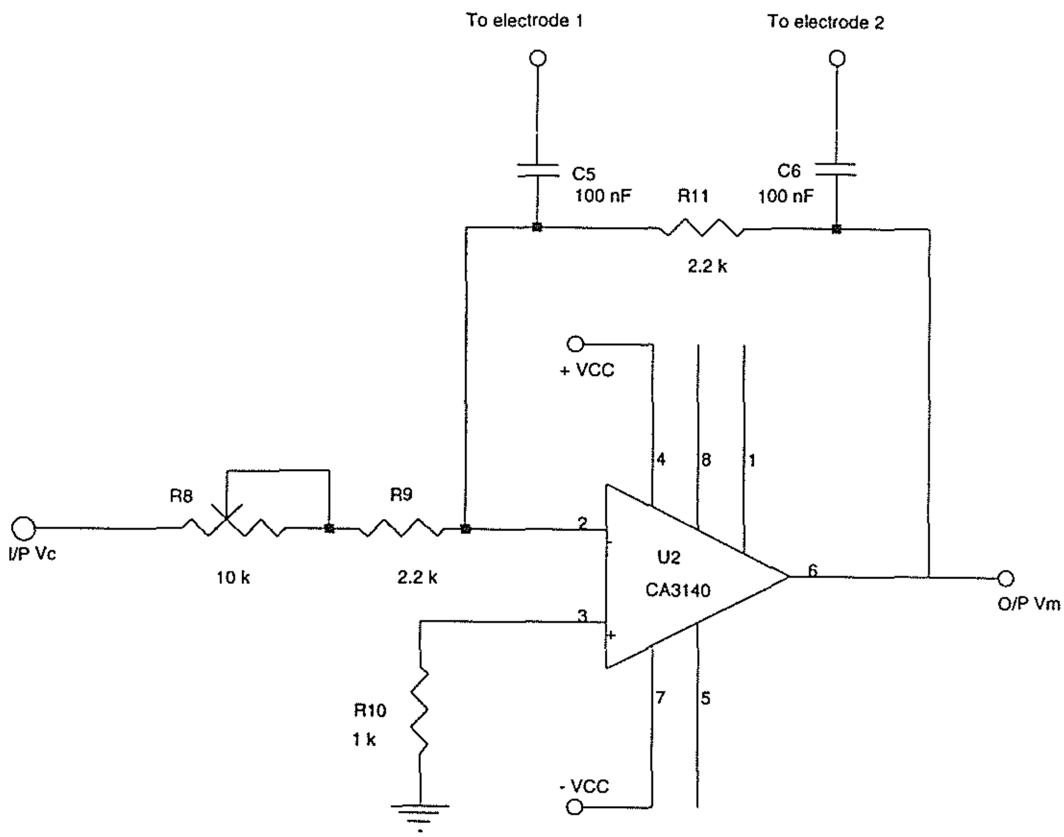


Fig. 4.5 Circuit diagram of voltage-to-current converter

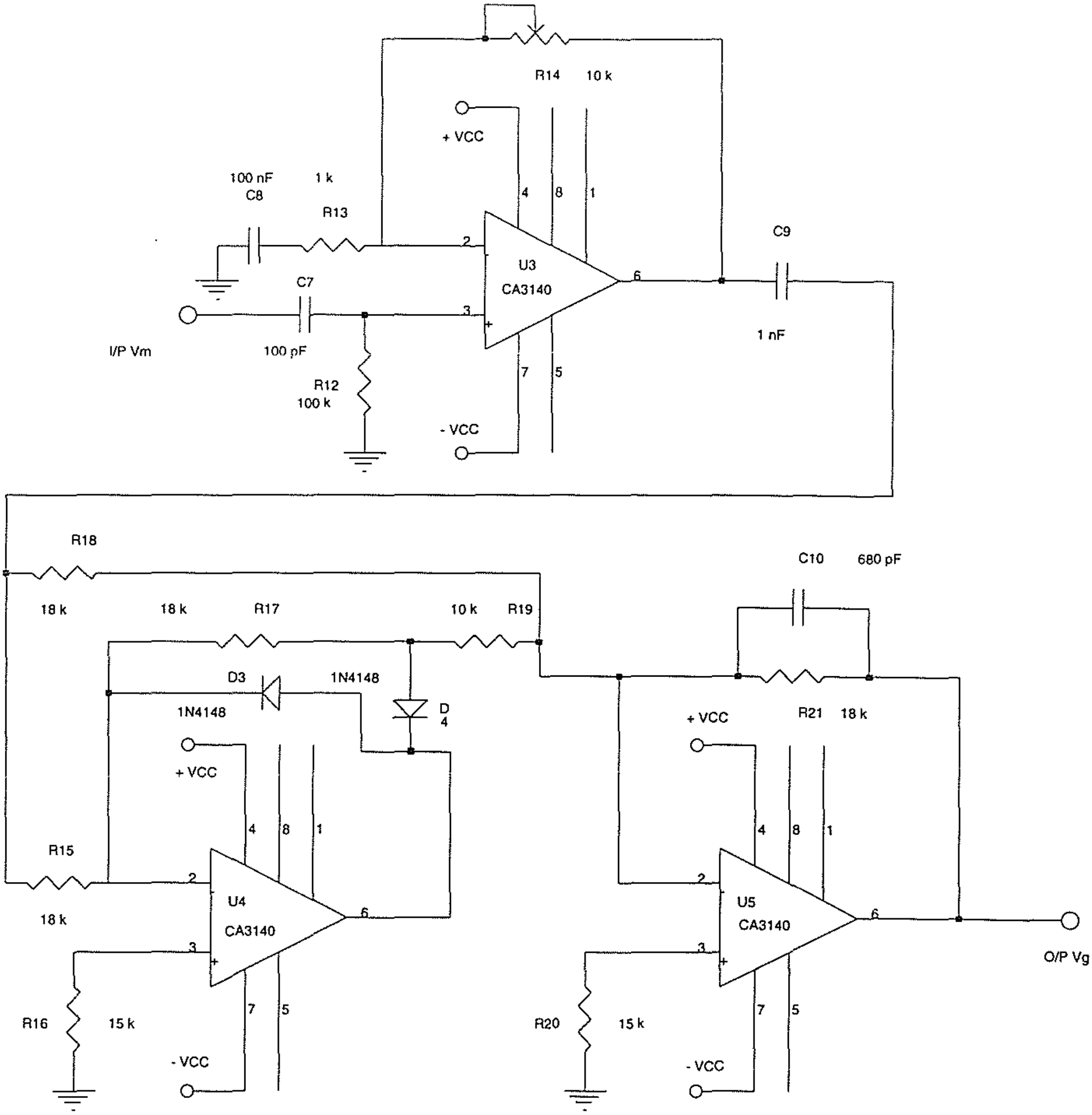


Fig. 4.6 Circuit diagram of demodulator

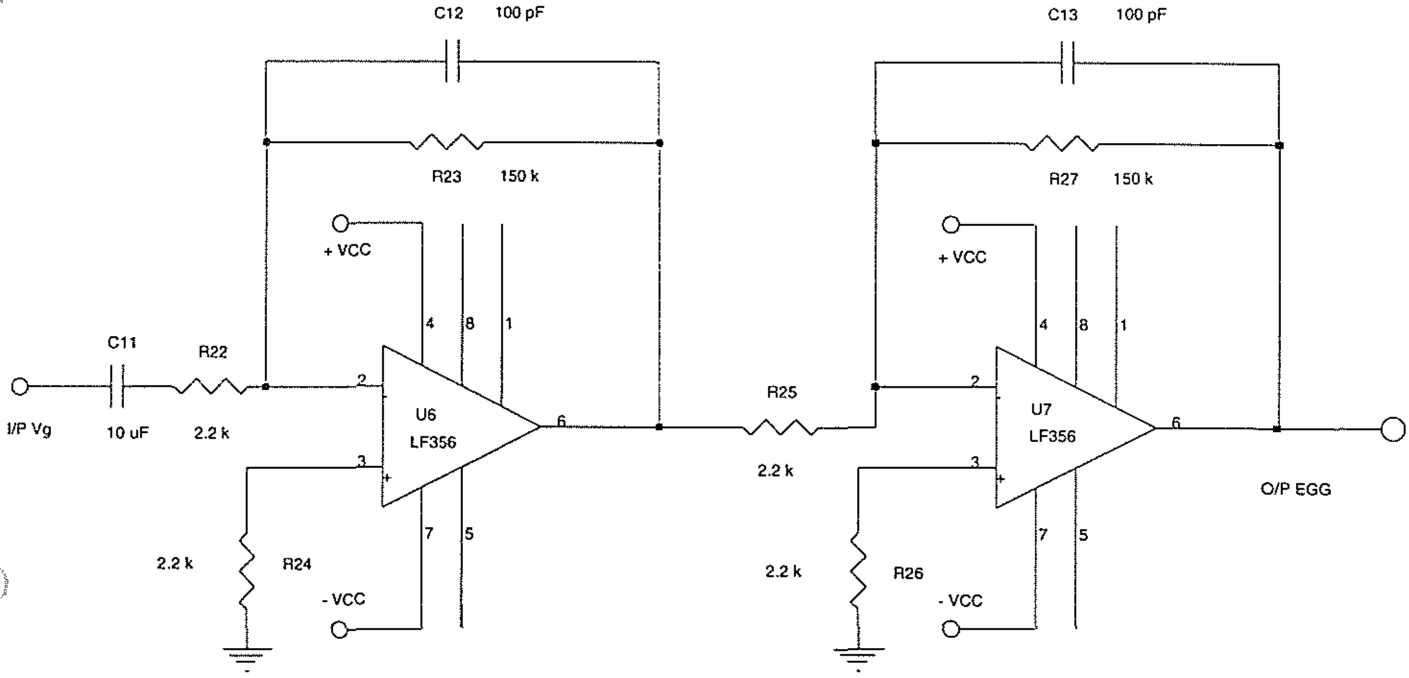


Fig. 4.7 Circuit diagram of amplifiers

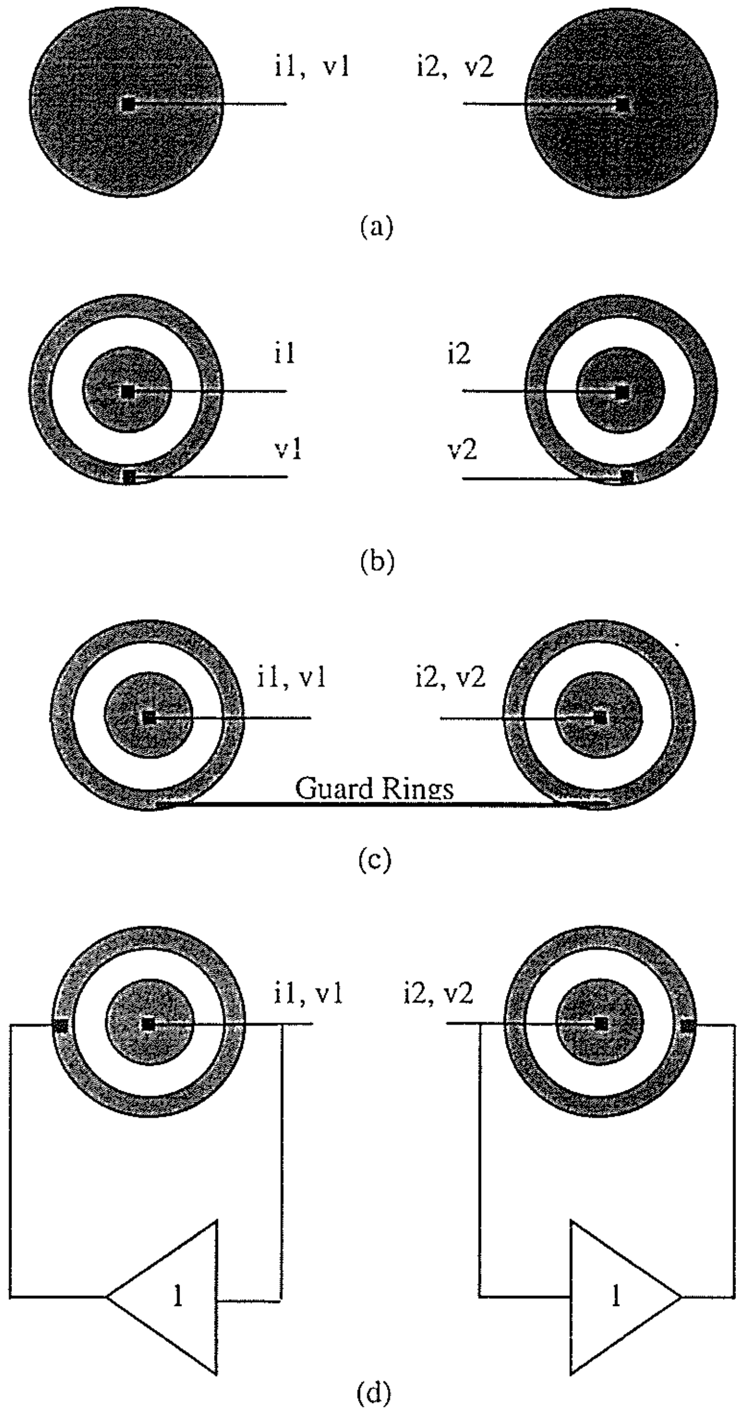


Fig. 4.8 Different electrode configurations (a) two electrode configuration, (b) four electrode configuration, (c) two electrode configuration with guard rings, (d) electrodes with driven guard rings

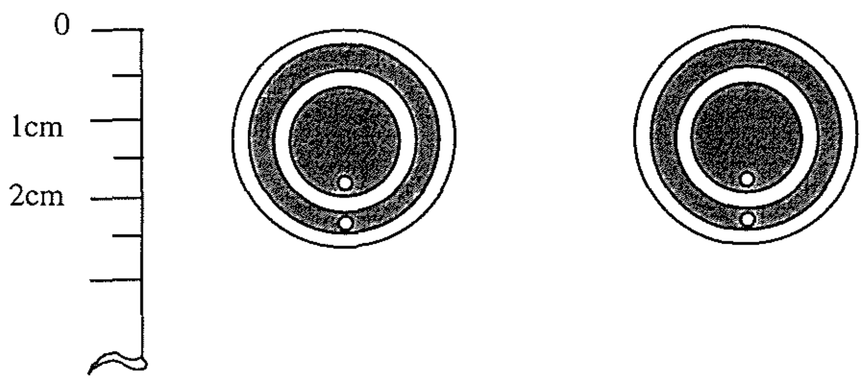


Fig. 4.9 Construction of the electrodes

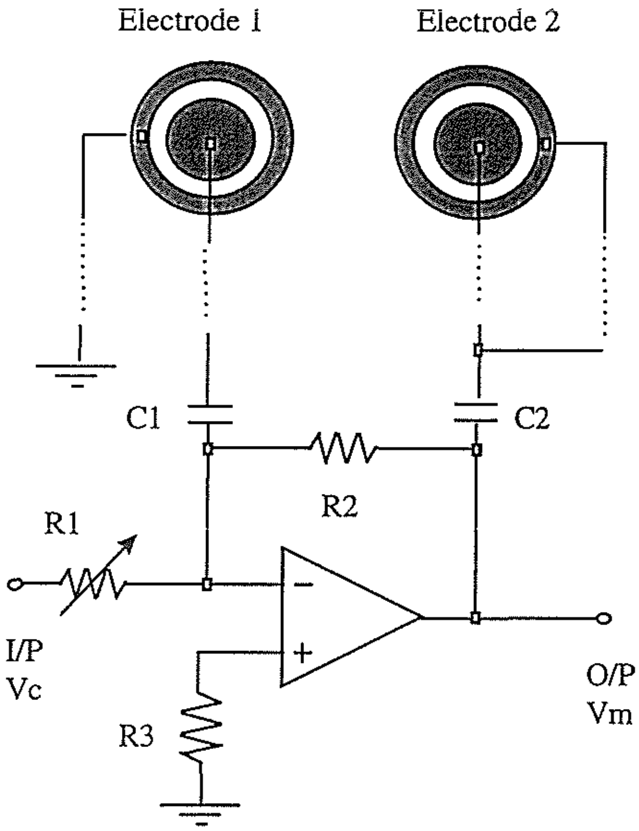


Fig. 4.10 Configuration of electrodes for maximum sensitivity

Chapter 5

THE SIGNAL ACQUISITION AND DISPLAY UNIT

The displaying of the impedance variation waveform can be done in the following three ways

- Interfacing the impedance detector module with a computer using a commercially available ADC card for capturing the impedance variation waveforms and then displaying the waveforms and the pitch on the computer screen.

- Designing a dedicated hardware for digitizing the impedance variation waveforms and then transferring the data to the computer through a parallel or serial port for further processing.

- Interfacing the impedance detector module to a graphics display by digitizing the impedance variation waveforms obtained. The digitizing of these waveforms can be done by building a dedicated ADC system using a microcontroller.

In the first option, the process of digitizing depends totally upon the ADC card available. The instrument loses its portability and also the cost of the overall system becomes high. In the second option, the displaying of the waveforms depends upon the rate of data transfer from the digitizing hardware to the computer, and hence we have to use enhanced printer port for interfacing. The third option makes the glottal pitch extractor completely portable. This option involves the design of the digitizing circuitry using an ADC and the displaying module using a graphics display controlled through a microcontroller. This option makes the glottal pitch extractor a real time system for monitoring the pitch an the impedance variation waveforms. Hence, the third option was chosen for implementing the glottal pitch extractor.

The impedance detector module is interfaced to a graphics display by digitizing the impedance variation waveforms obtained. The digitizing of these waveforms is be done by building a dedicated ADC system using a microcontroller. The idea is to develop a 8-channel, inexpensive digital storage oscilloscope which can display any waveform with a frequency of upto 5 kHz.

The design of the Signal Acquisition and Display unit (SAD) is explained in the following sections.

5.1 DESIGN OF THE SAD UNIT

The SAD unit has been designed and developed such that it can function as an inexpensive 8-channel digital storage oscilloscope. Fig. 5.1 shows the overall block diagram of the SAD unit. The design is divided into two parts : analog section and digital section. The analog section consists of an ADC 0809 from National Semiconductor, with built-in 8-channel multiplexer. An input overvoltage protection circuitry is designed for the same.

The digital section is built using a microcontroller ATMEL 89C55-24PC with 20 K internal EEPROM for program memory, 256 bytes RAM, and four 8-bit I/O ports. In addition to this, it consists of a 240x128 graphics display (Oriole OGM-24011), 2 K bytes RAM data memory, and a 4x4 keypad. The microcontroller controls the analog section for digitizing the inputs, with a maximum sampling rate of 20 k samples/sec. The digitized signal are displayed on the graphics display and can be stored into the data memory. All the operation of the SAD unit can be controlled by the keypad. A RS-232 serial interface has been provided for downloading the digitized samples to a computer for further processing and also a HP-7475 plotter can be connected to the SAD unit through this interface for taking a hardcopy of the stored results.

5.2 ANALOG SECTION

The analog section consists of ADC0809 for digitizing the incoming signal under the control of a microcontroller. The ADC0809 is an 8-bit successive approximation type

ADC with conversion time of approximately 100 μs (with a successive approximation register (SAR) clock frequency = 600 kHz), ± 1 lsb error, single supply operation (5 V), and a current drain of 5 mA. It has an internal 8-channel multiplexer with address logic and can be easily interfaced to 8-bit microprocessors bus [21].

Fig. 5.2 shows the interfacing of the A/D converter with rest of the circuitry. The data bus of the ADC is connected to port #0 of the microcontroller and the address and control lines are distributed among ports #2 and #3. The reference voltage of the ADC is 5 V. The clock for the ADC is given through the ALE line of the microcontroller. The microcontroller operates at a clock of 16 MHz generating an ALE signal of 2.6 MHz. The operation of the ADC was tested with the clock frequency of 2.6 MHz and it was found to be working satisfactorily. The maximum sampling frequency is 20 kHz in single channel operation, since the ADC conversion time is approximately 50 μs. The sampling frequency reduces to 8 kHz in dual channel operation. This is because of an additional overhead of the ADC multiplexer switching time of 10 μs.

