Inter-Channel Parameters for the Diagnosis of the Laryngeal Functions

Jong Min Choi, °Seung-Ho Choi, *Myung-Whun Sung, **Kwang Suk Park

Interdisciplinary Program, Medical and Biological Engineering Major,
*Department of Otorhinolaryngology Head and Neck Surgery, College of Medicine,
**Department of Biomedical Engineering, College of Medicine,
Seoul National University, Seoul, Korea
°Department of Otorhinolaryngology Head and Neck Surgery, College of Medicine,
Hallym University, Seoul, Korea

Abstract - **Speech, airflow and electroglottogram are widely used for the diagnosis of the laryngeal function. But, in most cases these signals were analysed separately and only combined subjectively by the reviewers. By simultaneous measurements and integration of these signals objectively, we can derive new diagnostic information quantitatively. Here, we propose a new inter-channel parameter CDR (Closure Delay Ratio) that is estimated from three-channel signals measured simultaneously; speech, EGG and airflow. This new parameter CDR makes the classification of patients with larynx diseases from normal subjects more confident.**
Keywords - **larynx disease, EGG, speech, airflow**

I. INTRODUCTION

Many people have vocal fold diseases. For example, laryngeal paralysis, vocal fold nodules and polyps, cancer in vocal folds are among them. These diseases can be diagnosed by measuring the bio-signals. We use several bio-signals for the diagnosis of the laryngeal function. Traditionally Speech is used most widely, electroglottogram (EGG) measuring the change of electrical impedance in vocal folds has been used for 20 years, airflow, subglottal air pressure are also examples. By using these signals, we can diagnose the function of the larynx noninvasively and effectively.

Each signal has some parameters and these are used for classification of laryngeal pathologies[1]. In case of speech, jitter using fundamental frequency and shimmer using peak-to-peak value are important parameters. In case of EGG, opening time and closing time are very significant features[2]. In case of aerodynamics, average flow and subglottal pressure are key values.

Even though speech, airflow and EGG have many important parameters, they were used separately only. In previous study, analysis using 'speech and EGG' [3],[4] or 'EGG and aerodynamic features'[5] was proposed, but this is neither a complete integrated analysis nor a real-time analysis. By simultaneous measurements of these signals, we can integrate information for the diagnosis of disease. So, we developed an instrument which can measure three signals simultaneously; speech, EGG and airflow. Simultaneous measurement makes it more possible to derive diagnostic parameters for the diagnosis of larynx.

II. THE DIAGNOSTIC SYSTEM OF LARYNX

We developed 3-channel laryngeal diagnostic system to measure speech, EGG, airflow simultaneously. This system is composed of hardware part and software part. Speech, EGG, airflow signals are obtained, amplified, and filtered in hardware part and then signals are transferred to computer for analysis. Software part has the function of real-time data acquisition, real-time data analysis.
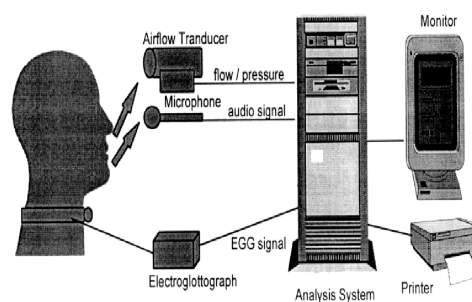

Figure 1. Schematic of the Multi-channel Laryngeal Diagnostic System

There are microphone for obtaining speech data, mesh(air flow transducer) for measuring airflow and EGG sensor in hardware part. Because obtained signals have small value, they are amplified. Also because signals are analogue, that is converted to digital signals by using 16bit resolution data

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle	Contract Number
Inter-Channel Parameters for the Diagnosis of the Laryngeal Functions	
	Grant Number
	Program Element Number

Author(s)	Project Number
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
Interdisciplinary Program, Medical and Biological Engineering Major Seoul National University Seoul, Korea	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group PSC 802 Box 15 FPO AE 09499-1500	
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.

