Proc. Bio Vision 2001 Int. Conf. Biomed. Eng., Bangalore, India, Dec. 21-24, 2001, pp. 1B.33-37

A Low Cost Impedance Glottograph and Glottal Pitch Analyzer

Parveen K. Lehana and Prem C. Pandey

***Abstract*–Measurement of the electrical impedance variation across the thyroid cartilage during phonation provides information about the dynamics of the closure of vocal chords and glottal pitch. A low cost impedance glottograph and glottal pitch analyzer is developed. It consists of glottal impedance detector module and a PC with sound card. In the hardware module, a 400 kHz low level current (< 1 mA) is passed through two circular disc electrodes (30 mm dia) placed across the thyroid cartilage of the larynx. The amplitude modulated voltage signal across the electrodes is demodulated and output as electro-glottogram. This is input to the line-in of the sound card on a PC for recording. The recorded signal is used by the glottal analyzer program for determining the instantaneous value of glottal pitch as a function of time and its histographic analysis for the study of voice disorders. The system permits simultaneous recording and analysis of speech through the second channel of the sound card.**

***Index Terms*–glottal pitch, electroglottography, laryngo-gram.**

I. Introduction

The vocal chords are set into vibrations during the production of voiced segments. The fundamental frequency of the vocal chords is called pitch. The vibration cycle of the vocal chords consists of three main phases: contact phase, the separation phase, and the open phase. The onset of opening and closing and the closure duration are significantly different across speech styles and physical condition of the vocal chords [1].

When vocal chords vibrate, the contact area between them changes, and it results in variation of electrical impedance across the thyroid cartilage of the larynx. An impedance glottograph provides an electrical signal proportional to the glottal impedance variation, termed as laryngogram (Lx waveform) or electroglottogram (EGG waveform) [2], [3]. Its analysis has applications in speech analysis, monitoring the status of vocal chords, diagnosis of voice disorders, training aids for the hearing impaired, and sensory aids as supplement to lip-reading [4], [5]. For generating natural sounding speech pitch contours are necessary. Glottograph can be used for extracting these contours.

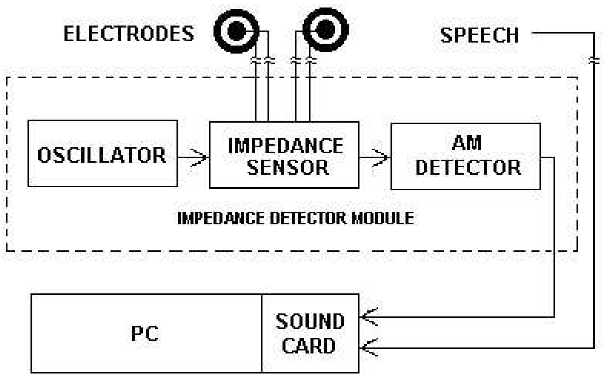
The glottal impedance variation can be measured by passing a high frequency (100 kHz to 5 MHz) low level current through the two electrodes placed across the thyroid cartilage of the larynx, and by detecting the AM voltage across the electrodes. The base impedance across the thyroid cartilage is approximately 500 Ω and the change due to vibration is less than 1 Ω [3]. The impedance sensing hardware should inject the current at a low level and provide adequate sensitivity and noise rejection. The glottal impedance detector of an earlier portable instrument [6] developed at IIT Bombay used 300 kHz sinusoidal current of 3 mA and disc electrodes with guard rings. The instrument had a microcontroller based signal acquisition and LCD unit for recording and displaying the waveform, and a serial port for transferring the digitized waveforms to a PC. In a modified version [7], the same sensitivity could be obtained with 1 mA current. Based on this glottal impedance detecting circuit, a low cost PC based impedance glottograph and glottal pitch analyzer has been developed.

II. HARDWARE

The hardware consists of the module for detecting the glottal impedance and a PC with sound card for signal acquisition, processing, and display, as shown in Fig. 1. The impedance detector consists of an oscillator with stabilized output; impedance sensor connected to two electrodes, and AM detector. The pair of circular electrodes (30-mm dia) is held in contact with the skin on both sides of the thyroid cartilage. Oscillator output voltage (400 kHz) is converted in current (adjustable < 1ma) and injected into the electrodes through ac coupling. The impedance across the thyroid cartilage varies in accordance with the contact area between the vocal chords, and consequently the voltage across the electrodes gets amplitude modulated. The AM demodulator circuit consists of band pass filter-amplifier,precision full-wave rectifier-averager, and band pass filter-amplifier. It provides the EGG or Lx

Department of Electrical Engineering, IIT Bombay, Powai, Mumbai 400076, India. Email: {lehana, pcpandey}@ee.iitb.ac.in

waveform. One line input of the sound card is used for acquisition of the Lx waveform. The other input may be used for simultaneous acquisition of speech signal.