The sampling of the incoming signal is done under the timer control of the microcontroller. By setting an appropriate count in the timer, the sampling frequency can be changed. The timer is operated as a 16-bit timer and hence it can not be used in auto reload mode. Thus it takes 5 machine cycles to reload and start the timer. This delay is thus to be subtracted from the required delay. The timer count can be related to the sampling frequency as follows

$$Tc = \left[\frac{Fc}{12 \text{ x } Fs}\right] - Td$$

$$Td = \frac{5 \text{ x } 12}{Fc}$$

$$\text{Timer count} = 65535 - T_c$$

where, F_s = sampling frequency, in S/s

F_c = crystal clock frequency, in Hz

T_d = time required for reloading the timer

Over voltage protection diodes are connected at the input of each channel as shown in Fig. 5.3, along with a series resistor for current protection.

The analog inputs are applied to the ADC with reference to a generated ground. Fig. 5.4 shows the circuit for the reference ground generation. The output of CA3140, which is at 2.5V, is used as the reference for the incoming signals. Thus the ADC becomes capable of digitizing bipolar input signals.

5.2 DIGITAL SECTION

A microcontroller from ATMEL, AT89C55-24PC [15], compatible with MCS-51 microcontroller family, is chosen for building the digital interface. Some of the main features and specifications of this microcontroller are listed below.

- 20K bytes of on-chip reprogrammable flash memory
- Fully static operation: 0Hz to 24 MHz
- 256 bytes of internal RAM
- 32 Programmable I/O lines
- Three 16 bit timer/counters
- Eight interrupt sources
- Low power consumption:

  Active Mode (12 MHz, Vcc=5V): 125 mW

  Idle Mode (12 MHz, Vcc=5V): 35 mW

  Power Down Mode (Vcc=5V): 100 μW

Fig. 5.5 shows the block diagram for the microcontroller interface. Port #0 is used as the data bus for the ADC, display, and the data memory. Port #1 is interfaced with the keypad and port #2 pins along with some pins from port #3 are used to generate the necessary control signals for rest of the circuitry.

A memory of 2kB (6116-3, access time 30 ns) is interfaced with the microcontroller for storing the digitized analog inputs. Each input channel is allocated a memory of 256 bytes.

Fig. 5.6 shows the interfacing of the memory with the microcontroller. The microcontroller puts the lower eight address lines on port #0 and the remaining three on port #2. A transparent (edge triggered by the rising clock edge) latch (74LS373) is used for demultiplexing the data and address bus for the memory interface [25].

Table 5.1 gives the port connection details of the microcontroller with the ADC, display, and the data memory interface.

A 240x128 graphics display (OGM-24011) [16] is used for implementing the waveform displaying system. It operates on +5V and -12V supply. The display is interfaced to the microcontroller through port #0. It has five control lines (RS, R/W, E, CS, RST) which are connected to port #2. The display uses an external inverter module which converts the 6V DC input into 200V AC for the backlight. The input to the inverter has been obtained from the 12 V supply through a potentiometer which can be adjusted for the desired brightness. Fig 5.7 shows the block diagram of the display interface with the microcontroller. An external potentiometer (50 kΩ) is used for controlling the display contrast. The display contrast voltage is kept approximately at 10 V for normal brightness.

A 4x4 matrix keypad is used for controlling the SAD unit. This keypad is interfaced to the microcontroller via port #1. Keypad scanning is accomplished through a conventional scanned row-column matrix technique, as shown in Fig. 5.8. Keys are scanned by outputting data at microcontroller port pins P1.0 through P1.3 to the rows, and then columns are scanned by reading back the data on port pins P1.4 through P1.7 for decoding the key pressed. Each row is made low, one at a time and the column pattern is read for that particular row. A pressed key has a unique row-column pattern of one row low, one column low. Multiple key presses are rejected by either an invalid pattern or a failure to match for three complete cycles. Each row is scanned at an interval of 1 ms for the entire keyboard. The 1 ms delay between scans is generated by timer T0 in an interrupt mode. The key decoding software incorporates a key debouncing logic with a debounce time of 10 ms.

The instrument can be interfaced with a computer or a serial printer with the help of the serial interface of the microcontroller using the RxD and TxD lines. However, this interface works with TTL levels and hence a TTL to RS-232 converter (ICL232) is used. ICL232 operates from a single +5V supply and generates the RS-232 logic levels (± 12 V). Fig. 5.9 shows the interfacing of ICL232 with the microcontroller.

5.4 CIRCUIT ASSEMBLY AND TESTING

The PCB layout was prepared using Circuit Maker and the circuit has been assembled on a PTH board of size 12x12 cm. The circuit diagram, the PCB layouts, and the connection details for the SAD unit are given in appendix B.

The software developed for the SAD unit is explained in the next section. The circuit operation with the software has been fully tested.

5.5 OPERATION MODES

Program has been written on the microcontroller to provide the operations of two channel real-time / storage scope and data logger. In real-time operation, triggering level can be selected. Stored waveforms can be recalled for display. Storage can be done with up to 100% pre-triggering. The operation features can be listed as the following

- Single / dual channel real-time display
- HOLD mode for measurement of voltage and time on the waveforms using cursors
- Channel store (two memory banks for each channel) with pre-trigger facility and recalling of the stored waveforms
- Data logger mode (sample, acquire, and output over serial port)
- Selection of single / dual channels from any of the eight channels
- Changing input voltage scale
- Changing time scale
- Shifting y-position of channel A
- Selection of triggering level (channel A)
- RS-232 serial interface for data transfer to the computer or HP-7475A plotter

The SAD unit has three main operating modes namely run, hold, and special function. These modes are explained below.

RUN mode : In this mode the unit continuously samples the incoming data and displays it depending upon the settings made for voltage scaling, time scaling, triggering level, and channel selection. At power up or after hardware reset, the unit is configured in this mode for single channel (channel A) operation. Channel #1 is displayed with a vertical scaling of 1 V/div and a time scale of 1 ms/div. The triggering level is set at 0 V on input channel #1. This modes provides the functions for setting the input voltage scale, time scale, triggering level, specifying pre-trigger data percentage, and mapping the input channels on channel A and B for display.

HOLD Mode : This is a special mode in which after a signal is captured and displayed, it can be HOLD till the user wants to return to the RUN mode. Two cursors are provided for the two channels displayed on the screen. These cursors can be moved individually along the waveforms. The cursor positions are displayed in the x and y directions. From these cursor positions, the voltage and time can be calculated using the current voltage and time scales, for making measurements on the waveform. This mode can be activated by pressing the HOLD / RUN key. Pressing the key again, restores the RUN mode.

SPECIAL FUNCTIONS Mode : This provides different special modes available namely store, recall, data logger, and plot. This mode is activated by pressing the SP. FUNCT. Key. When this key is pressed a menu appears displaying all these modes. The MENU key can be used for selecting a required mode along with the ENTER / EXIT key. These modes are explained below.

STORE : The unit can store upto 1 kB of data. The data can be stored for a single or dual channel. Each channel i.e. channel A and channel B are allocated 256 bytes of data memory. When this option is selected from the special functions menu, a sub-menu appears on the display giving an option of Mem 1 or Mem 2 for storing the waveforms currently present on the display.

RECALL : The waveforms stored in Mem 1 and Mem 2 can be recalled using this option. When a memory is recalled the unit waits for the ENTER / EXIT key to be pressed for returning back to the RUN mode.

DATA LOG : This mode is used for continuously transferring the digitized incoming signals to a computer through the RS-232 serial interface with 9600 baud, 8 data bits, no parity, and 1 stop bit. It has a facility to transfer single or dual channel depending upon the settings done for the channel selection in the RUN mode. The voltage and time scaling settings remains the same as for RUN mode. No triggering is used in this mode. When this option is selected, the unit asks for the time duration for data transfer. The default time duration is one minute and the maximum time duration is of 10 minute. The time duration can be cycled in steps of 1 minute with the help of the TIME/CUR key. After setting the time, the ENTER/EXIT key can be pressed for starting the data logger mode. After initiating this mode, the unit sends the number of channels, the sampling rate, and the total number of samples before transmitting the actual data. Each sample is sent as one byte. After the specified time duration, the RUN mode is started.

PLOT : The unit can be interfaced to a HP-7475A plotter over the 3-wire serial interface, for taking a hardcopy of the stored waveforms from any of the two memory banks available. The data is transferred at 1200 baud to the plotter without any handshaking. When the plotting is over the RUN mode is automatically started with the settings made before plotting the data.

Fig. 5.10 - 5.17 indicates the operation of the SAD unit in different modes.

5.6 FUNCTIONS ASSIGNED TO THE KEYS

Fig. 5.18 shows the keypad layout for the SAD unit. The functions of the keys available is explained below.

CH. A : In the RUN mode, the input channel to be connected to display channel A can be selected using this key. It cycles from 1 to 8.

CH. B : In the RUN mode, the input channel to be connected to display channel B can be selected using this key. It cycles from N, 1 to 8; where N means not connected (i.e. single channel operation).

VOLT (Up / Down) : In the RUN mode, these two keys are used for scaling the input. The different scales available are x1, x2, x5, and x10. The display indicates the voltage scales as Y_d = 0.2 V, 0.5 V, 1 V, and 2 V indicating volt / division.

TIME / CUR (Left / Right) : In the RUN mode, these two keys are used for changing the time base of the inputs. The available scaling factors are X_d = 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000 ms/div. In the HOLD mode, these keys are used for shifting the cursors along the channel waveforms. In the DATA LOGGER mode these keys are used for setting the time duration for data transfer.

MOVE (Up / Down) : In the RUN mode, these two keys are used for shifting the y-position of display channel A. Position of channel B can not be changed.

TRIG (Up / Down) : In the RUN mode, these two keys are used for changing the triggering level for input connect to display channel A. The available triggering levels are from +2.5V to -2.5V in steps of 0.5V. The level is not affected by voltage scale (VOLT) setting.

PRE TRIG : In the RUN mode, a pretrigger data percentage can be set using this key. The incoming data is captured by the SAD unit and stored temporarily in 256 bytes of RAM before displaying it. By setting the pretrigger percentage, the triggering instant can be set such that the signal data is displayed with the amount of pretrigger data specified. The options available are 0%, 25%, 50%, 75%, and 100% pretrigger data. Depending on the option selected, the triggering of the SAD unit changes and a dotted line is displayed for indicating the triggering instant of the SAD unit on channel A.

HOLD / RUN : This key can be used for changing the operation between RUN and HOLD mode.

SP. FUNCT. : This key is used for selecting the special modes available.

MENU (Up / Down) : These keys are used for selecting a particular option in the special functions menu and the resulting sub-menus. A cursor is provided for indicating the choice of an option within a menu.

ENTER / EXIT : This key is used for selecting the options given in the special functions menu and all other sub-menus. The EXIT function is activated only in the RECALL and DATA LOGGER mode.

Table 5.1. Port connections of the microcontroller

I/O Port lines	Function assigned
P0.0 to P0.7	Data bus of the ADC, display, and memory, AD0..AD7 for memory
P1.0 to P1.3	Write to the keypad row lines
P1.4 to P1.7	Read from the keypad column lines
P2.0	Display control line, RS, Memory Add. A8
P2.1	Display control line, R/W, Memory Add. A9
P2.2	Display control line, E, Memory Add. A10
P2.3	Display control line, CS
P2.4	Display control line, RST
P2.5	ADC status line, EOC
P2.6	ADC control line, START / ALE
P2.7	ADC control line, OE
P3.0	Serial interface, RxD
P3.1	Serial interface, TxD
P3.2	ADC select line, A2
P3.3	ADC select line, A1
P3.4	ADC select line, A0
P3.5	Not used
P3.6	Memory control line, WR
P3.7	Memory control line, RD