Abstract

Subject Terms

Report Classification	Classification of this page
unclassified	unclassified

Classification of Abstract	Limitation of Abstract
unclassified	UU

Number of Pages
3

acquisition board(SD-201, Comizoa, Korea). And then real-time data viewer shows speech, EGG, airflow data in CRT monitor and analyses parameters. Our real-time data viewer coded by Delphi 5(Borland, USA).
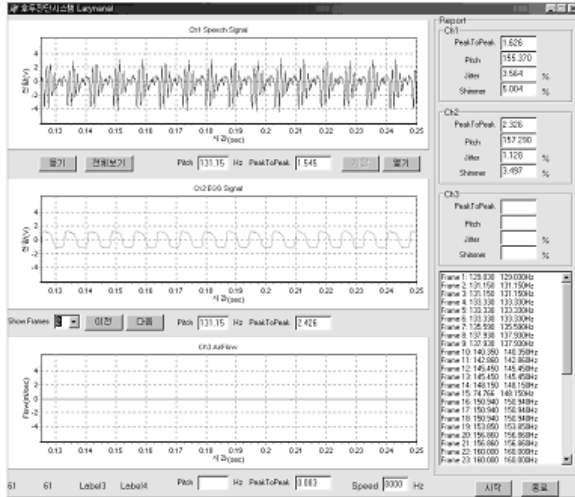

Figure 2. Example of analysis using our real-time synchronous data viewer

III. PARAMETERS

The most important parameter in speech signal and EGG is fundamental frequency(F0). In patients, when he or she pronounces /a:/ in comfortable tone fundamental frequency varies irregularly in each frame. So, jitter estimated by F0 is also important. Likely, amplitude is an important feature and shimmer estimated by peak-to-peak value in each frame is essential for analysis.

In speech signal processing, fundamental frequency is referred to as pitch. In this paper, it is assumed that pitch is equal to fundamental frequency. Pitch detection is a core work in speech, EGG signal processing because accurately estimated pitch makes the classification of larynx disease better. And also we consider this work is real-time signal processing. So, we select pitch detection method using Average Magnitude Difference Function(AMDF)[6]. This method needs smaller amount of calculation than other method – autocorrelation method, FFT etc. The formula of AMDF is as follows.

$$AMDF(t) = \frac{1}{L} \sum_{i=1}^{L} | s(i) - s(i-t) |$$

After pitch detection and peak-to-peak value extraction in each frame, jitter and shimmer are estimated. Jitter means how often changes the frequency of patient's pronunciation over all frames, and shimmer means how much changes the amplitude of patient's pronunciation. These parameters have been used for last 50 years in speech pathologies classification, and until now are significant. The formula of jitter and shimmer is as follows.

$$Jitter = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} | T_0(i) - T_0(i+1) |}{\frac{1}{N} \sum_{i=1}^{N} T_0(i)} * 100(\%)$$

$$Shimmer = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} | A(i) - A(i+1) |}{\frac{1}{N} \sum_{i=1}^{N} A(i)} * 100(\%)$$

In airflow analysis, mean flow per second is a valuable feature. In normal subjects, mean flow per second is almost constant when pronounce /a:/ but patients does not pronounce like it.

Figure 2 shows these parameters of speech, EGG, airflow in one frame in real-time.

IV. EXPERIMENTS

We propose new parameter Close Delay Ratio(CDR). In previous study, bio-signals have only their parameters and they are analysed respectively. These trends have problems that cannot reflect the relation among bio-signals. CDR parameters can reflect the properties of speech and EGG signal obtaining synchronously.
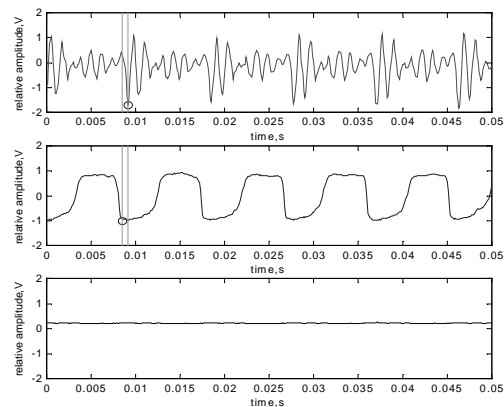