III. SIGNAL PROCESSING

The glottal pitch analysis program is developed for acquisition and display of the speech and Lx waveforms, extraction of instantaneous glottal pitch "Fx" waveform by processing Lx, and histographic analysis of Fx. Fig. 2 shows the block diagram for signal processing.

The selected Lx waveform segment is filtered by 257-coefficient FIR band pass filter (100 Hz–900 Hz). The filtered waveform $x(n)$ is converted to a rectangular waveform by the hysteresis comparator and this waveform is used for pitch period estimation. The thresholds θ_L and θ_U for the hysteresis comparator are dynamically set on the basis of average $a(n)$, peak $p(n)$, and valley $v(n)$.

$$\theta_L(n) = a(n) - 0.1(p(n) - v(n)) \quad (1)$$

$$\theta_U(n) = a(n) + 0.1(p(n) - v(n)) \quad (2)$$

The average $a(n)$ is calculated by using a first order lowpass filter.

$$a(n) = \alpha a(n-1) + (1-\alpha)x(n) \quad (3)$$

Peak $p(n)$ and valley $v(n)$ are obtained by lowpass filtering the local peaks and valleys with a fast attack response (towards the peak/valley) and slow release response (towards the average).

$$p(n) = \begin{cases} \beta p(n-1) + (1-\beta)x(n) & \text{for } x(n) \geq p(n-1) \\ \gamma p(n-1) + (1-\gamma)a(n) & \text{otherwise} \end{cases} \quad (4)$$

$$v(n) = \begin{cases} \beta v(n-1) + (1-\beta)x(n) & \text{for } x(n) \leq v(n-1) \\ \gamma v(n-1) + (1-\gamma)a(n) & \text{otherwise} \end{cases} \quad (5)$$

After experimenting with several set of values, it was found that α = 0.997, β = 0.5, and γ = 0.995 give satisfactory performance. The response times corresponding to these values of α, β, and γ are 86.2 ms, 0.4 ms, and 40.8 ms respectively. Response time is calculated by measuring the time during which the step response rises from 10% to 90% of its maximum value.

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

$$y(n) = \begin{cases} 1 & \text{if } x(n) \geq \theta_U \\ -1 & \text{if } x(n) \leq \theta_L \\ y(n-1) & \text{otherwise} \end{cases} \quad (6)$$

Pitch period is measured by counting the number of samples in each cycle of $y(n)$ and the value is updated at zero crossing of $y(n)$. These values are used for calculating pitch frequency waveform Fx and for obtaining histograms.

For calculating histogram, the frequency range from 50 Hz to 1000 Hz is logarithmically divided into bins. After detection of a new pitch period, the content of a bin corresponding to the pitch is incremented. The increment to be added is 1 for frequency-histogram, and equal to the number of samples in the pitch period for time-histogram The histogram is obtained by normalizing the content of each bin by the summation of the contents of all the bins. For *m*- pitch period histograms, the content of a particular bin is incremented only when *m* successive pitch periods have pitch frequencies in same bin range. A strong peak is indicative of the stability of the pitch frequency with respect to time.

The user interface of the program is menu driven. There are five display windows, which can be used for displaying input waveforms, intermediate processed waveforms, or the final histograms. The displays have cursors and facility for scaling. A segment of a waveform can be selected or playback and processing.

Proc. Bio Vision 2001 Int. Conf. Biomed. Eng., Bangalore, India, Dec. 21-24, 2001, pp. 1B.33-37

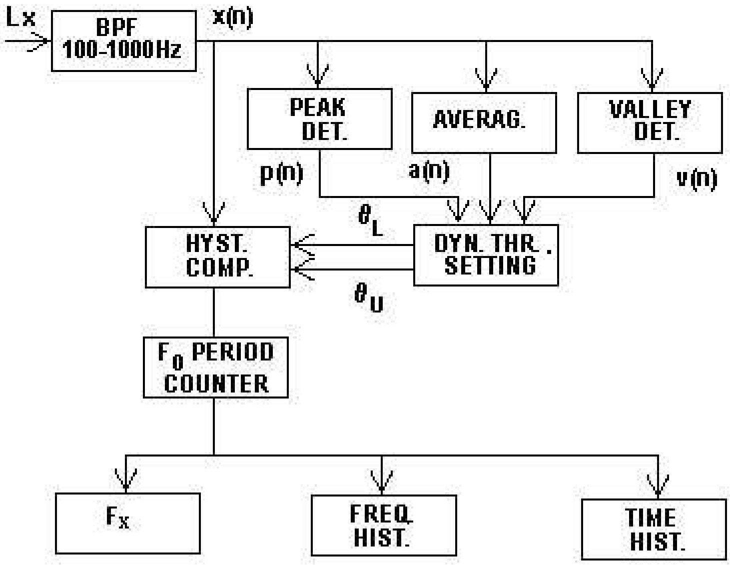


Fig. 2. Signal processing.