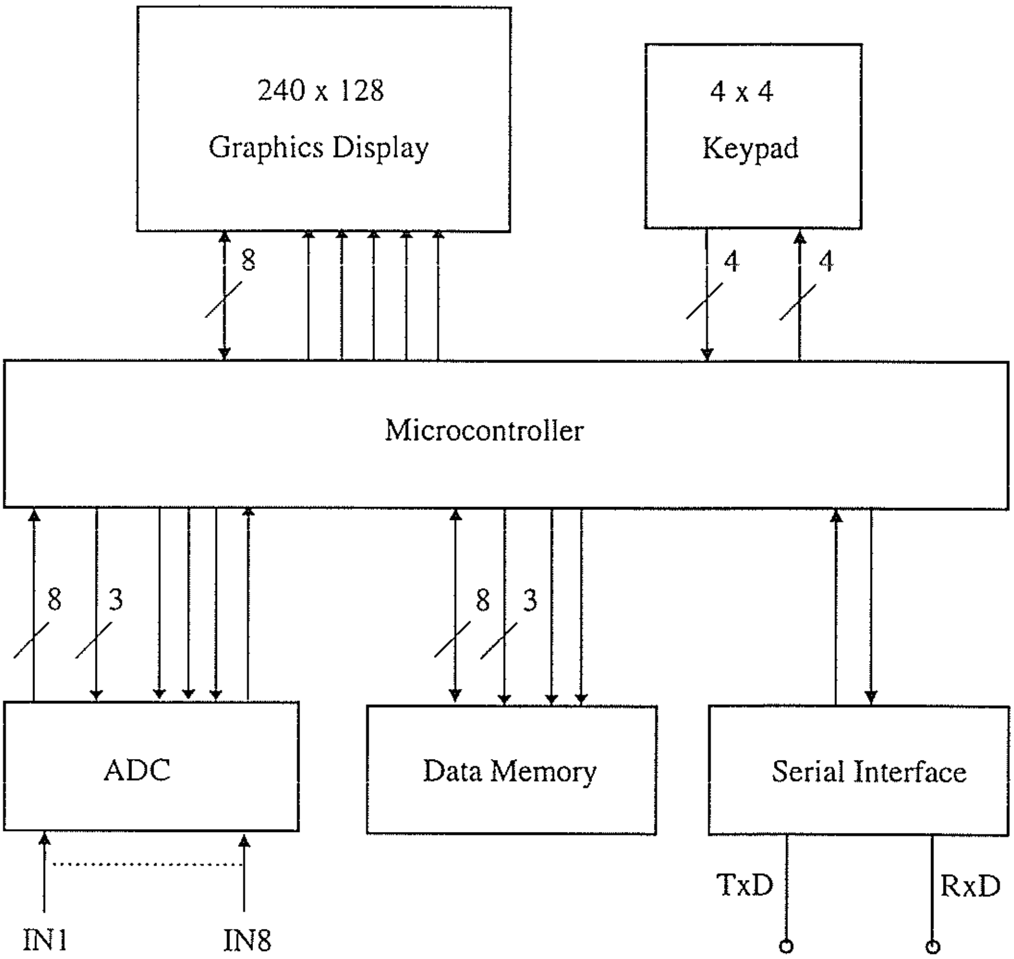


Fig. 5.1 Block diagram of the SAD unit

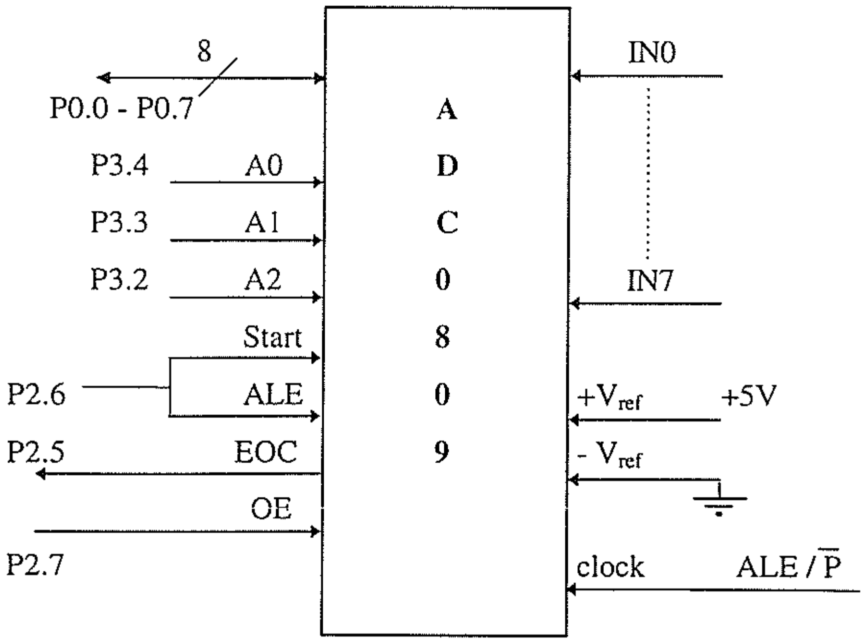


Fig. 5.2 A/D Converter

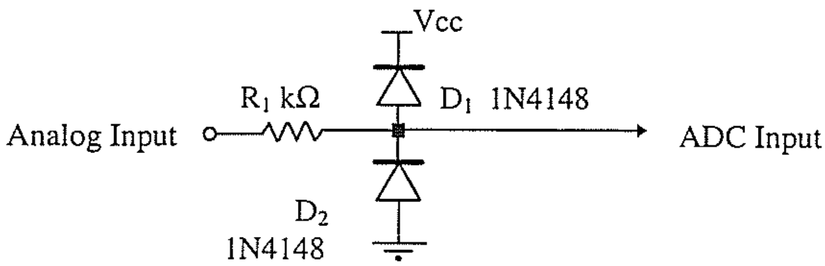


Fig. 5.3 Overvoltage protection for the ADC input

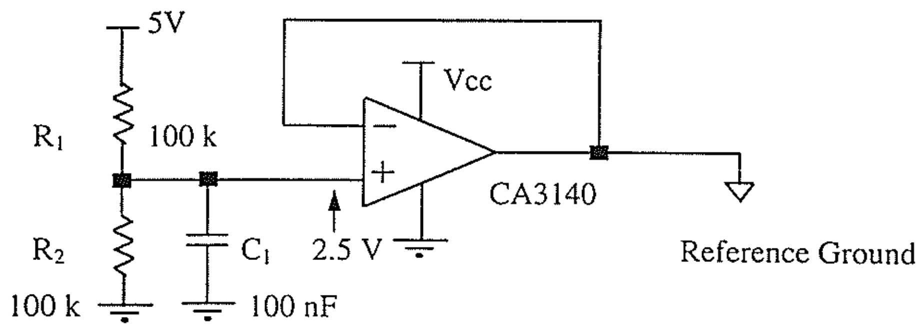


Fig. 5.4 Circuit for generating reference ground

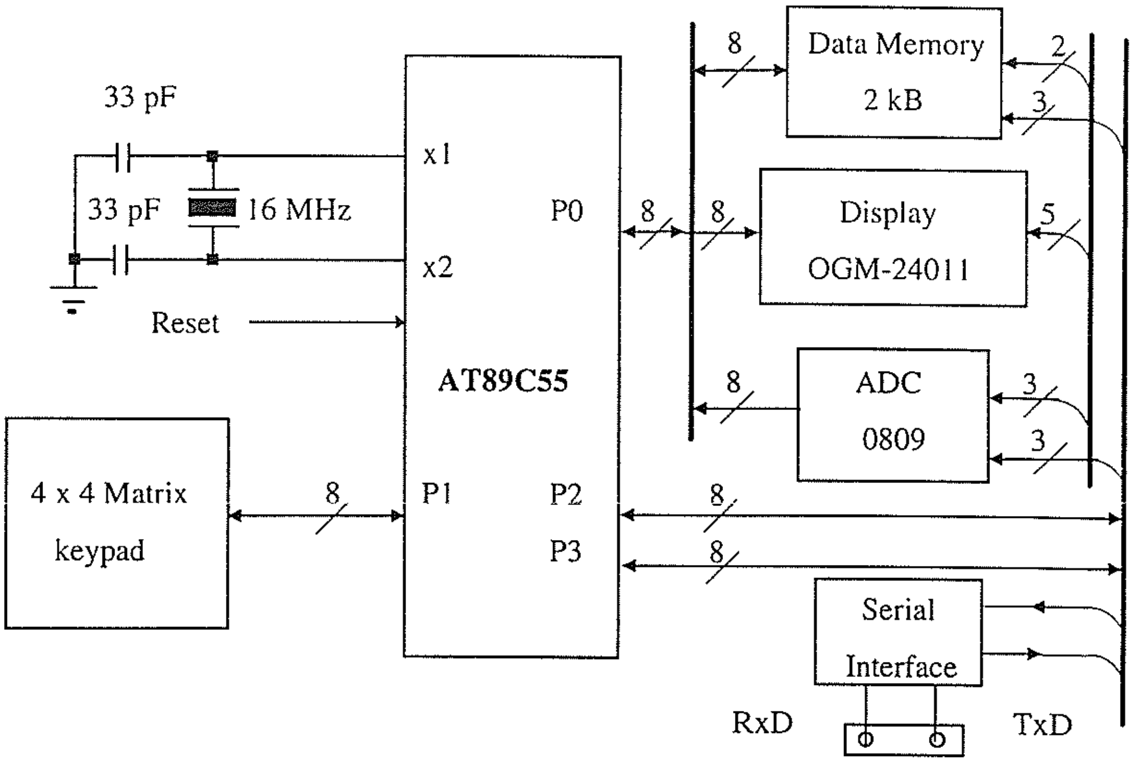


Fig. 5.5 Microcontroller Interface

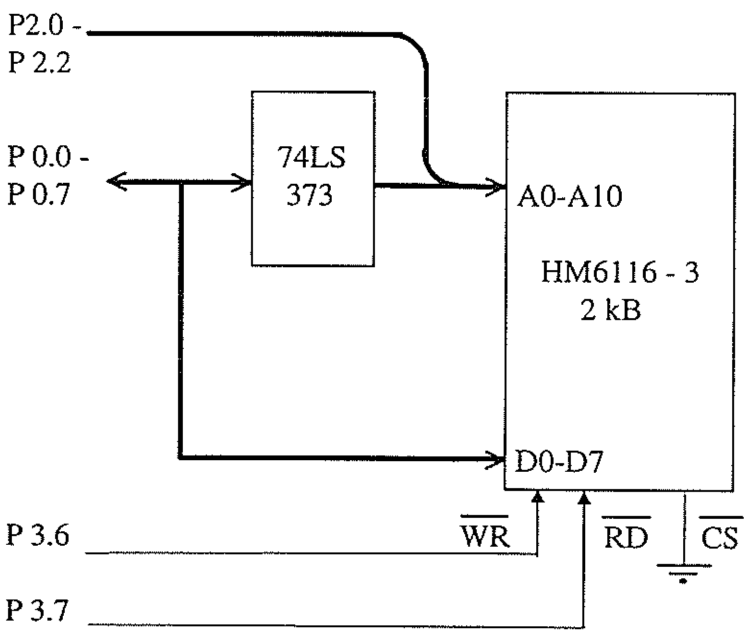


Fig. 5.6 Memory Interface

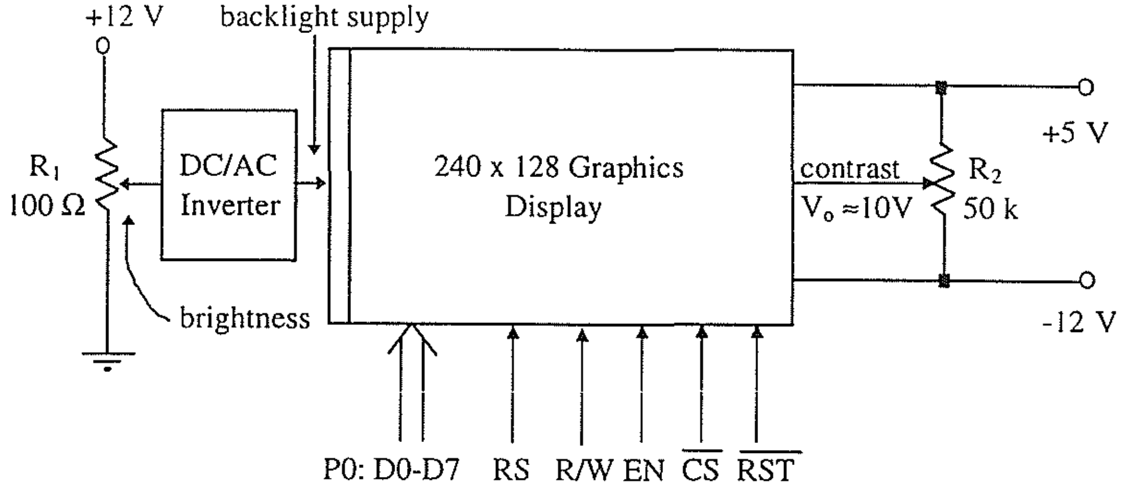


Fig. 5.7 Display Interface

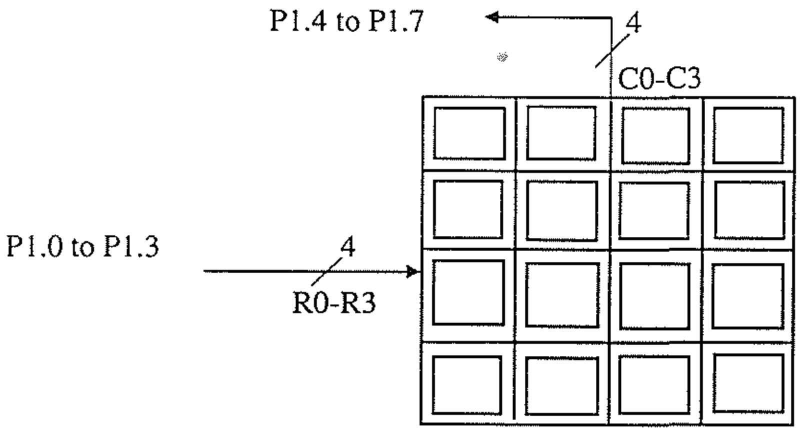


Fig. 5.8 4 x 4 Matrix keypad

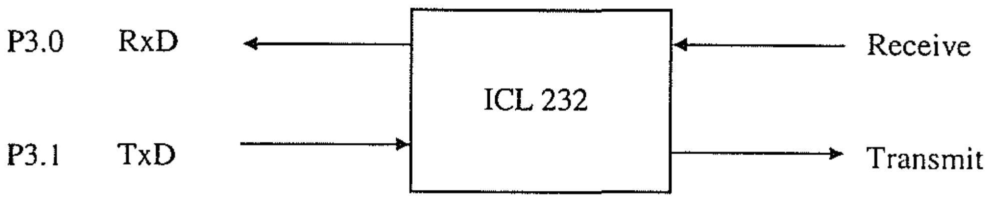


Fig. 5.9 TTL to RS-232 level converter

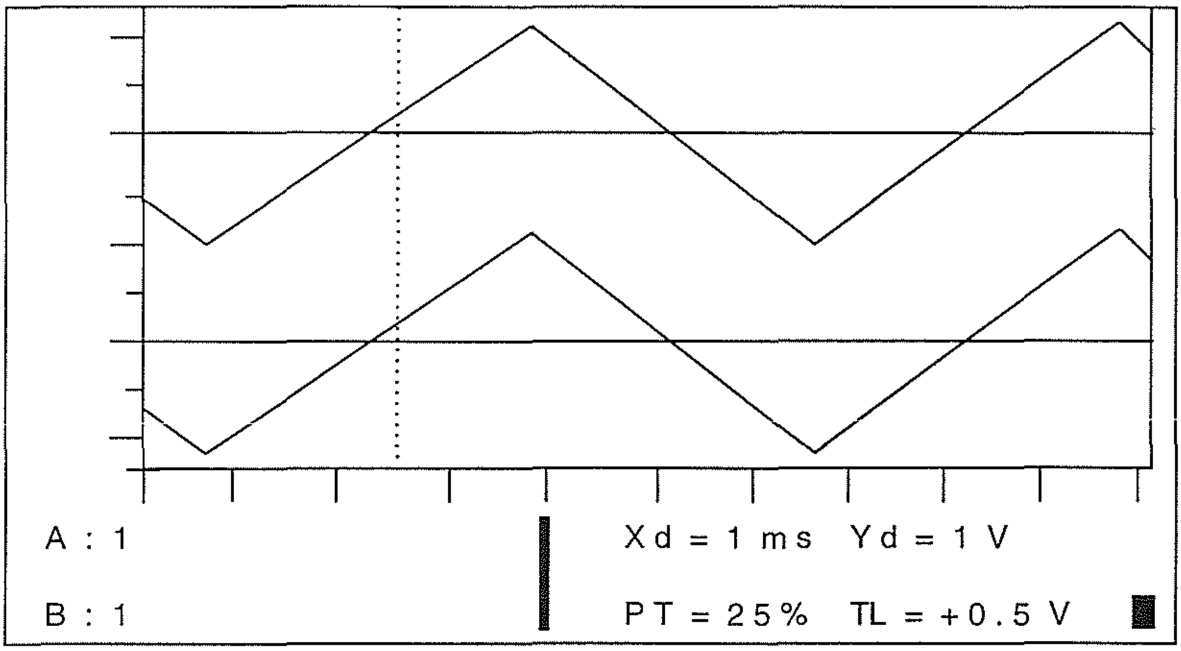


Fig. 5.10 RUN Mode

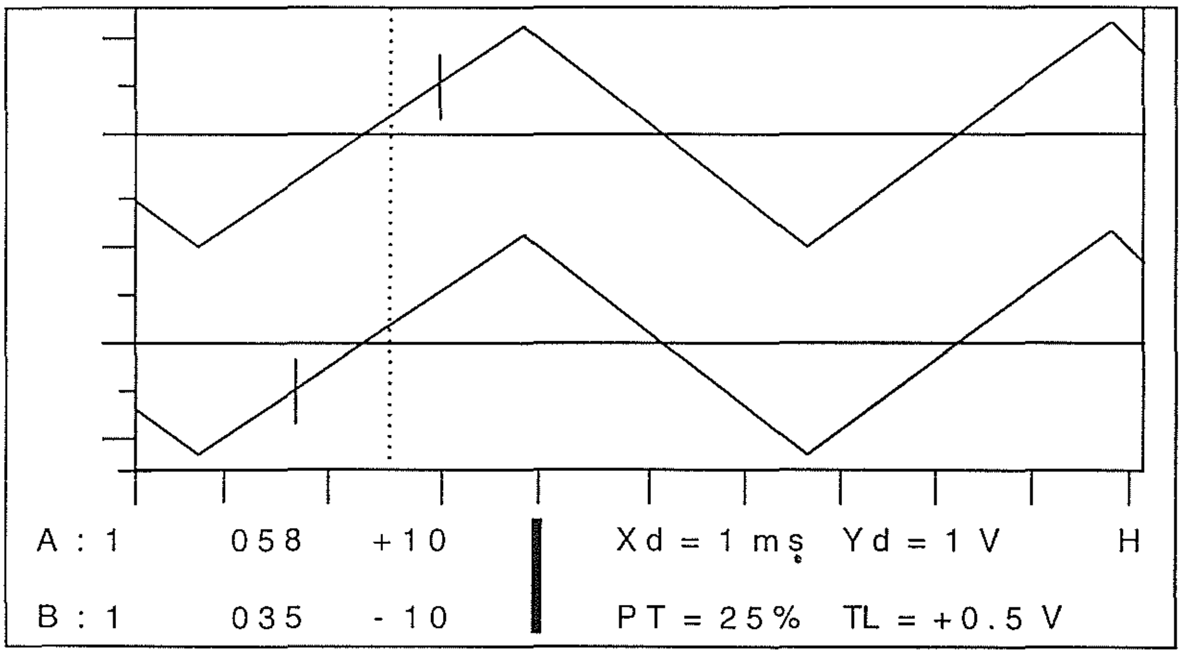


Fig. 5.11 HOLD Mode

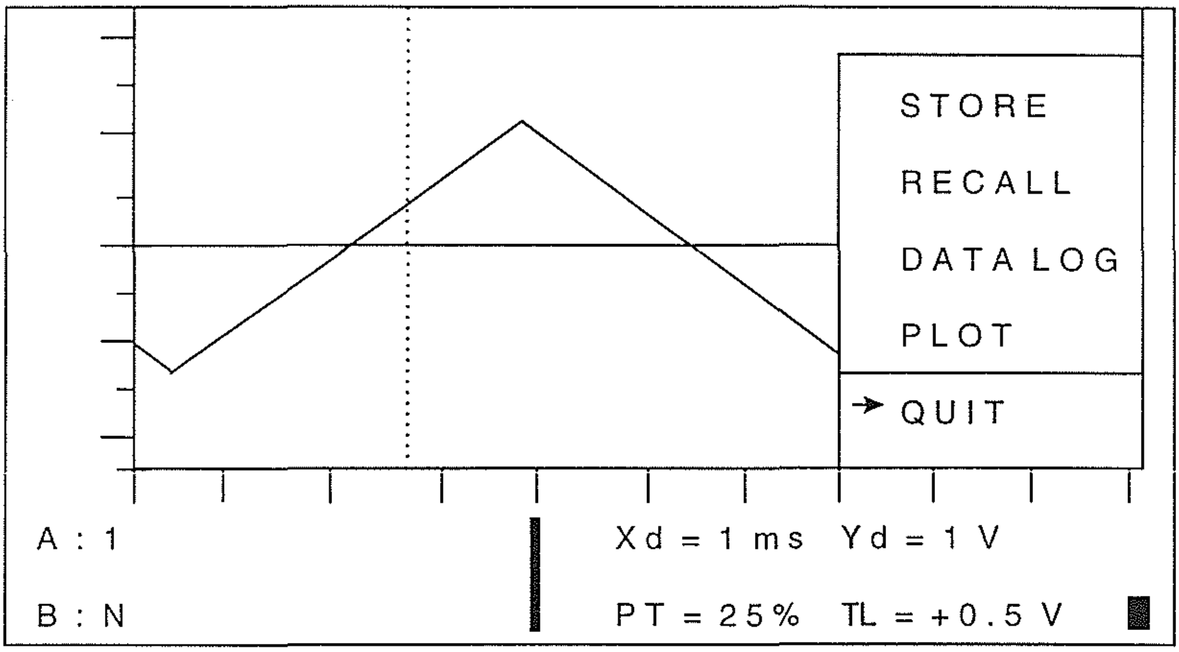


Fig. 5.12 SP. FUNCTION Mode

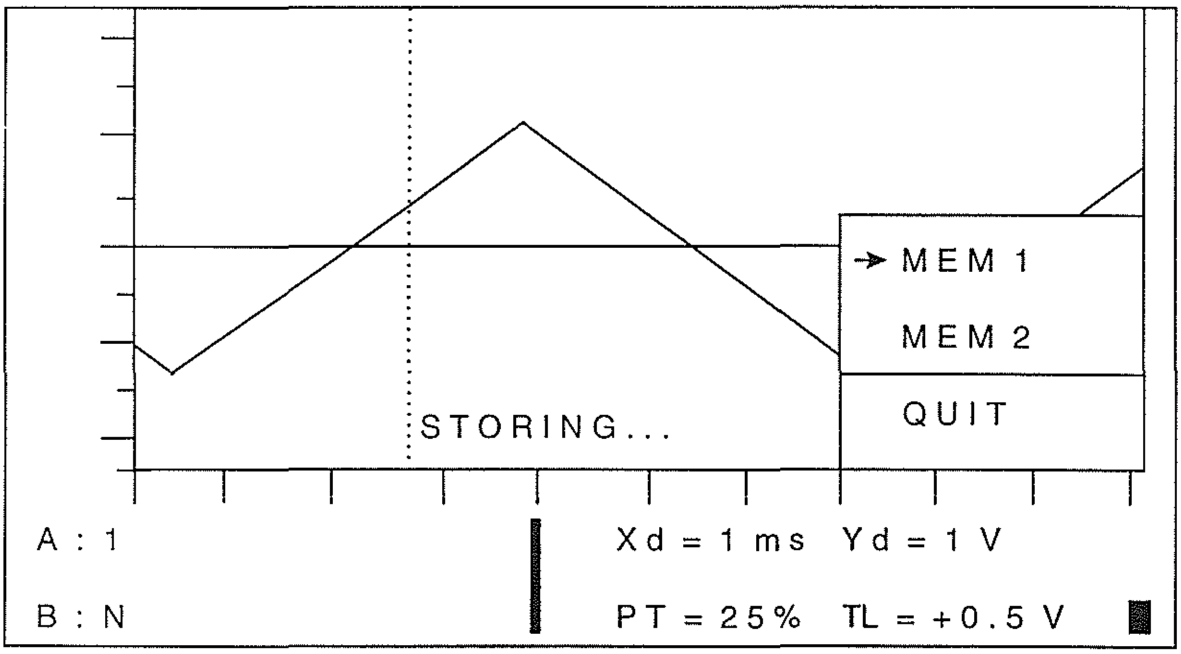


Fig. 5.13 Sp.function mode with STORE option selected

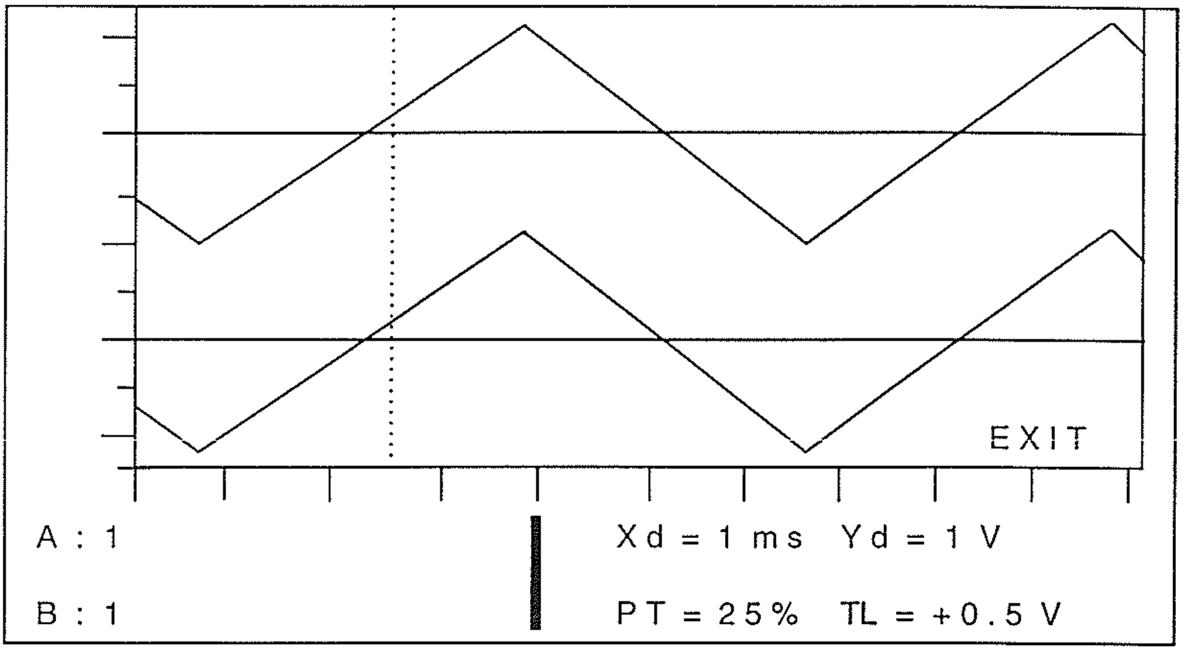


Fig. 5.14 Sp. function mode with RECALL option selected

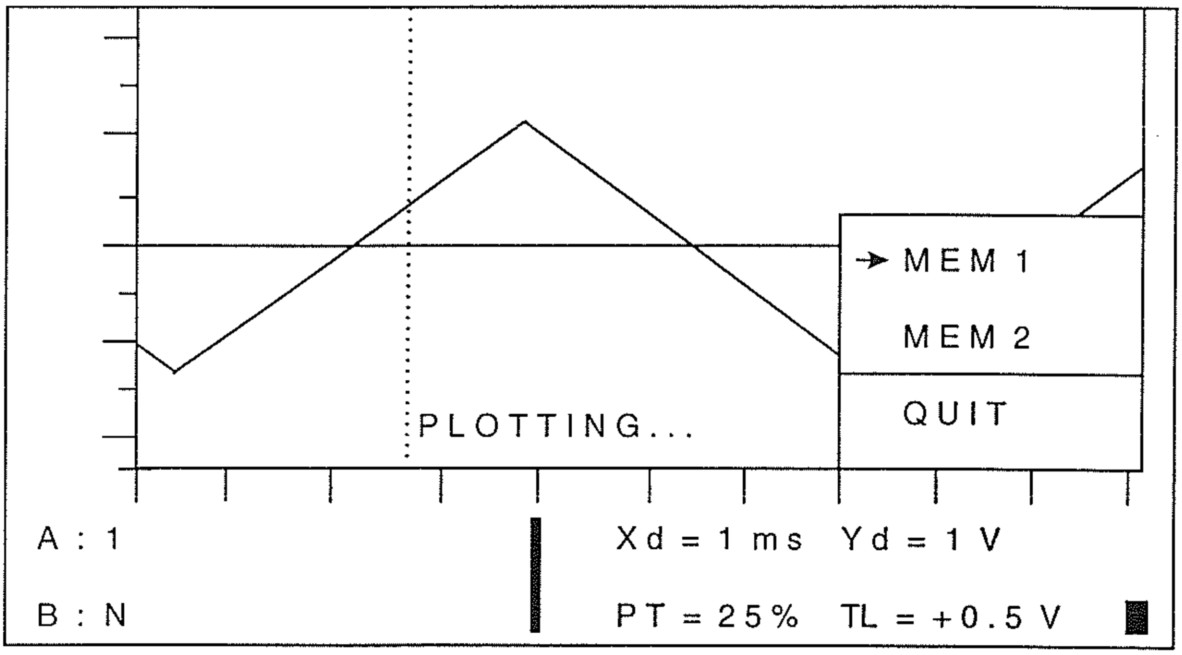


Fig. 5.15 Sp. function mode with PLOT option selected

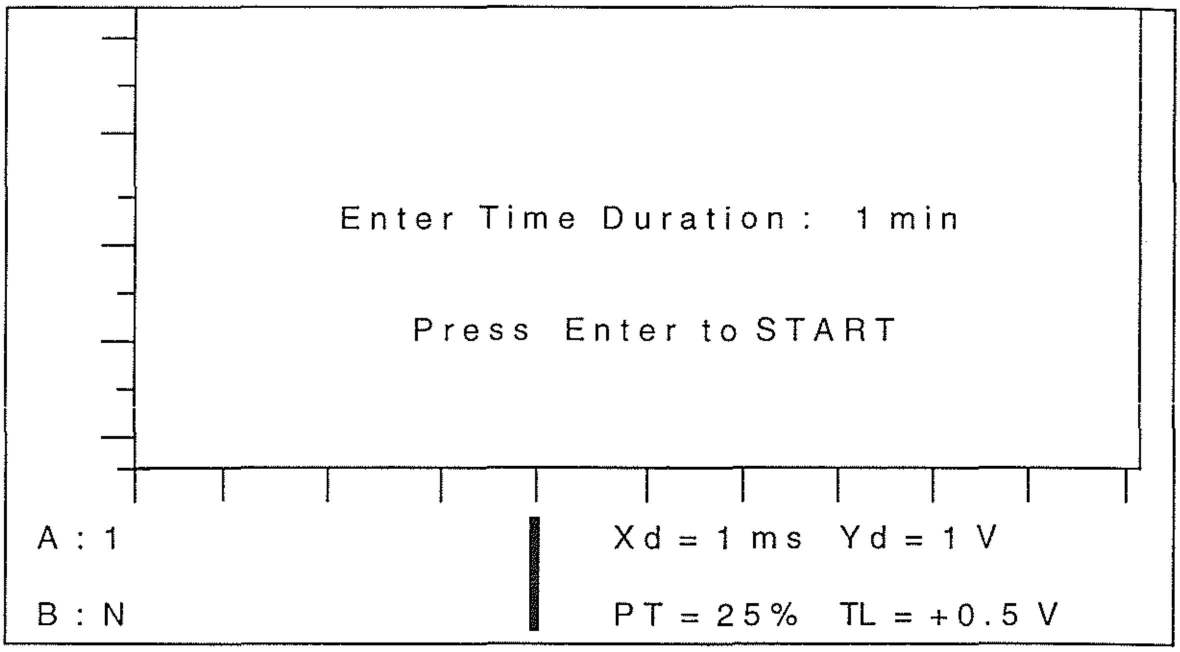


Fig. 5.16 Sp. function mode with D LOG option selected

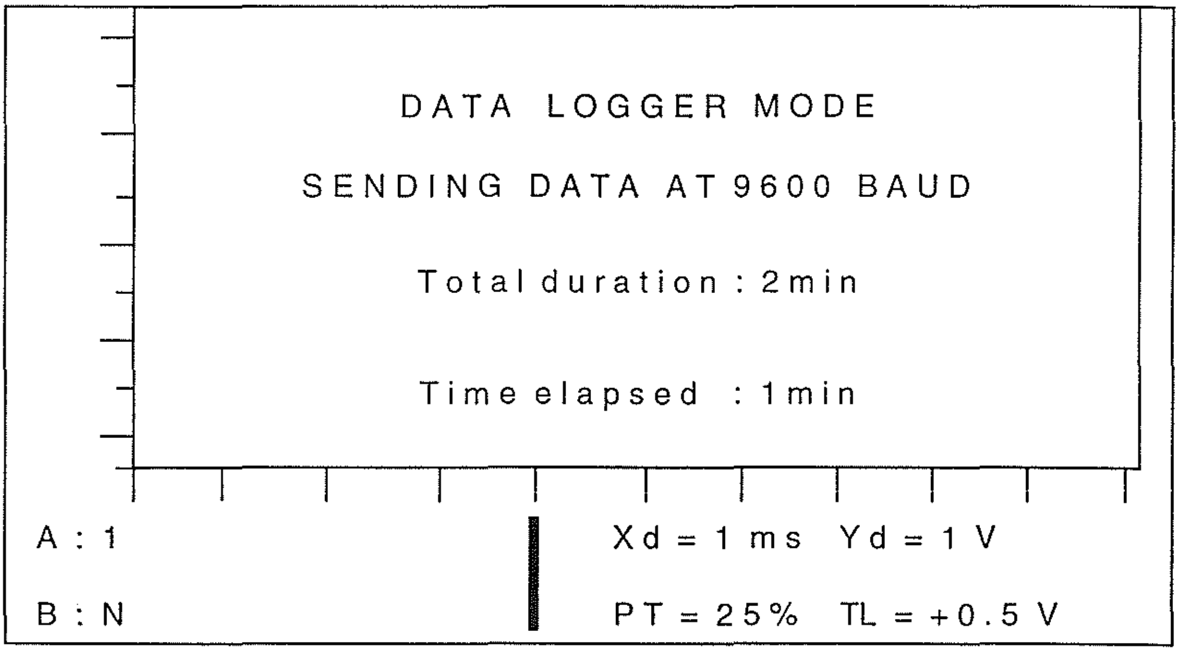


Fig. 5.17 Data logger mode

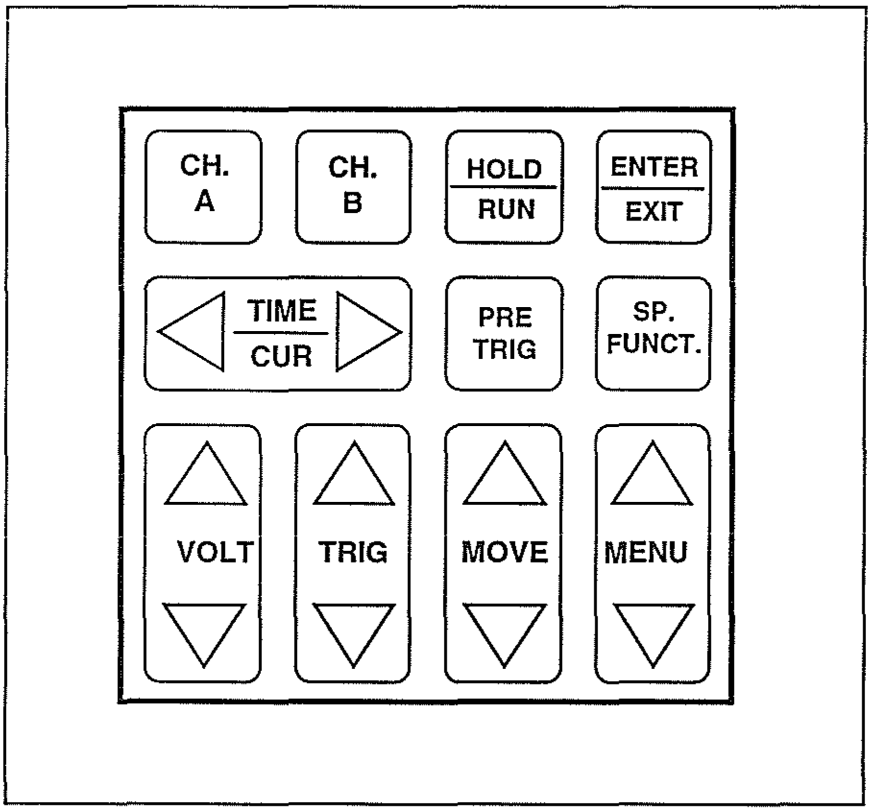


Fig. 5.18 Keypad layout of the SAD unit

Chapter 6

SIGNAL ACQUISITION AND PITCH EXTRACTION

The SAD unit has a limited amount of data memory and hence can not be used for storing large amount of data, especially for capturing the impedance variation waveform for a longer duration. Hence, it can not be directly used for calculating the pitch and the F_x histograms, which require large amount of data capture and processing. Thus in order to calculate the pitch and F_x histogram, a signal acquisition and pitch extraction software is developed on a notebook PC for interfacing DAQ-700 card from National Instruments as a signal acquisition system. The card is interfaced to the PC through a PCMCIA (Personal Computer Memory Card Interface Association) port.

This system is developed in such a way that it can be used for recording signals from two channels for various time durations; upto a maximum of 10 minutes with a variable sampling rate between 1 kS/s to 20 kS/s. The data transmitted by the SAD unit in the data logger mode can also be captured by this system for further processing and display. The details of this system are explained in this chapter.

6.1 HARDWARE USED

The system is developed on a notebook 486DX-2 PC with 4 MB RAM and two PCMCIA slots. The main component used here is the DAQ-700 data acquisition card. DAQ-700 is a low-powered, digital, and timing I/O card for computers equipped with a PCMCIA type II slot. The card contains a 12 bit successive approximation A/D converter with a conversion time of 1 μs, ± 1 lsb accuracy, 8 differential or 16 single ended inputs, 8 lines of digital inputs, and 8 lines for digital output. The card also has a clock generator and a programmable counter / timer for timing I/O. The card interfaces with the inputs and outputs through a 50 pin connector.

The card is software configurable. It can be configured for single ended or bipolar input configuration. The maximum sampling rate of 100 kS/s. The input range is also programmable from ± 2.5 V to ± 10 V. The developed system uses a fixed input range of ± 10 V and a programmable sampling rate with a maximum of 20 kS/s.

6.2 SOFTWARE DETAILS

The software is written in C on a Turbo C++ compiler. The overall features provided by the software are as follows

- Data capture from the PCMCIA port with variable sampling rate and capture duration
- Data capture from the RS-232 serial port interfaced with the SAD unit
- Off-line waveform display with scroll, zoom, and channel #1 y-position shift facility
- Real-time data capture and display for two channels
- Analysis of the captured data which includes calculation of the pitch up to 10 kHz, single or triple period F_x histogram, and F_x plot

The software has three main operating modes namely data capture, waveform display, and analysis. The details of these modes are given below

DATA CAPTURE Mode

The data capture mode is subdivided into two modes namely; data capture from the PCMCIA port that is the DAQ-700 card and the data capture from the RS-232 serial port.

Data Capture from PCMCIA Port

When the PCMCIA port option is selected, the card is detected and configured automatically. If the card connection is not proper, an error message is displayed and the user is not allowed to proceed further. After the proper initialisation of the card, the user is provided with different settings for data capture. The number of input channels (single / dual), channels polarity (single ended / differential), sampling rate (1 kS/s to 20 kS/s), time duration (maximum of 10 min. depending upon the sampling rate), and the file name for storing the data can be specified by the user. When all the settings are done, the system waits for a key to be pressed in order to start the data capture. During data capture, if the time specified is more than 1 sec., it is displayed on the screen along with the number of samples captured and the sampling rate. The data file is stored in binary format. The number of channels, sampling rate, and the total number of samples are stored in the beginning of the data file. Each sample is stored as two bytes. When the data capture is over, the system returns to the main modes selection menu.

Data Capture from Serial Port

In the serial capture mode, a special software for serial interface named PROCOMM, is initiated for capturing and storing the data to a file specified by the user. The data is captured from serial port 2 (COM2) with a setting of 9600 baud, 8 data bits, no parity, and 1 stop bit. The SAD unit also transfers the data in the same format. When the transfer is initiated from the SAD unit, it first transfers the number of channels followed by the sampling rate and the total number of samples. The samples are then sent continuously as one byte per sample. The downloaded data from the SAD unit is stored in a file in ASCII format. After the data is stored, again the main modes selection menu appears on the screen.

WAVEFORM DISPLAY Mode

The waveform display mode is also subdivided into three modes namely; off-line display from the PCMCIA port, off-line display from the serial port, and real-time display from the PCMCIA port.

Off-line Display

The data captured from the PCMCIA or the serial port can be displayed using the off-line display modes. Since the formats of the data files for the PCMCIA and serial port are different, two off-line display modes are provided. The user has to specify the file name to be displayed in these modes. The validity of the file specified is checked before displaying it for proper operation. If the file specified is invalid, the user is given an option of continuing further or returning to the main menu. When a proper file is specified, the data is displayed on the screen. The user is provided with the following options for specifying the display characteristics

UP arrow	:	shift the channel #1 upwards
DOWN arrow	:	shift the channel #1 downwards
LEFT arrow	:	scroll the display to the right of the screen
RIGHT arrow	:	scroll the display to the left of the screen
N	:	starting sample number of the display
X	:	zoom factor for the time axis