Figure 3. Example of CDR parameter in Normal Subject

Like Figure 3, CDR parameter is defined as a value that divide a difference value between closing

instance of vocal folds in EGG and end point of speech signal by pitch period, so this definition include the advantage of synchronous measurement.

$$CDR = \frac{|GCI - EP(speech)|}{period} * 100(\%)$$

Speech, EGG, airflow signals of 5 normal subjects and 5 patients with vocal folds polyp are acquired synchronously, and analysed the parameters in real-time processing. Data were recorded during sustained phonation of the vowel /a:/ at comfortable pitch and loudness level. Sampling rate for data acquisition is 8kHz per channel.

V. RESULTS

In patients, each signal is unstable than that of normal subject. Hence, we expect that CDR is larger. Figure 4 shows this aspect.
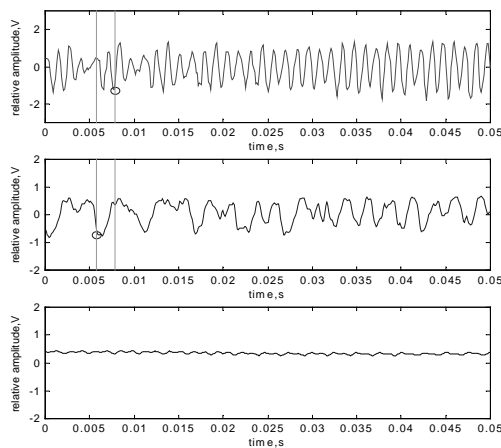

Figure 4. Example of CDR parameter in patient with vocal fold polyp

CDR of patients is larger than that of normal subjects. Table 1 shows this result. CDR of patients is almost twice than that of normal subjects. Mean value of CDR parameter in 5 patients and 5 normal subjects respectively.

	In Patients	In Normal Subject
CDR	23.5%	9.1%

Table 1. Mean value of CDR parameter

CDR parameter seems to be a good parameter because it succeeds in classifying the patients into the normal subject as seen above. But more studies and more experiments needs.

VI. DISCUSSION & CONCLUSION

This multi-channel laryngeal diagnostic system has many advantages compared with existing instruments. First, it is possible to measure speech, EGG and air flow signal simultaneously. It is significant that we can obtain more information of patient's disease in several aspects. Second, our system makes real-time data processing possible. For that, the patients cannot waste their time, and for only several pronunciations they can get the information of their larynx immediately.

In the future, some parameters using speech, EGG, and airflow signal to measure simultaneously must be proposed because more significant parameters needs for more precise classification of the laryngeal diseases.

References

[1] Marcelo de O Rosa, Jose C. Pereira, Marcos Grellet, "Adaptive Estimation of Residue Signal for Voice Pathology Diagnosis", IEEE Trans. Biomed. Eng., vol. 47, pp.96-104, January 2000;

[2] Jack J. Jiang, Shuangyi Tang, Michel Dalal, Chi-haur Wu, David G. Hanson. "Integrated Analyzer and Classifier of Glottographic Signals", IEEE Trans. Rehab. Eng., vol. 6, pp.227-233, June 1998

[3] Ashok K. Krishnamurthy, Donald G. Childers, "Two-Channel Speech Analysis", IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, pp.730-743, August 1986

[4] Donald G. Childers, Keun Sung Bae, "Detection of Laryngeal Function Using Speech and Electroglottographic Data", IEEE Trans. Biomed. Eng., vol. 39, pp.19-25, January 1992

[5] Jeff Kuo, Eva B. Holmberg, Robert E. Hillman, "Discriminating Speakers with Vocal Nodules Using Aerodynamic and Acoustic Features", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, 1999, pp. 77-80

[6] Hae Young Kim, Jae Sung Lee, Myung-Whun Sung, Kwang Hyun Kim, Kwang Suk Park, "Pitch detection with average magnitude difference function using adaptive threshold algorithm for estimating shimmer and jitter", in Proc. IEEE Int. Conf. Medicine, Biology, 1998, pp. 3162-3165