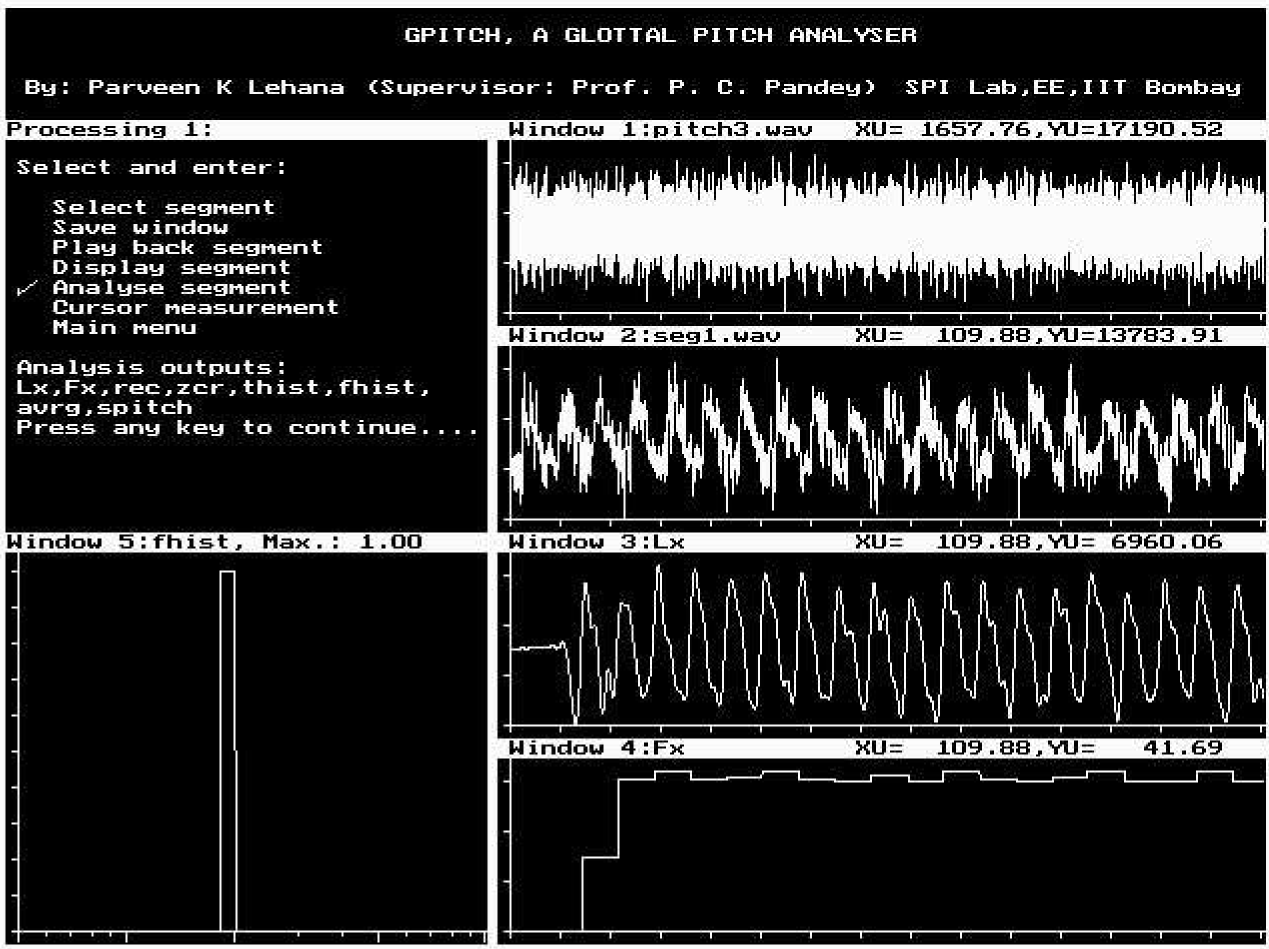


Fig. 3. Graphical user interface screen of the glottal pitch analyzer.