ESC	:	return to the main menu

Real-time Display

In the real-time display mode, the DAQ-700 card is detected and initialised in the bipolar differential input mode, $\pm$ 10 V input range, dual channels with a sampling rate of 20 kS/s. After the proper initialisation of the card, the data from the first two channels i.e. channel #0 and channel #1 of the card is captured and displayed continuously like in an oscilloscope. The mode continues till the user presses the ESC key for returning to the main menu. The functions assigned to different keys in this mode are as follows

UP arrow	:	increase voltage scaling
DOWN arrow	:	decrease voltage scaling
LEFT arrow	:	decrease sampling rate
RIGHT arrow	:	increase sampling rate

ESC	:	return to the main menu

ANALYSIS Mode

In this mode the data captured is processed for calculating the pitch and displaying the F_x plot and the F_x histogram.

F_x Plot

In this option, all the instantaneous pitch values v/s time are plotted. The functions assigned to different keys in this mode are as follows

LEFT arrow	:	scroll the display to the right of the screen
RIGHT arrow	:	scroll the display to the left of the screen
ESC	:	return to the main menu

F_x Histogram

The user is given an option of plotting single or triple period F_x histograms. The F_x histogram is plotted on the screen with the pitch on the x-axis and the % probability on the y-axis. The user can return to the main menu by pressing ESC key.

6.3 PITCH CALCULATION AND HISTOGRAM PLOTTING

The algorithm developed for calculating the pitch uses period by period analysis method for detecting the pitch periods. The number of samples N, between one zero crossing and its second successive zero crossing are calculated. The pitch is then given as pitch = sampling rate / N. This method is computationally less intensive and hence suited for on-line processing. The algorithm uses a window length of 1000 samples for calculating the instantaneous pitch values. Every time 1000 samples are read from the data file and are processed.

The zero crossings are detected using the centre clipping method as follows

a) The mean of the maximum and the minimum levels of the captured data is found. This is taken as the threshold.

b) One bit quantisation of the sequence $\{y(n)\}$ is done

$$y(n) = 1 \qquad \text{if } y(n) > \text{threshold} + |\,10\% \text{ of maximum level}\,|$$

$$y(n) = -1 \qquad \text{if } y(n) < \text{threshold} - |\,10\% \text{ of minimum level}\,|$$

$$y(n) = 0 \qquad \text{otherwise}$$

c) Zero crossings are detected at the changeover from 0 to 1 or 0 to -1

Every time the pitch value is calculated, it is stored in a file for obtaining the F_x histogram and the F_x plot. This method was found to be appropriate since for calculating the F_x plot and the F_x histogram all the instantaneous pitch periods are required. Thus, here we calculate the time duration between two successive zero crossings for obtaining the instantaneous pitch periods.

The F_x plot is obtained by plotting the instantaneous pitch values v/s time. The pitch axis is calibrated linearly from minimum pitch to maximum pitch.

In case of the single period and triple period histograms, the frequency range of interest is divided into a number of equal bins. The default values of the frequency range is 0 to 1 kHz, and the number of bins is set to 50. Each bin is incremented when a pitch value falls within its frequency range. The content of each bin divided by the total number of pitch values is a measure of the probability distribution. The histogram obtained by this method is known as a single period F_x histogram. In triple period F_x histogram, each bin is incremented only if three successive pitch values fall within its range [8].

Chapter 7
TEST RESULTS

The glottal impedance detector unit and the signal acquisition and display unit have been tested for proper functioning. The demodulator unit of the impedance variation detector was tested with a modulated signal and it was found that for a base impedance of 500 Ω, impedance variation down to 0.6 Ω can be extracted.

The electrodes were made on a glass epoxy double sided PCB and the circuit was tested for extracting the impedance variation waveform. The circuit was tested using a two electrode method with various configurations of the connection of the individual conductors of each probe. Maximum sensitivity was obtained with driven guard rings.

The glottal pitch extractor has been used for recording the speech waveform and glottal impedance variation waveforms of different subjects. Recordings were made for subjects with normal vocal folds as well as subjects with vocal fold disorders. The recordings obtained are given in Figures 7.1 to 7.9. Here, $S(t)$ indicates the speech waveform and $L_x(t)$ indicates the impedance variation waveform. Figures 7.1 to 7.3 shows the recordings done on a male subject (MC) with normal vocal folds for three vowels */a/*, */i/*, and */u/*. For all the three vowels, we see that glottal pulses are qualitatively similar to those reported by others earlier [5], [6], [8], [9], [10], [17], and to the model waveform of Fig. 3.1. The glottal pulses in the vowel are synchronized with the valley of the impedance waveform. The pitch period is estimated as 5 ms. Figures 7.4 to 7.7 shows the recordings done on a female subject (DS) with normal vocal folds for the same vowels. We see that waveforms are qualitatively similar, but the pitch period has decreased to 4 ms. Figures 7.7 to 7.9 shows the recordings done on a male subject (AK) undergoing treatment for paralysis of right vocal fold. We see that the average pitch period for this subject is 6 ms. We see that the waveshape of the L_x waveform for this subject is very different from those for the other two subjects. The top of the L_x

waveform is relatively flattened, and a comparison with the model in Fig. 3.1 suggests that this could be due to the vocal folds remaining apart for a longer duration, which could be caused by the paralysis of one of the vocal fold. We see that this method can only tell us about the net state of the contact phases, and not about the movement of the individual folds.

Fig. 7.10 shows a speech waveform $S(t)$, along with its wideband spectrogram for the utterance "...*atma amar hai*...". The spectrogram shows various formants present in the speech waveform. The vertical lines, known as striations, shows the variation in the pitch periods. Fig. 7.11 shows the impedance variation waveform $L_x(t)$, along with its wideband spectrogram for the same utterance. Here we see that the formants present in the speech waveform are absent and only the fundamental frequency is obtained. There is however a continuous noise during the unvoiced period but there is no variation in the striations.

The software developed for signal acquisition and analysis was used for capturing the impedance variation waveform from a male subject with normal vocal folds. Fig. 7.12 shows the single period F_x histogram obtained for a male subject (MC), reading continuous text for a time duration of 1 minute. The average value of pitch for this subject is found to be approximately 150 Hz. Fig. 7.13 shows the F_x histogram for a female subject (DS) reading the same text for a time duration of 1 minute. Here the average value of pitch is increased to 210 Hz. Fig. 7.14 shows the F_x histogram for a male subject (AK) having a paralysis of the right vocal fold. We see that the average pitch is decreased as compared to the other male subject (MC) and it is 130 Hz.

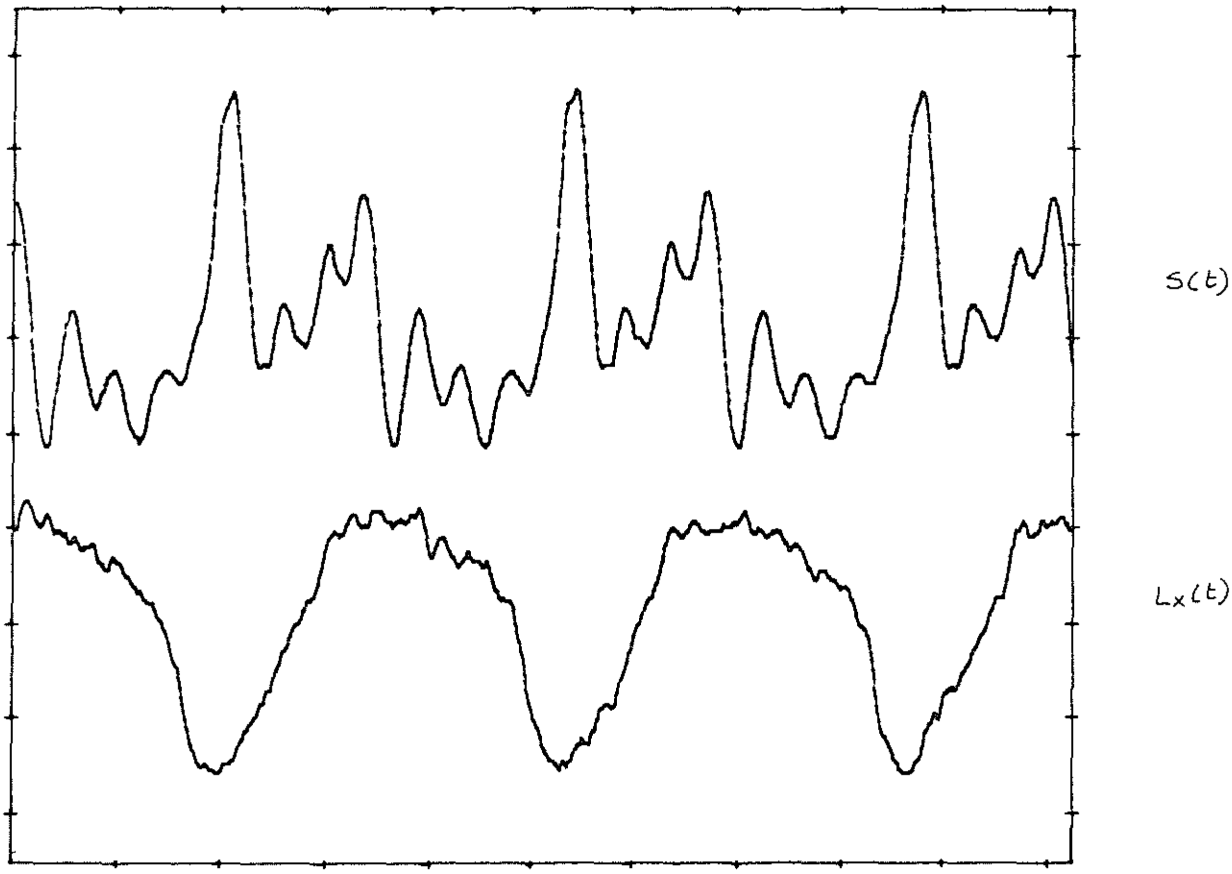
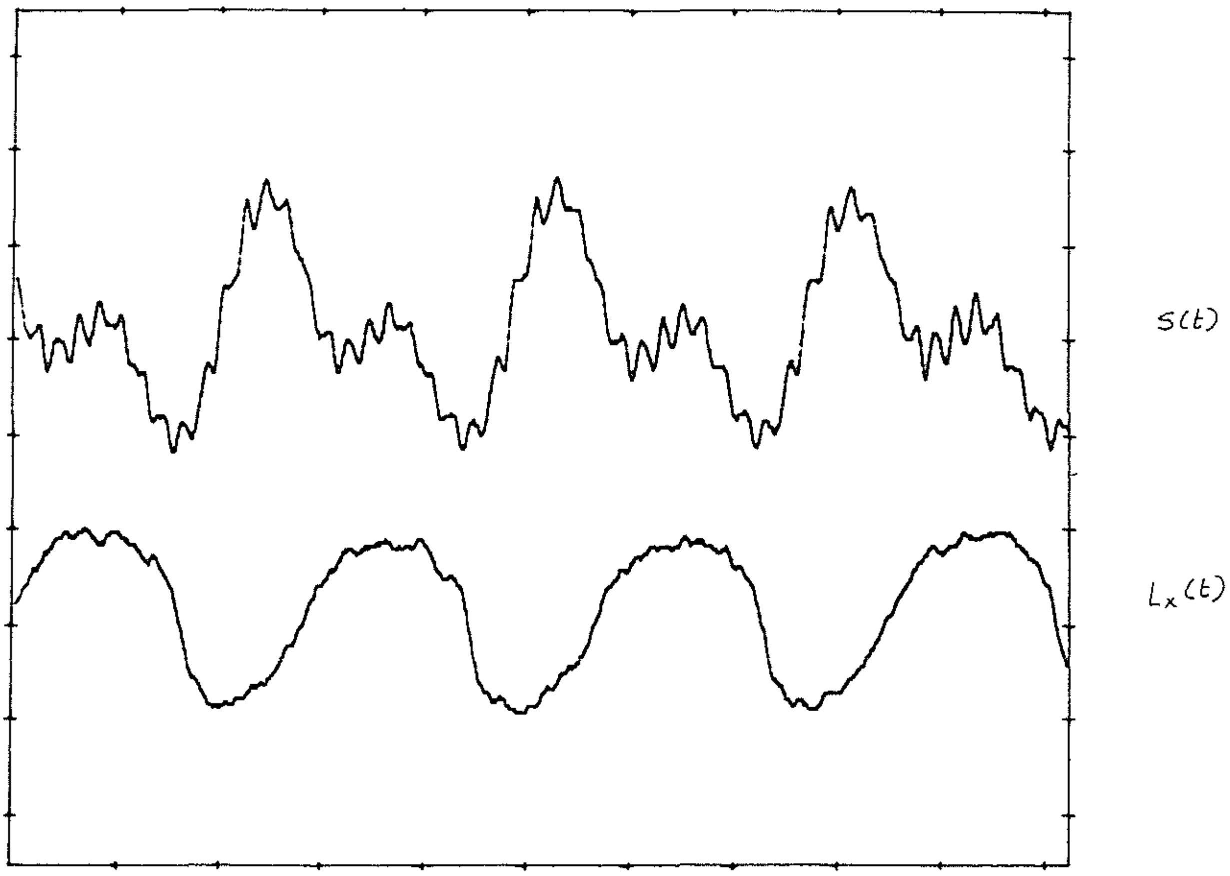
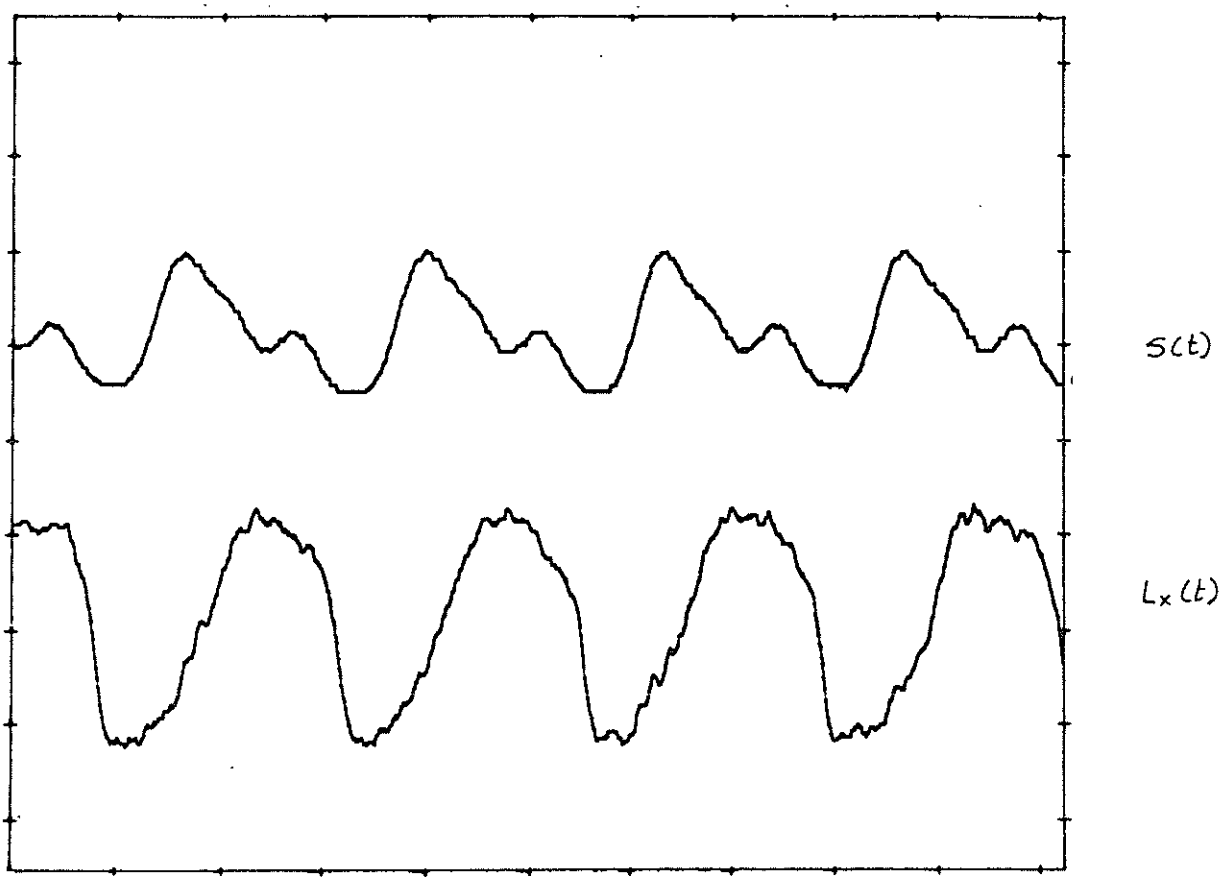


Fig. 7.1 Recordings of speech and the impedance variation waveform during the utterance of vowel /a/ by a male subject (MC) with normal vocal folds.



2 ms/DIV

Fig. 7.2 Recordings of speech and the impedance variation waveform during the utterance of vowel /i/ by a male subject (MC) with normal vocal folds.



2 ms/DIV

Fig. 7.3 Recordings of speech and the impedance variation waveform during the utterance of vowel /u/ by a male subject (MC) with normal vocal folds.

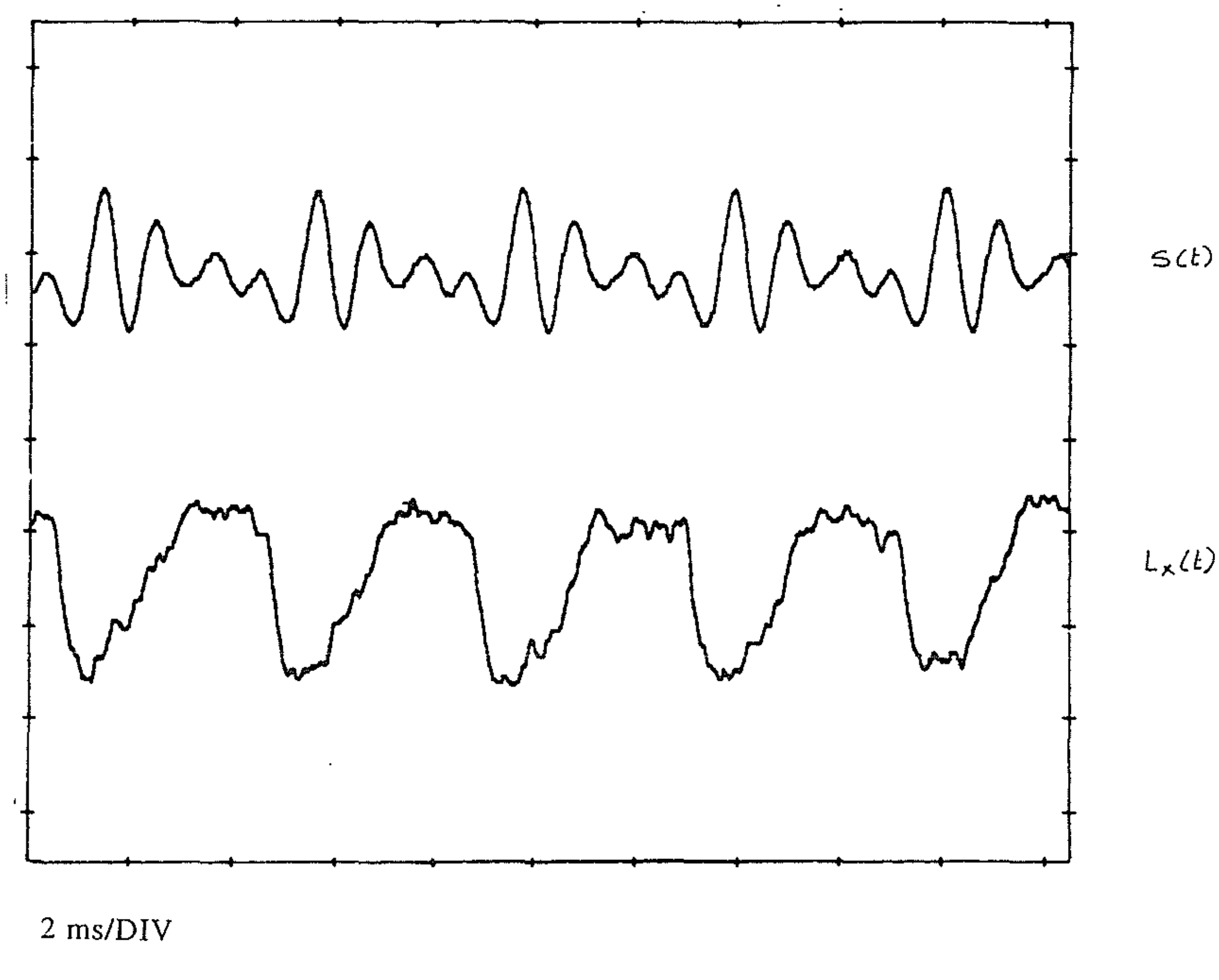
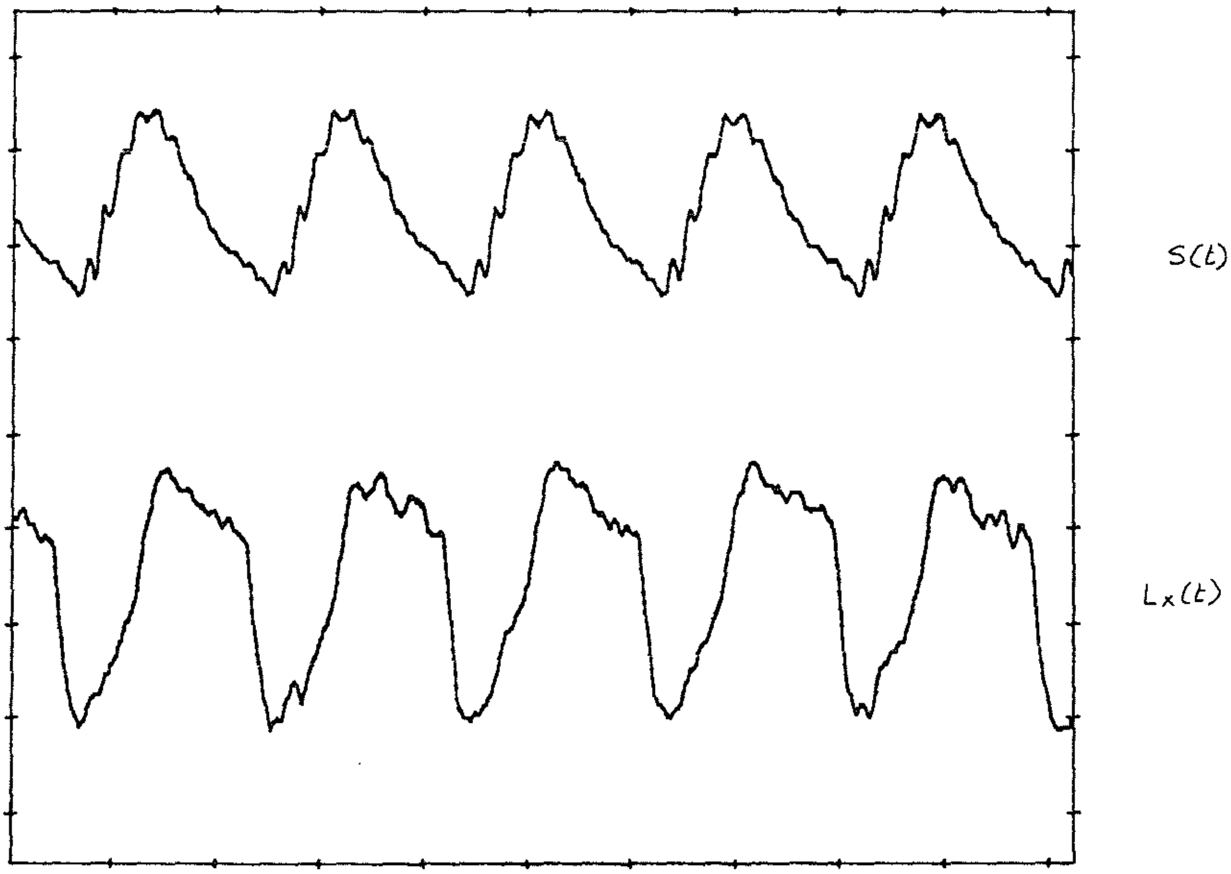


Fig. 7.4 Recordings of speech and the impedance variation waveform during the utterance of vowel /a/ by a female subject (DS) with normal vocal folds.



2 ms/DIV

Fig. 7.5 Recordings of speech and the impedance variation waveform during the utterance of vowel */i/* by a female subject (DS) with normal vocal folds.