Proc. Bio Vision 2001 Int. Conf. Biomed. Eng., Bangalore, India, Dec. 21-24, 2001, pp. 1B.33-37

IV. RESULTS & DISCUSSION

The sensitivity of the impedance detector module was tested using a glottal impedance simulator circuit, with 0.5 Ω switching resistance superimposed on 500 Ω base impedance. A current of 1 mA was injected into simulator terminals and this resulted in 0.275 V peak-to-peak demodulated output, indicating a sensitivity of 550 mV/Ω. The signal-processing set-up was tested for satisfactory operation using frequency swept inputs.

Fig. 3 shows the display screen with the glottal impedance waveform, obtained from first author's thyroid cartilage, for the sustained vowel /*a*/. Sampling rate used was 11.025 kHz. The first window shows the acquired waveform for duration of about 2 seconds. Second window shows a selected segment of the waveform and the bandpass filtered waveform is shown as Lx in the third window. It shows the Fx waveform and single period frequency histogram in window 4 and window 5 respectively. The histogram is obtained with the F_0 range of 50 Hz to 1000 Hz logarithmically divided in 50 bins. The Lx waveform has nineteen pitch periods and Fx waveform shows tracking of pitch frequency F_0. Over this short segment, the histogram shows a narrow distribution of pitch frequency.

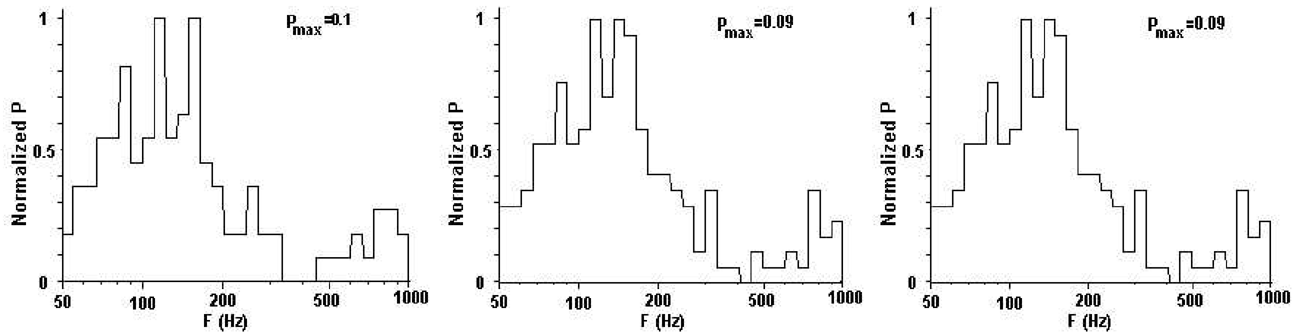
For histographic analysis, a 10-minutes long Hindi story was read by a male speaker (32 years) with normal voice. Speech and glottal signals were recorded. The whole Lx waveform, so obtained was divided into four segments. Three of these segments were of 3 minutes duration and were selected from beginning, middle, and end of the waveform, respectively. Fourth segment selected was of 9 minutes duration, discarding 0.5 minutes at ends. Single period and triple period histographic analysis was performed on these segments for obtaining frequency- and time-histograms. The histogram obtained after analysis are shown in Fig. 4.

Fourcin [2] has indicated that healthy larynx characteristically has sharp edges to its frequency distribution and modal peaks in the histograms are characteristic of the speaker. These peaks result from the auditory control of his prosody and are ordinarily stable. Fig. 4a shows single period frequency-histograms for the four segments. We see that histogram shapes for the four segments are very similar, as expected for a speaker with normal voice.

Fig. 4b shows single period time- and triple period time- as well as frequency-histograms for the 9-minute segment. Time- and frequency-histograms look almost similar except low frequency emphasis in case of time-histograms. Triple period histograms shows stability of the pitch frequency.

V. CONCLUSION

In summary, a low cost impedance glottograph has been developed. A glottal pitch analyzer has been realized by interfacing the glottograph to the sound card and graphical user interface and analysis program on a PC. The instrument can be used for studying voice disorders as well as speech analysis and synthesis applications.



(A) (B) (C)

Fig. 4a. Single period frequency histograms for the three segments each of 3-minutes duration.