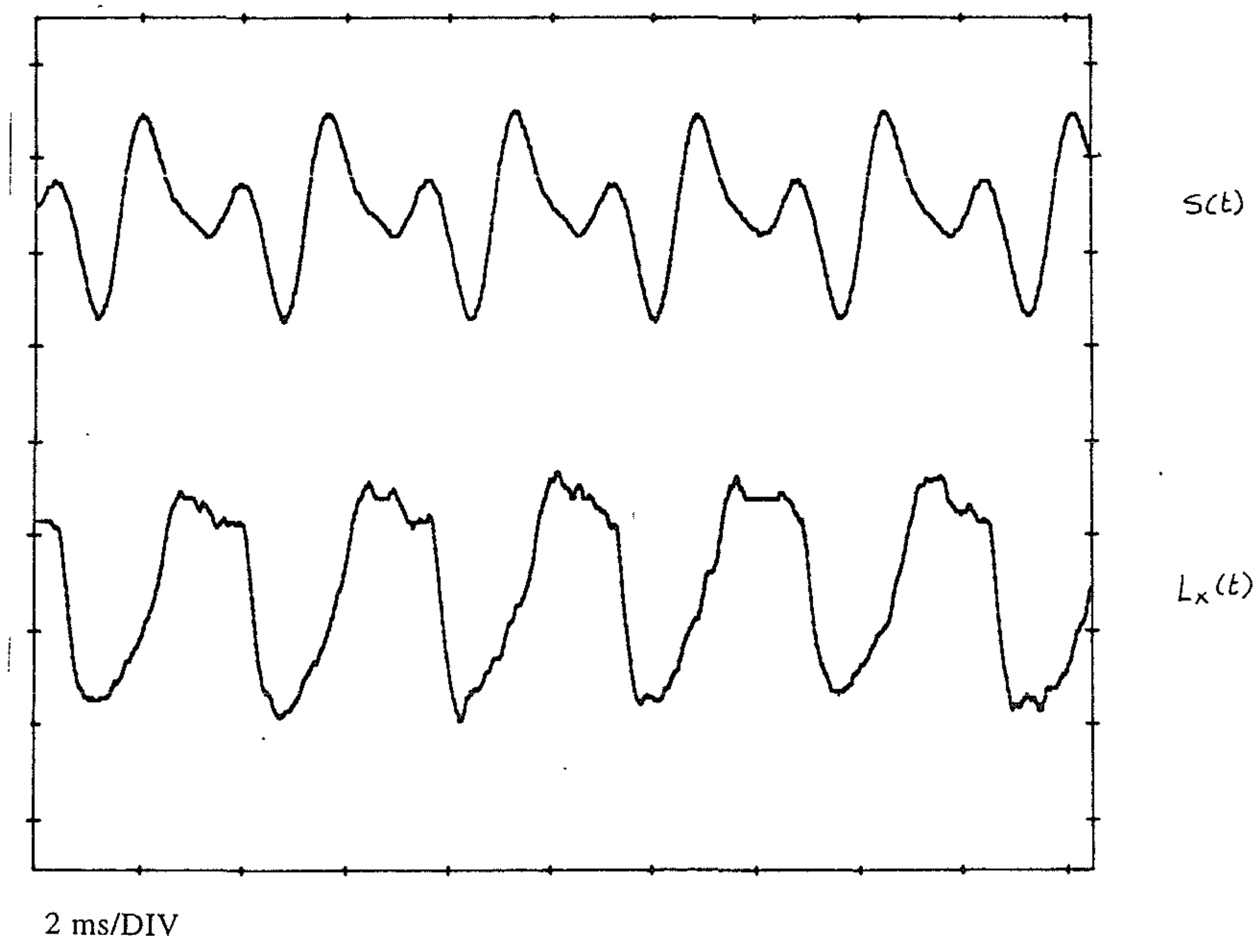


Fig. 7.6 Recordings of speech and the impedance variation waveform during the utterance of vowel /u/ by a female subject (DS) with normal vocal folds.

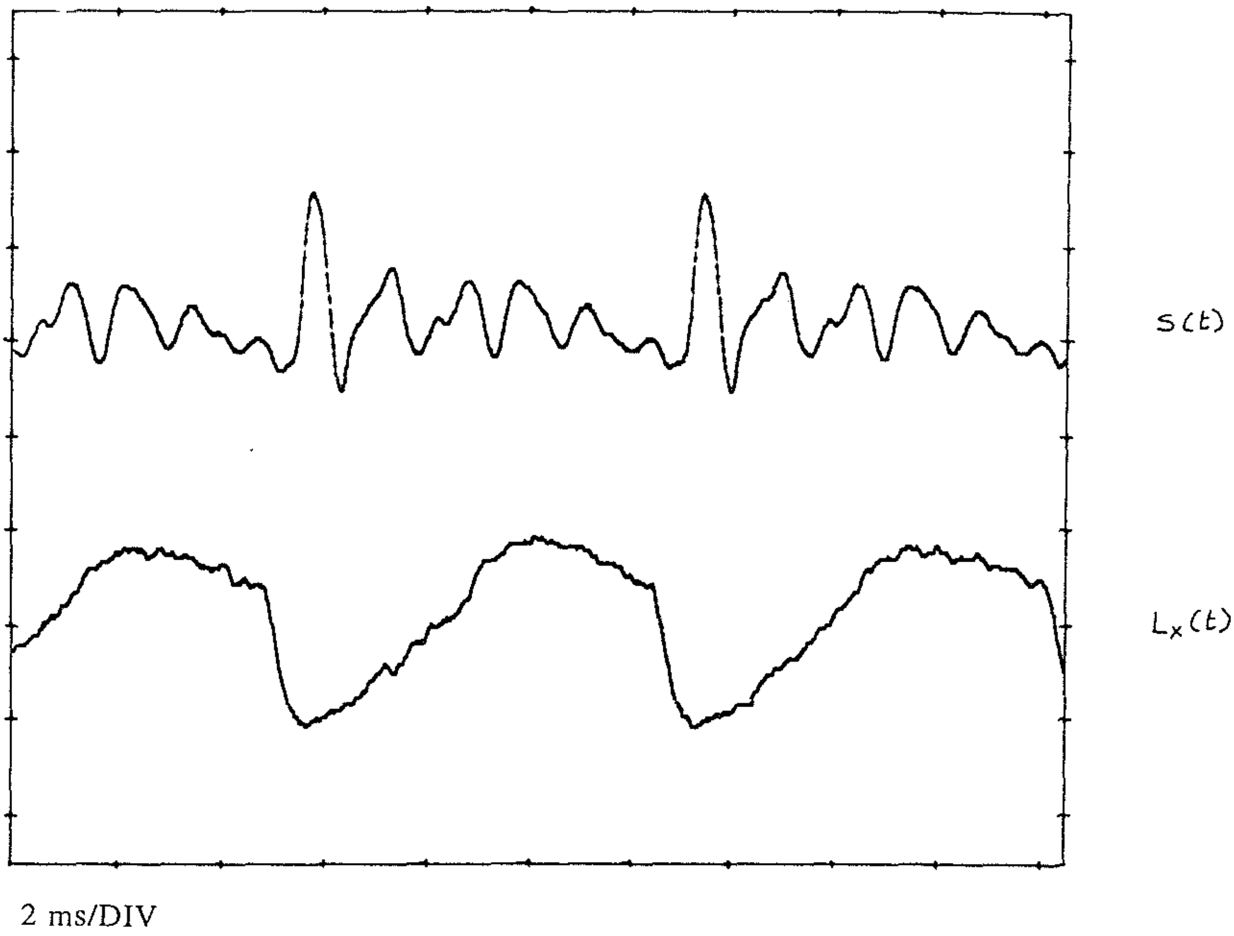


Fig. 7.7 Recordings of speech and the impedance variation waveform during the utterance of vowel /a/ by a male subject (AK) undergoing a treatment for the paralysis of the right vocal fold.

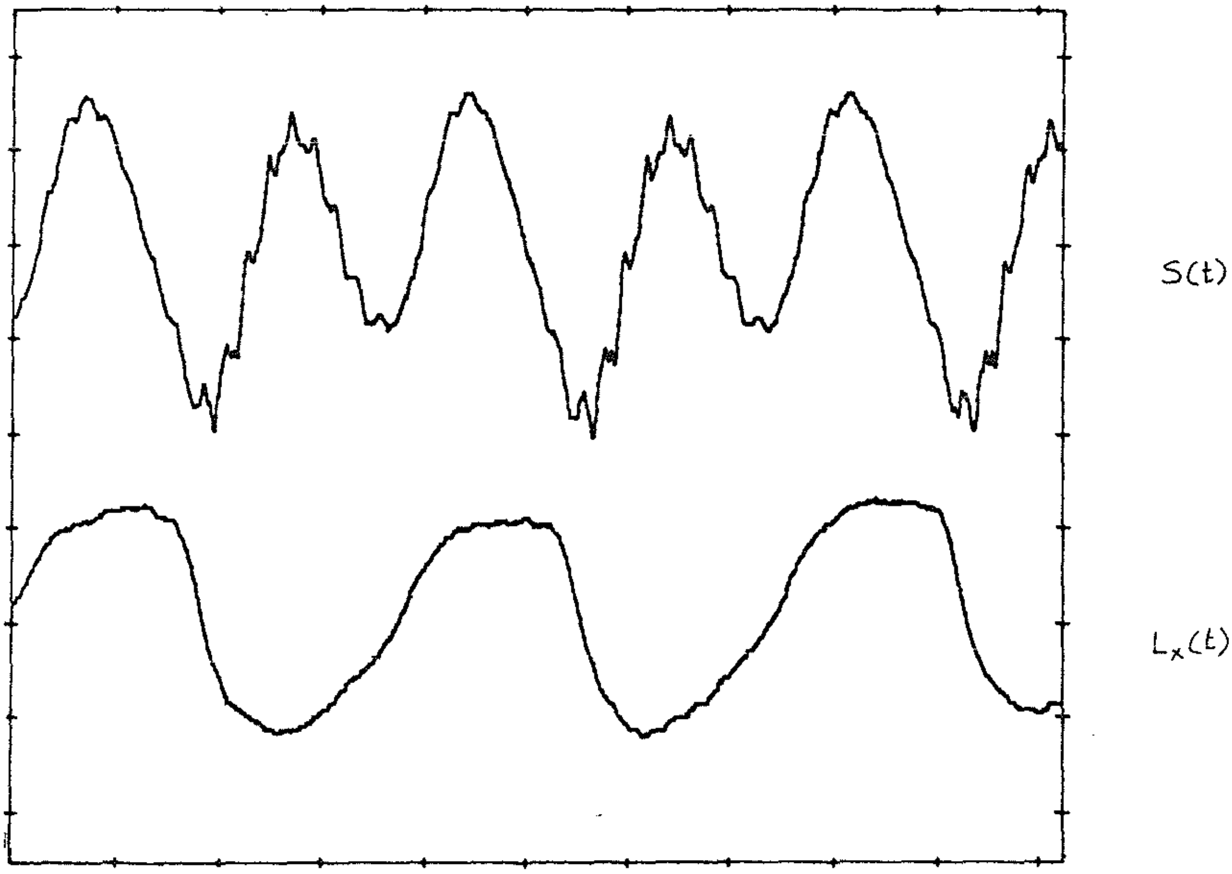


Fig. 7.8 Recordings of speech and the impedance variation waveform during the utterance of vowel /i/ by a male subject (AK) undergoing a treatment for the paralysis of the right vocal fold.

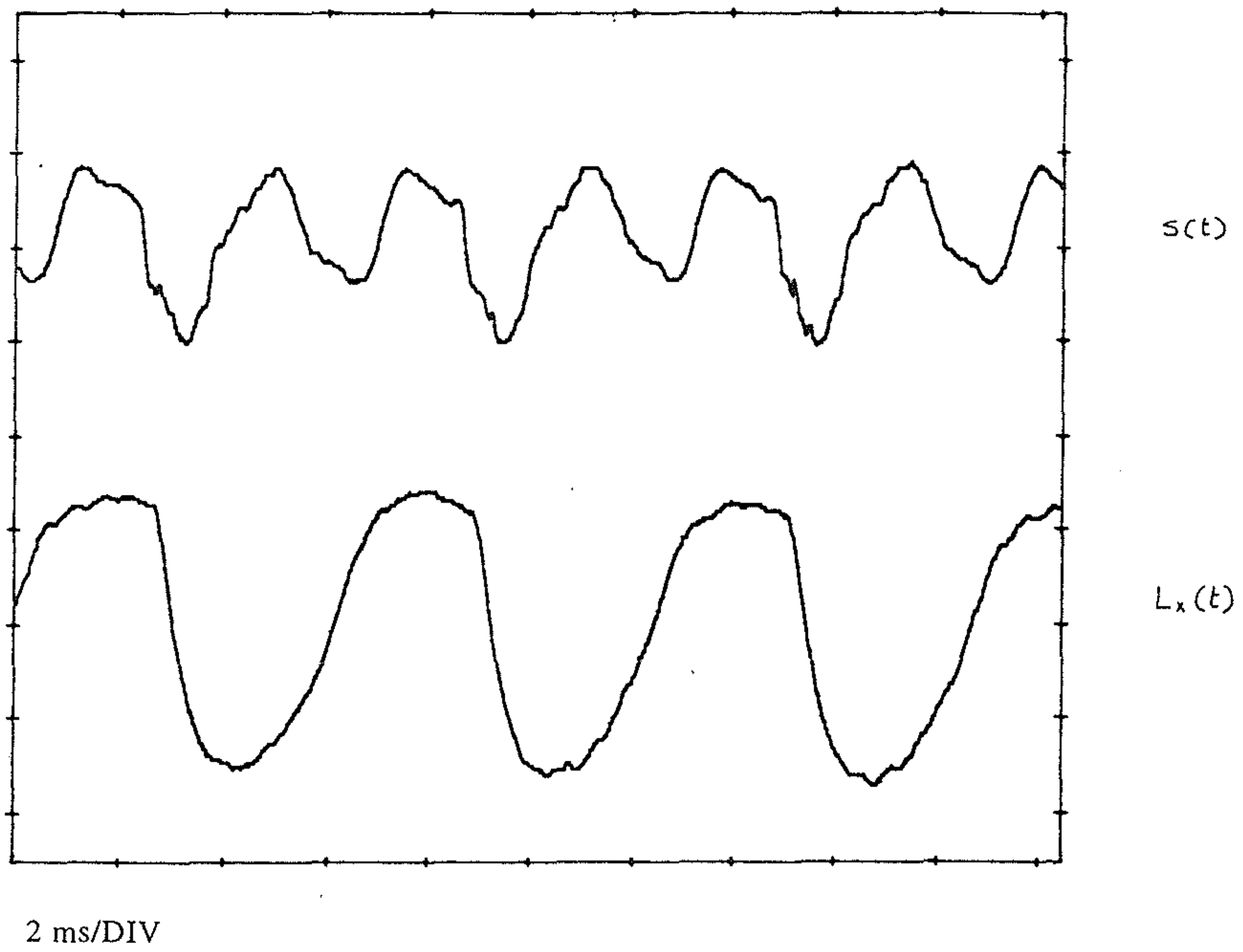


Fig. 7.9 Recordings of speech and the impedance variation waveform during the utterance of vowel /u/ by a male subject (AK) undergoing a treatment for the paralysis of the right vocal fold.

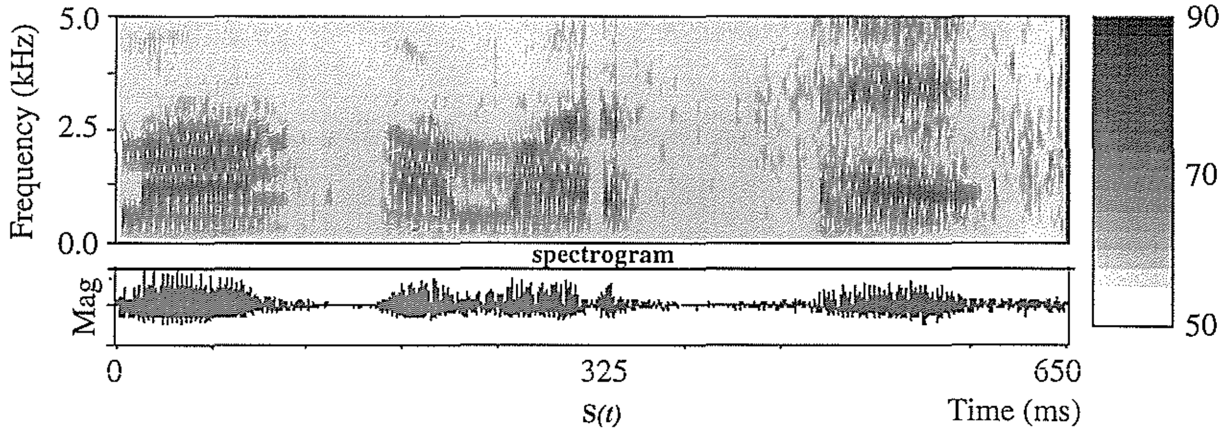


Fig. 7.10 Wideband spectrogram and the speech waveform $S(t)$ for the utterance "*...atma amar hai...*".

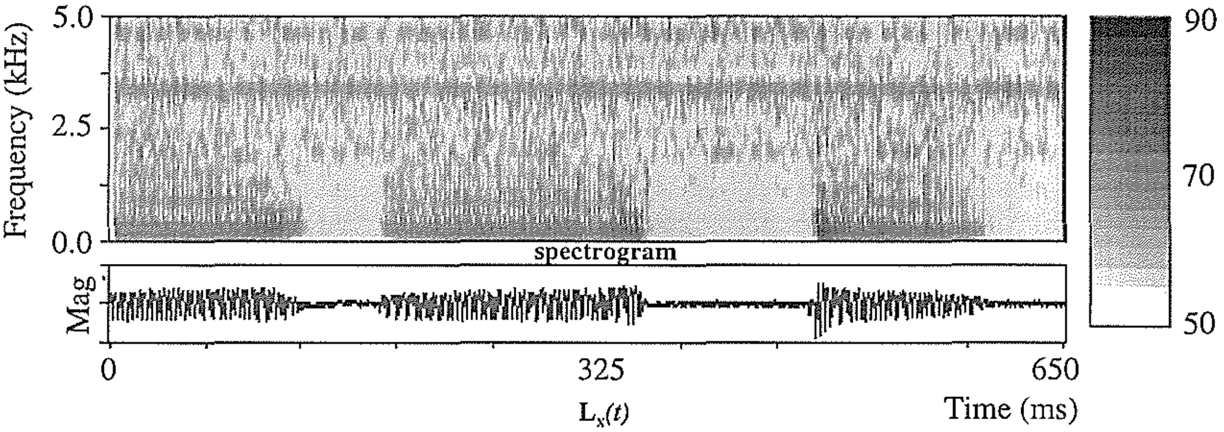


Fig. 7.11 Wideband spectrogram and the impedance variation waveform $L_x(t)$ for the utterance "*...atma amar hai...*".

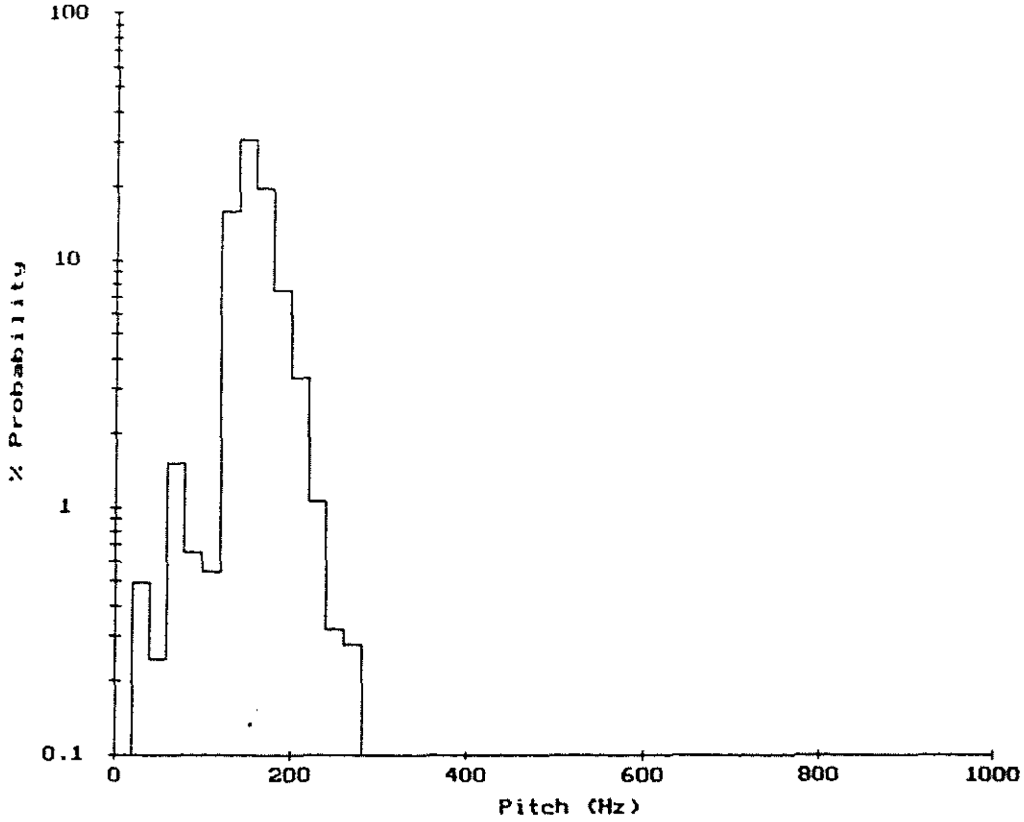


Fig. 7.12 Single period F_x histogram for a male subject (MC), reading continuous text for a time duration of 1 minute.

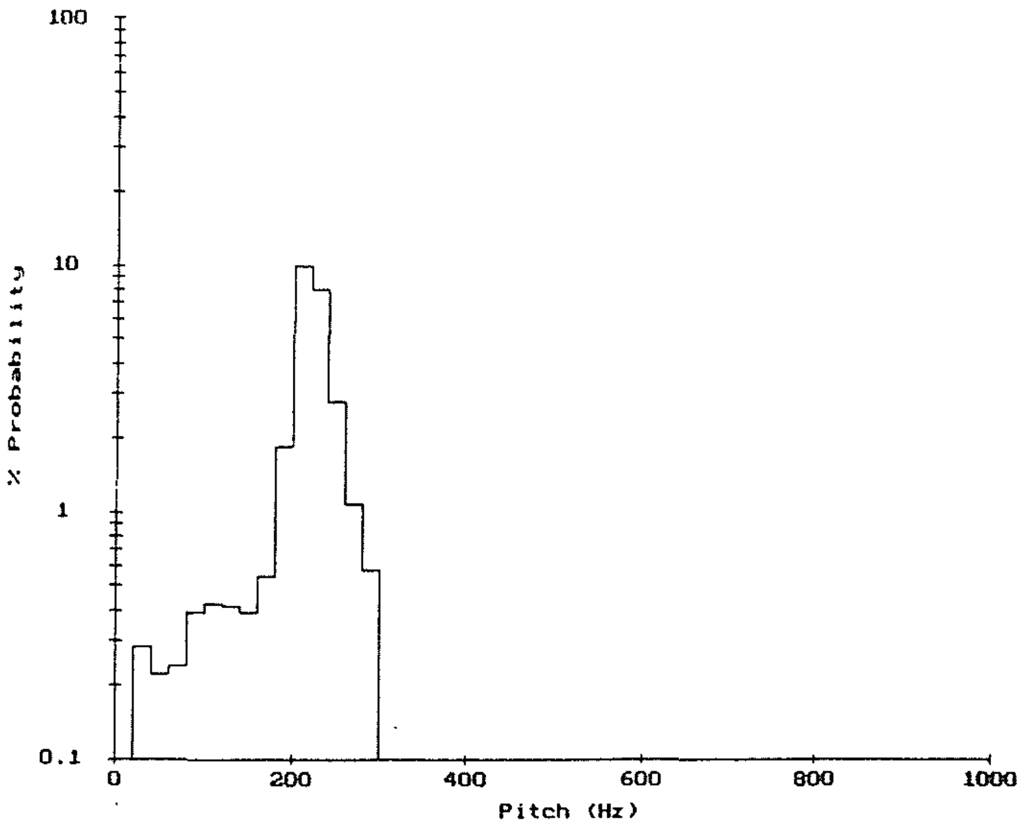


Fig. 7.13 Single period F_x histogram for a female subject (DS), reading continuous text for a time duration of 1 minute.

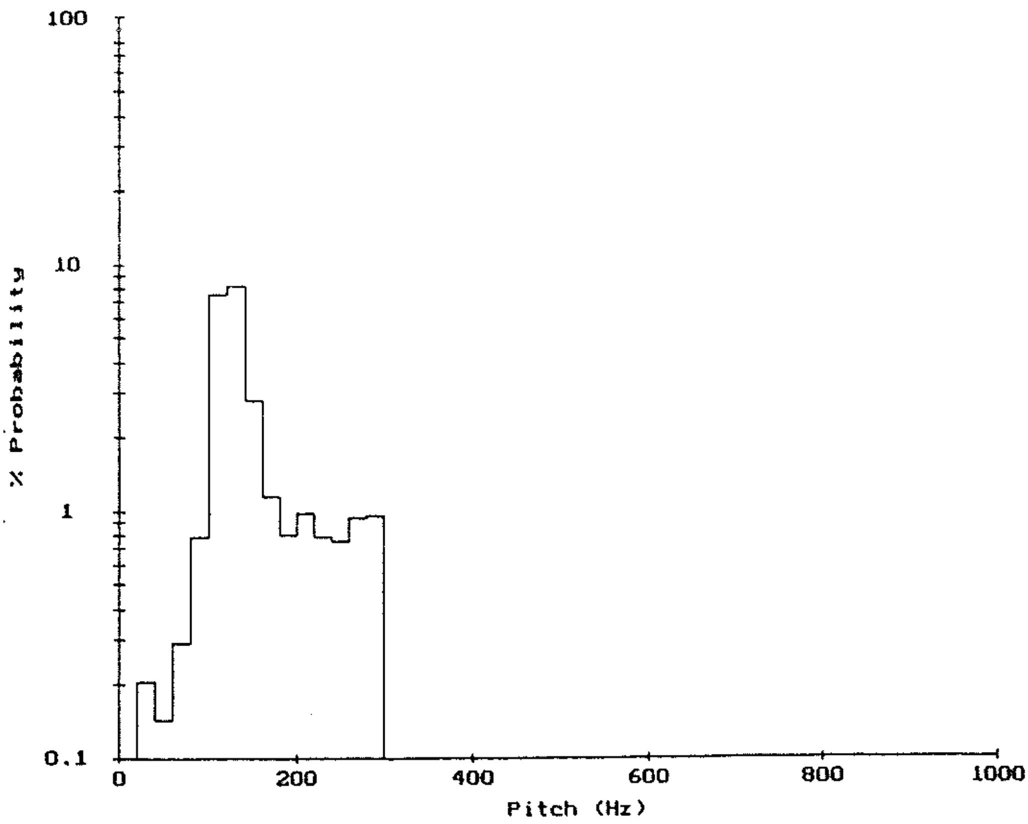


Fig. 7.14 Single period F_x histogram for a male subject (AK) undergoing a treatment for the paralysis of the right vocal fold, reading continuous text for a time duration of 1 minute.

Chapter 8
SUMMARY & CONCLUSION

8.1 WORK DONE

The aim of this project was to design and develop a low-cost, battery operated glottal pitch extractor to detect these impedance variations across the human larynx. A low cost microcontroller based signal acquisition, analysis, and LCD graphics display unit was also to be developed as a part of this instrument for displaying the impedance variation waveforms, along with a RS-232 serial interface for downloading the measurement results on to a computer and for taking a hardcopy of the stored waveforms on a plotter.

The basic circuit for extracting the impedance variation has been developed and tested with simulated impedance variation.

The electrodes were made on a glass epoxy PCB and the circuit was tested for extracting the impedance variation waveform. The circuit was tested using a two electrode method with various configurations of the connection of the individual conductors of each probe. Maximum sensitivity was obtained with driven guard rings.

The signal acquisition and display unit has been developed and tested for operation in various modes. All the eight channels are tested for various inputs and are working satisfactorily.

A signal acquisition and analysis software is developed on a notebook PC with a PCMCIA based A/D card for capturing the impedance variation waveforms and plotting the F_x histograms. All the modes of the software were tested for capturing various signals and the software is found to be working satisfactorily.

The glottal pitch extractor was used for recording the impedance variation waveforms of different subjects. Subjects with normal vocal folds and subjects with vocal fold disorders were tried out. F_x histograms were plotted for male subject with normal vocal folds for different vowels as well as for continuous text.

8.2 SUGGESTIONS FOR FURTHER DEVELOPMENT

The impedance variation detector unit requires ± 8 - 12 V supply, and at present it operates from two 9 V batteries. It will be desirable to operate it from less expensive batteries. This can be done using two 1.5 V cells, a 3 V to 5V high efficiency converter (Philips TEA1204) and 5V to ± 12 V, 100 mA DC-DC converter module (such as Newport Components NMH0512D, NMXD0512).

The signal acquisition and display unit works on 5V, 180 mA and ± 12 V, 60 mA supply. Therefore this also can be obtained using the DC-DC converters similar to the one used in the impedance variation detector.

The impedance variation detector circuit has been built using cascading of high pass and low pass filters for obtaining a bandwidth of 10 Hz to 10 kHz. It is desirable to use a single higher order filter for the same.

Impedance variation waveforms of subjects with different disorders of the vocal folds should be tried out for carrying out a detailed analysis of the vocal fold disorders and its effect on voice pitch.

APPENDIX A

A.1 SPECIFICATIONS OF THE IMPEDANCE VARIATION DETECTOR

The portable impedance variation detector unit has one pair of gold-plated electrodes, along with a suitable neckband and two 9 V rechargeable nickel-hydride batteries. Other important specifications of this unit are as follows.

Electrodes	: Neck electrodes
Electrode area	: Circular gold plated glass epoxy PCB, 30 mm in diameter
Electrode distance	: Adjustable
Neck band	: Velcro strap
Electrode voltage	: Below 1 V
Electrode current	: 3 mA maximum
Carrier frequency	: 300 kHz
Bandwidth for impedance variation detector	: 10 Hz to 10 kHz
Impedance variation sensitivity	: 2 V/Ω for $Z_b = 500\ \Omega$
Power source	: Two PP3 type rechargeable Nickel-Hydride batteries. 9 V / 120 mAh
Current drain	: 30 mA
Use with one charge	: 1 hour of continuous use
Size	: 190 (L) x 140 (W) x 45 (H) mm

A.2 FEATURES AND SPECIFICATIONS OF THE SIGNAL ACQUISITION AND DISPLAY UNIT

Features

- Single / dual channel display.
- 20 k samples-per-second for single channel operation.
- Menu-driven operator interface for easy operation.
- Digital storage for four waveforms distributed among two memory banks.
- Simultaneous display of two waveforms (any combination amongst the eight channels) allows for useful comparisons.
- Hold mode for measurement of voltage and time on the waveforms using cursors.
- Data logger mode for continuous acquisition of signals.
- RS-232 serial interface for data transfer to the computer and for plotting on HP-7475A plotter.

Specifications

<u>General</u>

No. of input channels	: Eight	
Max. sampling rate	: 20 k samples/sec	single channel
	8 k samples/sec	dual channel
Input quantization	: 8 bit	
Accuracy	: ± 1 LSB	
Record length	: 200 samples in RUN mode	

User specified in DATA LOGGER mode

Storage memory	: two memory banks of 512 bytes each
Vertical resolution	: 10 levels per division
Horizontal resolution	: 1 ms per division
Move control	: steps of 1 division for channel A
Input voltage	: 0-5 V_{pp} / ± 2.5 V_{pp}
Voltage scaling	: 0.2, 0.5, 1, and 2 V/div.
Time scaling	: 1, 2, 5, 10, 20, 50, 100, 200, 500, and 1000 ms/div.
Triggering	: -2.5 - 0 - +2.5 V in steps of 0.5 V
Pre-trigger	: 0, 25, 50, 75, and 100 %
Data logger mode	: continuous data capture under user control and transfer at 9600 baud on RS-232 serial interface
Plotting	: interface to HP-7475A plotter
Power supply and current drain	: + 5 V - 180 mA, ± 12 V - 60 mA
Ground reference	: true ground for circuit operation, generated ground of 2.5 V for capturing bi-polar signals

Connectors	: BNC female with generated ground for two channels 25 D for other six channels with generated as well as true ground stereo socket (female) for serial interface
Size	: 290 (L) x 240 (W) x 45 (H) mm

Display

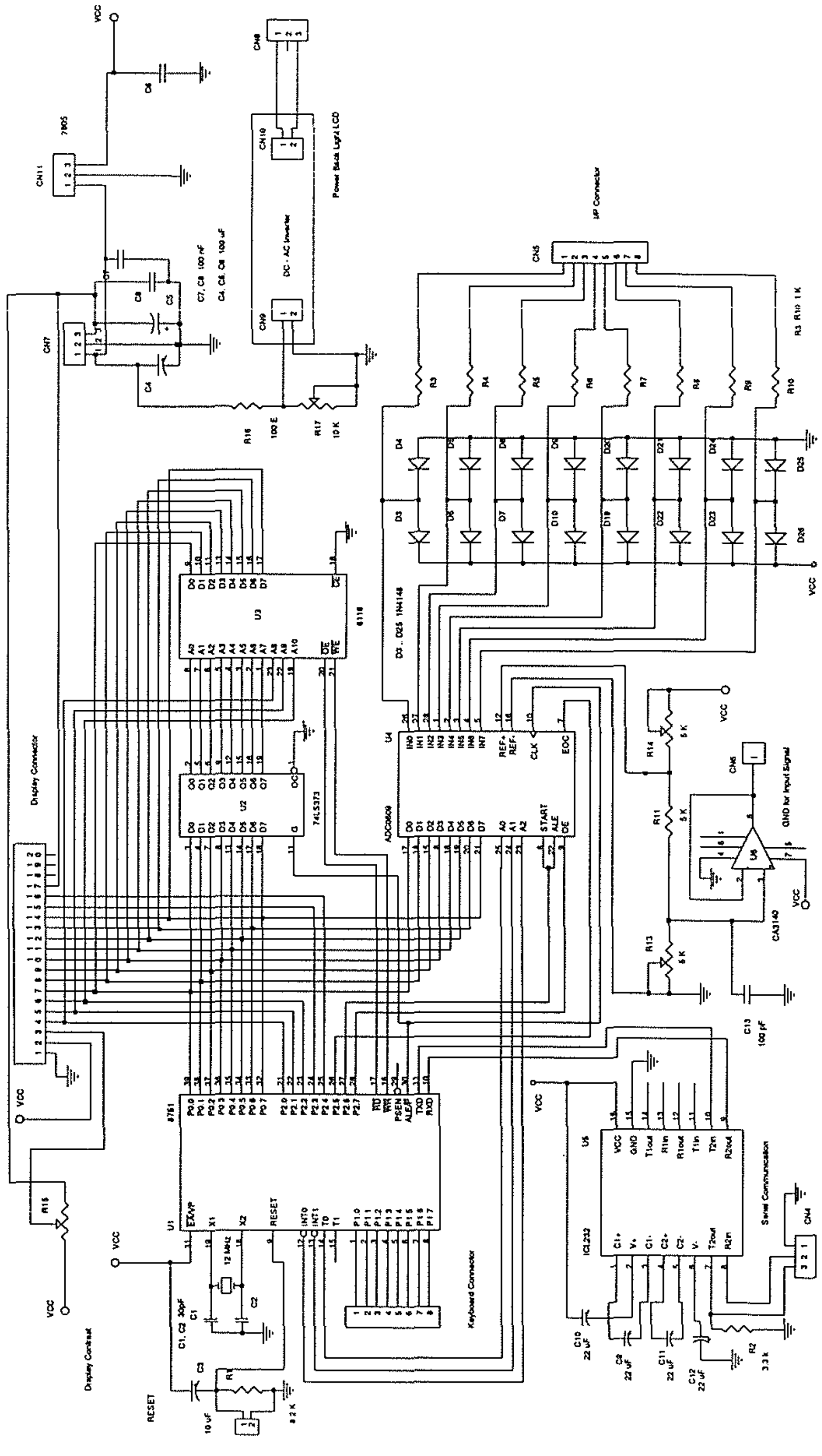
Type	: Dot matrix liquid crystal display (OGM-24011) with backlight
Active area	: 240 x 128 pixels
Trace area	: 200 x 100 pixels
Power supply	: + 5 / 30 mA, ± 12 V / 10 mA
Dimensions	: 144 x 104 mm

Keypad

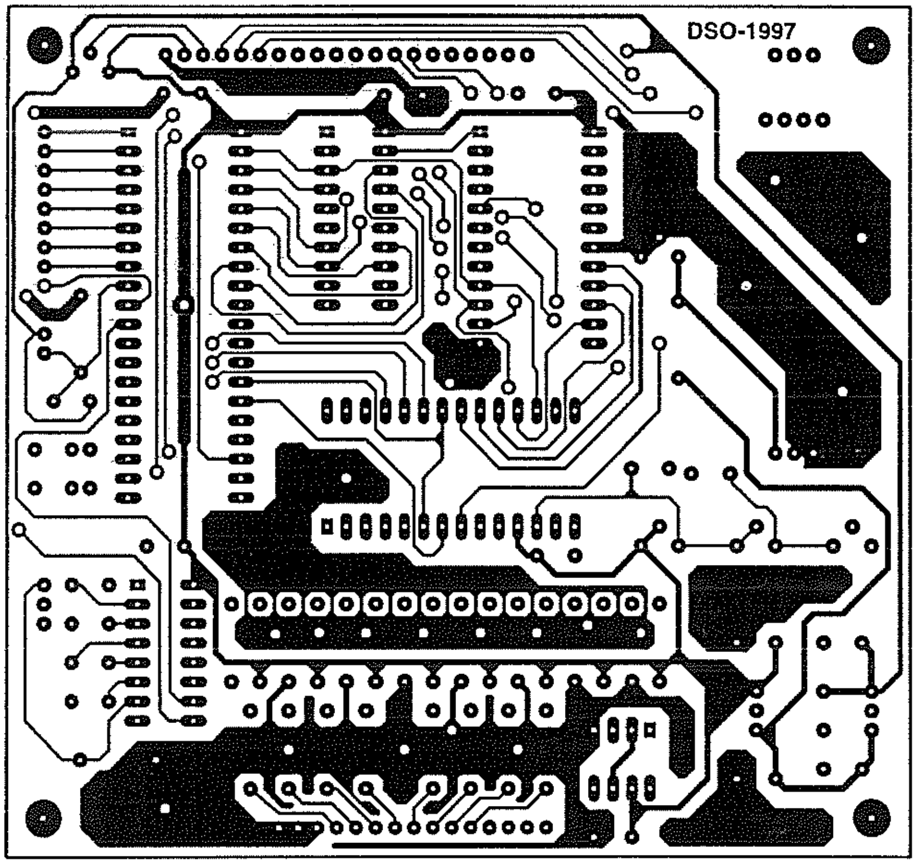
4 x 4 Contact type matrix keypad (90 x 90 mm)