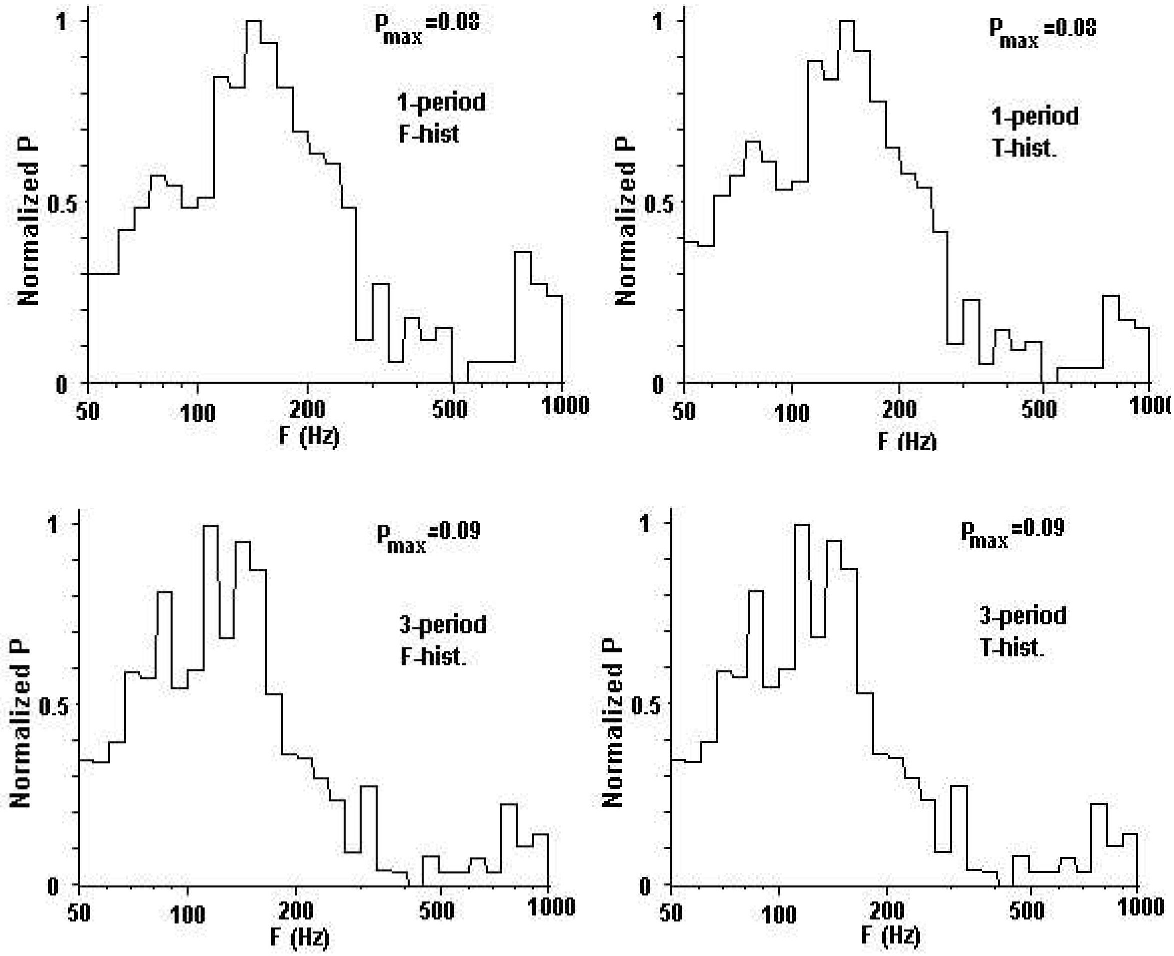


Fig. 4b. Single period and triple period frequency- & time-histograms for 9-minute segment.

References

[1] K.E. Cumming and M.A. Clements, “Improvements to and applications of analysis of stressed speech using glottal waveforms,” *Proc. IEEE ICASSP*, vol. 2, pp. 25-28, 1992.

[2] A.J. Fourcin, “Laryngographic assessment of phonatory function,” *ASHA Reports*, vol. 11, 1981.

[3] *F-J Electronics A/S, Electrograph (type EG 810)*, Vangede Bygade 114, DK-2820 Gentofte, Denmark.

[4] R.H. Colton and E.G. Conture, “Problems and pitfalls of electro-glottography,” *Journal of Voice*, 4, pp. 10-24, 1990.

[5] D.G. Childers, "Assessment of the acoustics of voice production," *NIDCD Monograph*, pp. 63-83, 1991.

[6] M.S. Chitnis and P.C. Pandey, “Extraction of voice pitch by measuring impedance variation across the thyroid cartilage,” *Proc. National Conf. Biomedical Engineering*, April 9-11, pp. II-10 to II-12, 1998.

[7] A.V. Patil, “A glottal pitch extractor,” *M. Tech. dissetation*, Dept. of Electrical Engineering, IIT, Bombay, Jan. 2000.