APPENDIX B

B.1 CIRCUIT DIAGRAM OF THE SAD UNIT

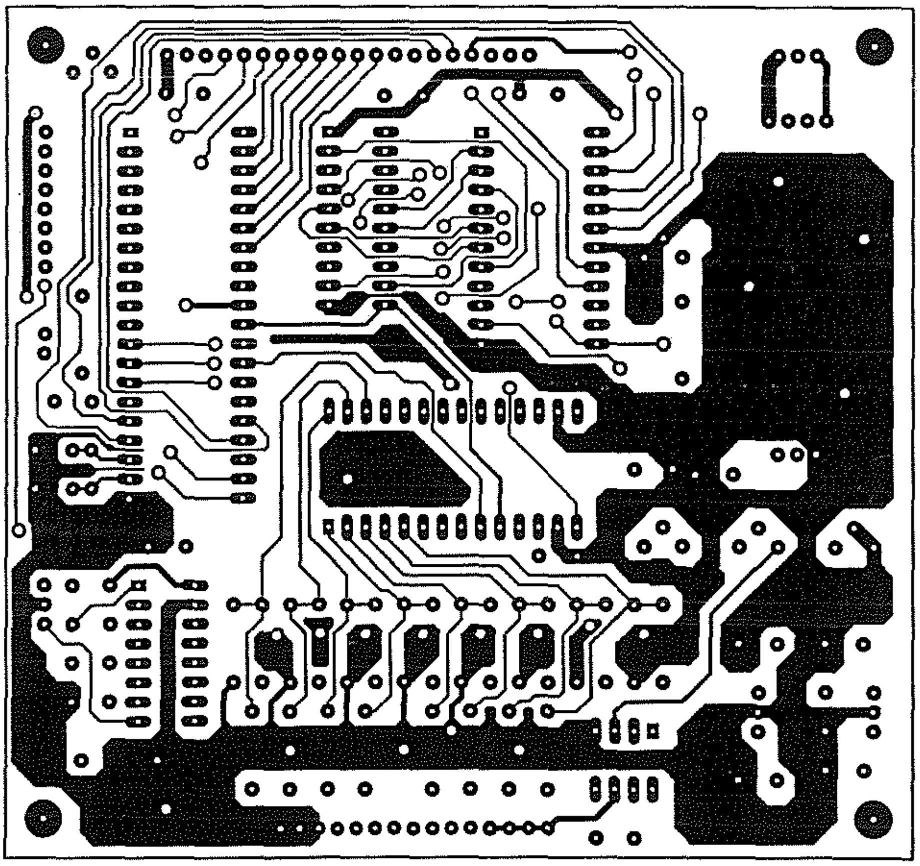


B.2 COMPONENT SIDE PCB LAYOUT OF THE SAD UNIT



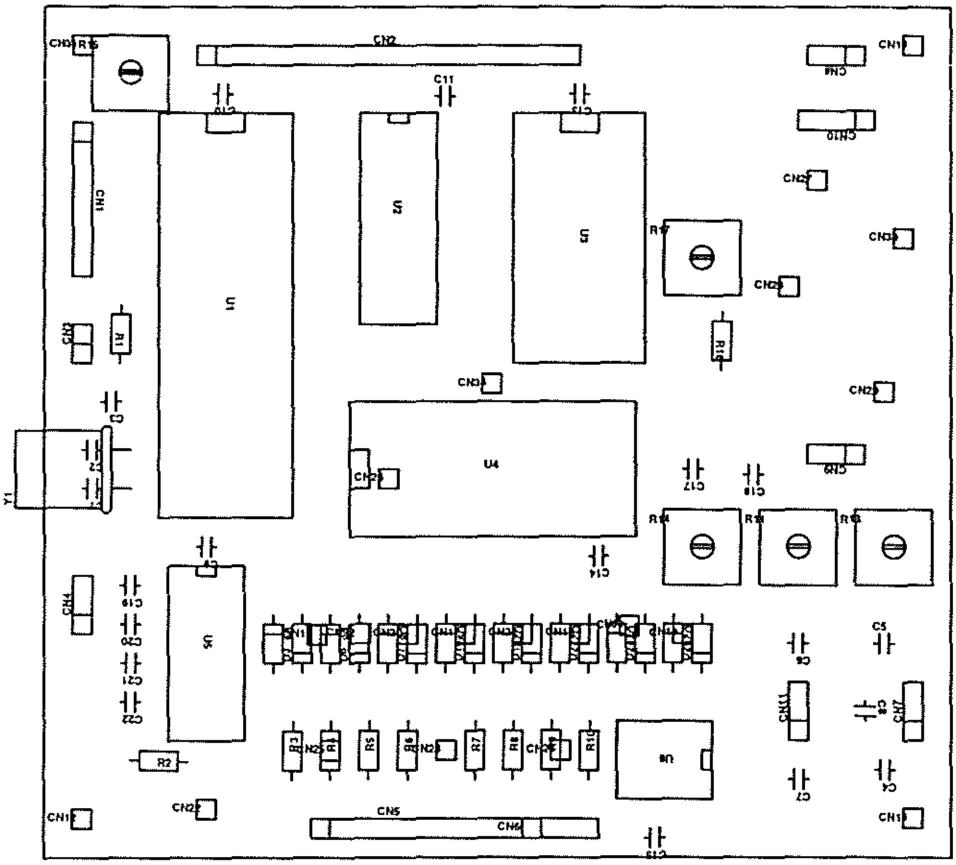
Component Side

B.3 SOLDER SIDE PCB LAYOUT OF THE SAD UNIT

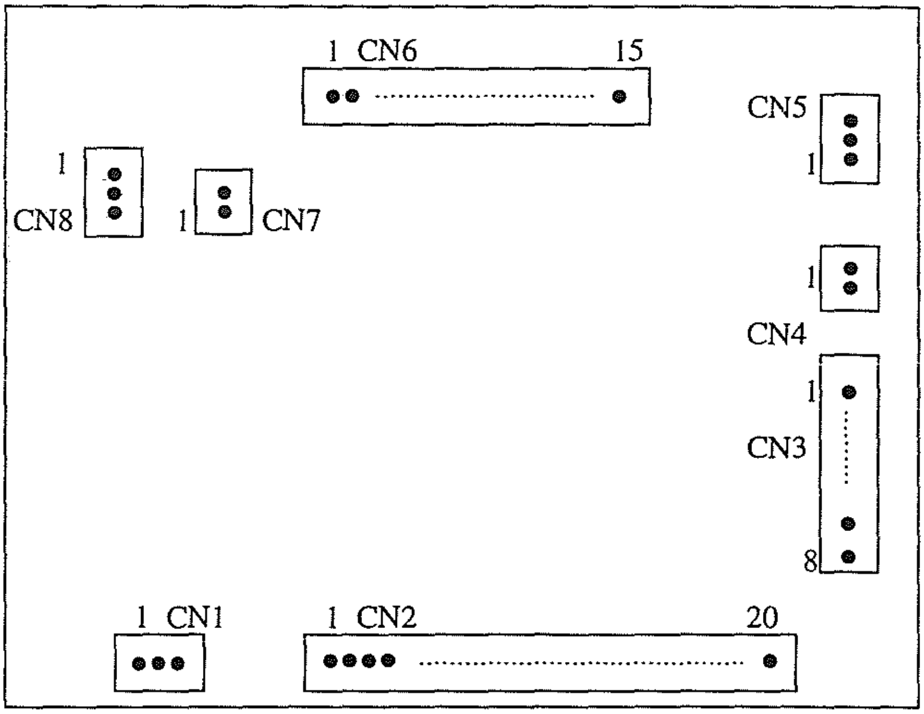


Solder Side

B.4 COMPONENT PLACEMENT PCB LAYOUT OF THE SAD UNIT

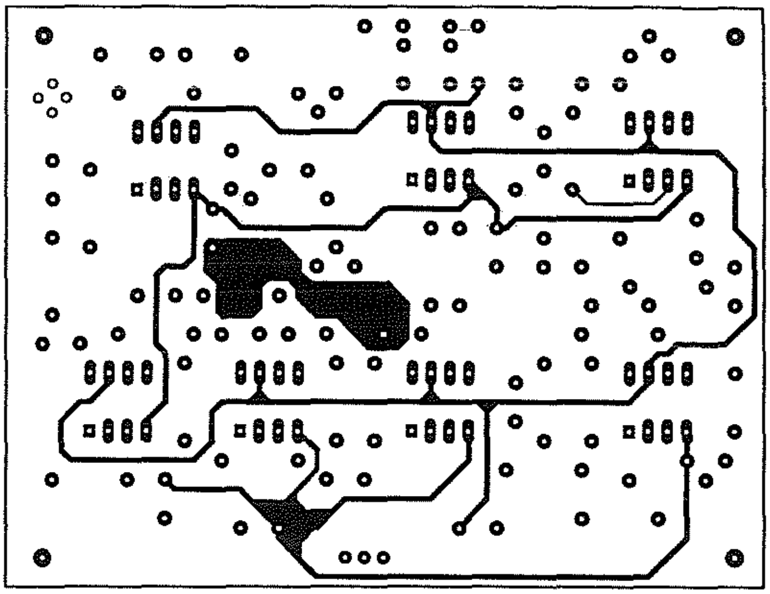


B.5 CONNECTION DETAILS OF THE SAD UNIT



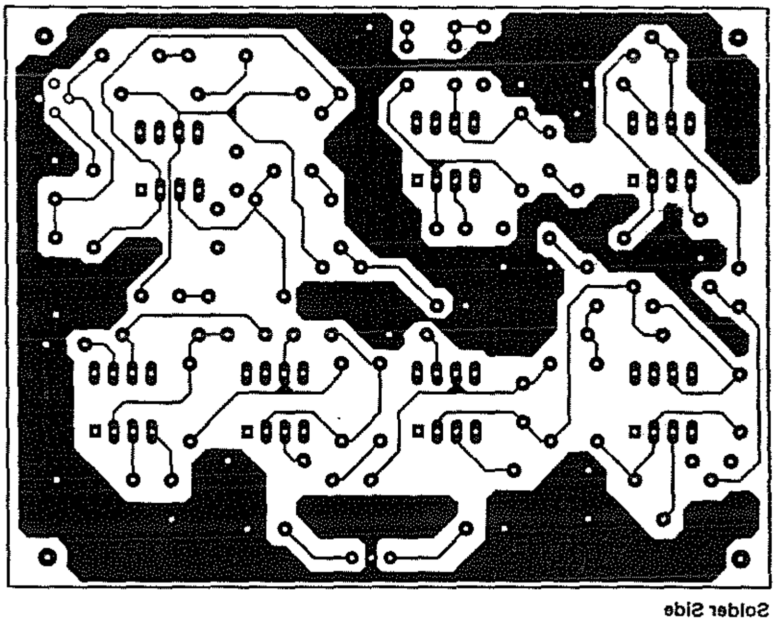
CN1 : Display backlight connector	1 - White	2 - NC	3 - White
CN2 : Display connector	1 - Vss	2 - Vdd	3 - Vo
	4 - RS	5 - R/W	6 - E
	7 - 14 DB0 - DB7		15 - $\overline{CS}$
	16 - $\overline{RST}$	17 - Vee	18 - 20 NC
CN3 : Keyboard connector	1 - Black	2 - Brown	3 - Red
	4 - Orange	5 - Yellow	6 - Green
	7 - Blue	8 - Violet	
CN4 : Reset switch connector	1 - Black	2 - Brown	
CN5 : Serial port connector	1 - Transmit	2 - Receive	3 - ground
CN6 : Analog input connector	1 - 4 Reference ground		
	5 - Ch 8	6 - Ch 7	7 - Ch 6
	8 - Ch 5	9 - Ch 4	10 - Ch 3
	11 - Ch 2	12 - Ch 1	
	13 - 15 Circuit ground		
CN7 : Power Supply connector	1 - +5 V	2 - Ground	
CN8 : Power supply connector	1 - +12 V	2 - Common	3 - -12 V

B.6 CONPONENT SIDE PCB LAYOUT OF THE IMPEDANCE VARIATION DEETCTOR

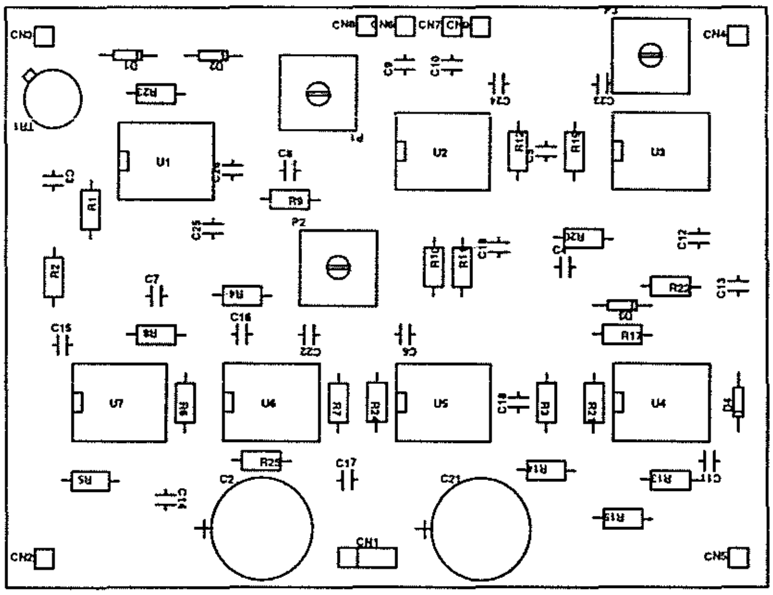


Component Side

B.7 SOLDER SIDE PCB LAYOUT OF THE IMPEDANCE VARIATION DETECTOR



B.8 COMPONENT PLACEMENT PCB LAYOUT OF THE IMPEDANCE VARIATION DETECTOR



APPENDIX C

C.1 COMPONENT LIST FOR IMPEDANCE VARIATION DETECTOR

Reference Designator	Part Number / Value	Part Description	Approximate Price per Part (Rs.)
C12, C13	680 Pf	Capacitor (ceramic)	1.00
C1, C2, C4, C7	100 pF	Capacitor (ceramic)	1.00
C5, C6, C8	100 nF	Capacitor (ceramic)	1.00
C3, C9	1 nF	Capacitor (ceramic)	1.00
C10	10 nF	Capacitor (ceramic)	1.00
C11	10 uF	Capacitor (electrolytic)	3.00
R1, R2	5.1 k	Resistor (MFR)	0.75
R3, R4, R6, R9, R11, R22, R24, R25, R26	2.2 k	Resistor (MFR)	0.75
R10, R13	1 k	Resistor (MFR)	0.75
R7, R8, R14	10 k	Trim pot	15.00
R19	10 k	Resistor (MFR)	0.75
R16, R20	15 k	Resistor (MFR)	0.75
R15, R17, R18, R21	18 k	Resistor (MFR)	0.75
R12	100 k	Resistor (MFR)	0.75
R23, R27	270 k	Resistor (MFR)	0.75
R5	470 k	Resistor (MFR)	0.75
D1	3.9 V	Zener Diode	3.00
D2, D3, D4	1N4148	Diode	2.00
U1, U2, U3, U4	CA3140	Op-Amp	20.00
U5, U6	LF356	Op-Amp	15.00
Q1	BFW11	FET	10.00
		Electrodes (one pair, gold plated)	15.00
		Neckband	10.00
		PCB (PTH, double sided)	200.00
B1, B2	9 V / 120 mAh	Nickel-Hydride battery	300.00
		Cabinet, connectors	100.00
		TOTAL ≈	1,150.00

C.2 COMPONENT LIST FOR THE SAD UNIT

Reference Designator	Part Number / Value	Part Description	Approximate Price per Part (Rs.)
C1, C2 C3 C4, C5, C6 C7, C8 C9, C10, C11, C12	30 pF 10 uF 100 uF 100 nF 22 uF	Capacitor (ceramic) Capacitor (ceramic) Capacitor (electrolytic) Capacitor (ceramic) Capacitor (electrolytic)	1.00 3.00 4.00 1.00 4.00
R1 R3…R10 R2 R11 R12, R13, R14 R15 R16 R17	8.2 k 1 k 3.3 k 5 k 5 k 50 k 100 E 10 k	Resistor Resistor (MFR) Resistor Resistor Trim pot Trim pot Resistor Trim pot	0.50 0.75 0.50 0.50 15.00 15.00 0.50 15.00
D1…D26	1N4148	Diode	2.00
U1 U2 U3 U4 U5 U6	89C55-24PC 74LS373 HM6116M-3 ADC0809 ICL232 CA3140	Microcontroller (Atmel) Latch 2 kB Memory ADC RS-232 converter Op-Amp	800.00 40.00 40.00 100.00 40.00 20.00
	OGM24011	LCD (Oriole) 4 x 4 Keypad PCB (PTH, double sided) Cabinet, connectors	8000.00 400.00 250.00 200.00
		TOTAL ≈	10,060.00

REFERENCES

[1] K. J. Ayala, *The 8051 Microcontroller, Architecture, Programming, and Applications*, West Publishing, St. Paul, MN 55164-0526, 1991.

[2] T. Baer, A. Löfqvist, and N. S. McGarr, "Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques," *J. Acoust. Soc. Am.*, vol. 73, pp. 1304-1308, Apr. 1983.

[3] R. Bhagwat , "Glottal pitch extractor," *B.Tech. Project Report*, IIT Bombay, 1990, Guide : Dr. P.C. Pandey.

[4] K. R. Botkar , *Integrated Circuits*, Khanna Publishers, Delhi, 1992

[5] D. G. Childers and J. N. Larar, "Electroglottography for laryngeal function assessment and speech analysis," *IEEE Trans. Biomed. Engg.*, vol. BME-31, pp. 807-817, Dec. 1984.

[6] D. G. Childers, D. M. Hicks, D. M. Moore, and Y. A. Alaska, "A model for vocal fold vibratory motion, contact area, and the electroglottogram," *J. Acoust. Soc. Am.*, vol. 80, No.5, pp. 1309-1320, Nov. 1986.

[7] J. J. Dubnowski, R. K. Schafer, and L. R. Rabiner, "Real time digital hardware pitch detector," *IEEE Trans. Acoust. Speech & Signal Process.*, vol. ASSP-24, pp. 2-8, Feb.1976.

[8] A. J. Fourcin, " Laryngographic assessment of phonatory function," *ASHA Reports*, vol. 11, 1981.

[9] F-J Electronics A/S, Electro-Glottograph (type EG830), Gentofte, Denmark.

[10] F-J Electronics A/S, Portable Electroglottograph (type EG80), Gentofte, Denmark.

[11] R. A. Gayakwad, *Op-amps & Linear Integrated Circuits*, 2nd ed., Prentice Hall India, New Delhi, April 1991.

[12] T. Haji, S. Horiguchi, T. Baer, and W. J. Gould, "Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation," *J. Acoust. Soc. Am.*, vol. 80, pp. 58-62, Jul. 1986.

[13] D. M. Howard, A. J. Fourcin, "Instantaneous voice period measurement for cochlear stimulation," *Electronic Letters*, vol. 19, 15th July 1983.

[14] M. Hirano, *Clinical Examination of Voice*, Disorders of Human Communications, vol. 5, Springer-Verlag Wien, New York, 1981.

[15] Intersil Inc., *Component Data Catalog*, Cupertino, CA 95014, 1987.

[16] E. J. Kreul and M. H. L. Hecker, "Descriptions of the speech of patients with cancer of the vocal folds. Part II: judgments of age and voice quality," *J. Acoust. Soc. Am.*, vol. 49, No. 4, pp. 1283-1287, 1971.

[17] Laryngograph Ltd., *Portable Electro-laryngograph*, London, England.

[18] K. K. Mahajan, "Glottal pitch extractor," *M.Tech. Mini Project Report*, IIT Bombay, 1995, Guide : Dr. P.C. Pandey.

[19] D. Moore and G. S. Berke, "The effect of laryngeal nerve stimulation on phonation: a glottographic study using an *in vivo* canine model," *J. Acoust. Soc. Am.*, vol. 83, pp. 705-715, Feb. 1988.

[20] T. Murry, "Some acoustic features of hoarseness," *J. Acoust. Soc. Am.*, vol. 58, Suppl. No. 1, pp. S111, Fall 1971.

[21] National Semiconductor Corp., *Linear Data Book*, vol. 2, Santa Clara, California, 1988.

[22] M. H. L. Hecker and E. J. Kreul, "Descriptions of the speech of patients with cancer of the vocal folds. Part I: measures of fundamental frequency," *J. Acoust. Soc. Am.*, vol. 49, No. 4, pp. 1275-1282, 1971.

[23] *Microcontroller Data Book*, Atmel Corp., 2125 O'Nel Drive, San Jose, CA 95131.

[24] Oriole Electronics Pvt. Ltd., *LCD Display Data Book*, Mumbai-86, 1997.

[25] D. O'Shaughnessy, *Speech communication, Human and Machine*, Addison-wesley publishing company, New York.

[26] P. C. Pandey, "Speech processing for cochlear prosthesis," Ph.D. Thesis, Department of Electrical Engineering, University of Toronto, 1987.

[27] S. M. Rosen, A. J. Fourcin, and B. C. J. Moore, "Voice pitch as an aid to lipreading," *Nature*, vol. 291, pp. 150-152, 14th May 1981.

[28] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, 1978.

[29] RCA Corp., *Guide to linear integrated circuits*, Santa Clara, CA.

[30] A. M. Smith and D. G. Childers, "Laryngeal evaluation using features from speech and the electroglottograph," *IEEE Trans. Biomed. Engg.*, vol. BME-30, pp. 755-759, Nov. 1983.

[31] A. Sriram, "Glottal pitch extractor," *B.Tech. Project Report*, IIT Bombay, 1991, Guide : Dr. P.C. Pandey.

[32] *TTL Logic Data Book*, vol. 2, Texas Instruments, Dallas, Texas, March 1988.

[33] A. V. I. Thajudin, "Glottal pitch extractor," *M.Tech. Mini Project Report*, IIT Bombay, 1994, Guide : Dr. P.C. Pandey.

[34] I. R. Titze, "Physiological and acoustic differences between male and female voices," *J. Acoust. Soc. Am.*, vol. 85, pp. 1699-1707, Apr. 1989.

[35] K. J. L. Wilson, *Anatomy and Physiology in Health and Illness*, ELBS with Churchill Livingstone, Hong Kong, 1994.

[36] M. Qu, Y. Zhang, J. G. Webster, "Motion artifact from spot and band electrodes during impedance cardiography," *IEEE Trans. Biomed. Engg.*, vol. BME-33, pp. 1029-1035, Nov. 1986.