33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

A Neck Mounted Interface for Sensing the Swallowing Activity based on Swallowing Sound

Masahiro Nagae and Kenji Suzuki

Abstract— **This study proposes a novel interface for sensing swallowing activities (based on the sound produced when swallowing) that is mounted on the back of the neck. Recently, owing to the aging of the population, the number of dysphagia patients has been increasing. Although various types of diacrisis and dysphagia testing methodologies already exist, several problems remain. For example, the videofluorography methodology requires extensive preparation; it presents risks of aspiration and demands knowledge about diacrisis, and it must be performed in a medical facility. This study introduces a wearable sensing system that can be used in daily life. As it gives direct, real-time information regarding swallowing movements, this device is also aimed at being used for rehabilitation purposes. The proposed system uses a new evaluation technique that is based on the duration of the swallowing action.**

I. INTRODUCTION

Recently, the number of cases of dysphagia has increased [1]. Dysphagia is caused by a diminished ability to swallow, owing to aging or symptoms related to paralysis and apoplexy. Dysphagia results in difficulties in swallowing and eating and carries significantly increased risks of aspiration. Aspiration in turn may cause asphyxia, recurrent pneumonia, and, in the worse cases, may even result in patient death. Therefore, the evaluation of the severity of the disorder is important when planning efficient rehabilitation.

The repetitive saliva swallowing test (RSST) and videofluorography (VF) are standard methods for diagnosis of diacrisis of dysphagia. RSST is a quick, simple test to determine whether there is an episode when swallowing. However, it is unsuitable for exact diacrisis [2]. VF can determine the condition of swallowing from the movement of a bolus. However, it is a long procedure and presents the risk of aspiration of barium. A number of studies have analyzed swallowing activities, for example, acoustical modeling of swallowing sound mechanisms [3] and biomedical analysis of swallowing behavior [4]. However, none of these techniques is suitable for home rehabilitation. Other studies aimed at evaluating swallowing through the use of electromyography; however, this method does not give information relative to the movement of a bolus [5].

Recently, cervical auscultation through the use of a stethoscope has also been proposed [6]. This auscultation technique is based on the detection of sound signals obtained from the neck. This method allows the estimation of the condition of the swallowing activity, but requires analysis from a skilled practitioner to obtain a detailed diagnosis. Related to this methodology, studies have been aimed at evaluating swallowing activity using frequency analysis of the emitted sound [7]. However, these approaches also require a large apparatus and specialized diagnostic skills. In addition, they do not give real-time feedback information to the patient and caretakers.

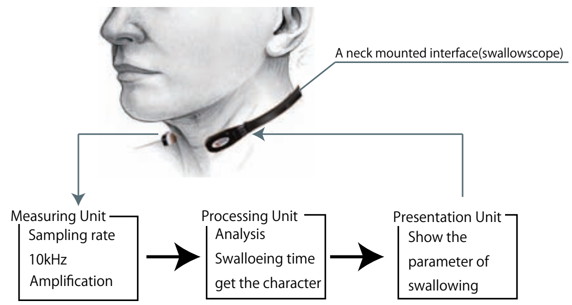

Fig. 1. System overview

In this paper, we introduce an interface that is able to measure the sound, allowing one to record the swallowing activity in real time. The interface is mounted on the back of the neck and it does not interfere with the patient's natural behavior. Information on whether swallowing proceeds normally or not is very important for efficient rehabilitation. The purpose of this work is twofold: (1) to evaluate the information on swallowing activity in real time and (2) to improve the rehabilitation of dysphagia by having patients evaluate their condition by themselves through tracking of an index of their swallowing activity.

II. SYSTEM DESIGN AND IMPLEMENTATION

A. Interface Configuration

Fig.1 shows the developed interface, which uses an electret condenser throat microphone. This microphone's advantage is its ability to select exclusively the sound surrounding the throat, without detecting other, environmental noises. Our system consists of three parts: (1) the measurement unit, which detects the swallowing sound using the microphone; (2) the processing unit, which evaluates swallowing activities; and (3) the presentation unit, which displays the data.

In the measurement unit, the microphone, which is placed at the back of the neck, samples sound at a rate of 10[kHz]. The microphone sensitivity is -66$\pm$3[dB]; the input voltage is amplified 200 times. In the processing unit, the sound signal extracted from the microphone is analyzed using MATLAB.

This work was not supported by any organization
Masahiro Nagae is with Depertment of Inteligent Technologies Graduate school of Systems and Information Engineering, University of Tsukuba. 1-1-1 Tennodai Tsukuba 305-8573, Japan masahiro@ai.iit.tsukuba.ac.jp
Kenji Suzuki is with Depertment of Inteligent Technologies Graduate school of Systems and Information Engineering, University of Tsukuba. 1-1-1 Tennodai Tsukuba 305-8573, Japan kenji@ieee.org

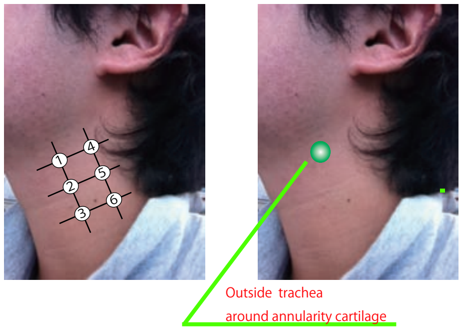

Fig. 2. The measuring location

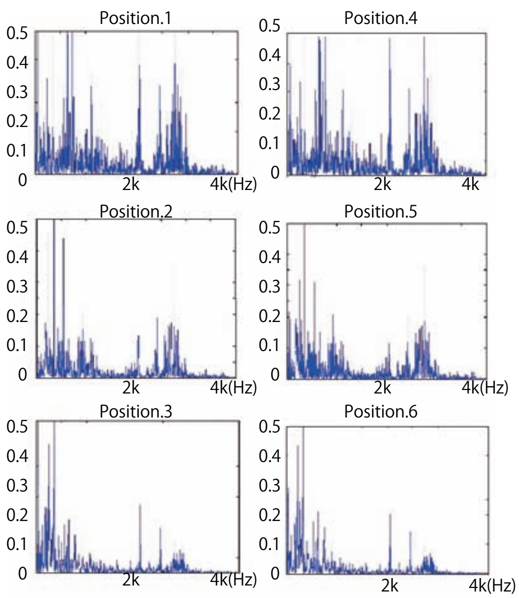

Fig. 3. The frequency response at each measuring location

B. Analytical Method

It is necessary to discriminate between sounds emitted by swallowing actions from sounds originating from coughing or vocalizing. Frequency analysis might be used for this discrimination. However, as a real-time response is an important feature of our system, wavelet-transform analysis with a Gaussian window was used. The wavelet-transform $C(a, b)$ and the basis function with a Gaussian window $\psi(t)$ are calculated via the following equations:

$$C(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t)\psi(\frac{t-b}{a})dt \quad (1)$$

$$\psi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} e^{i\omega t} \quad (2)$$

where $f(t)$ is the original signal data, a is a factor related to the frequency scale, and b is a factor related to the shift in the time scale. In this research, a was set to a 32-point scale.

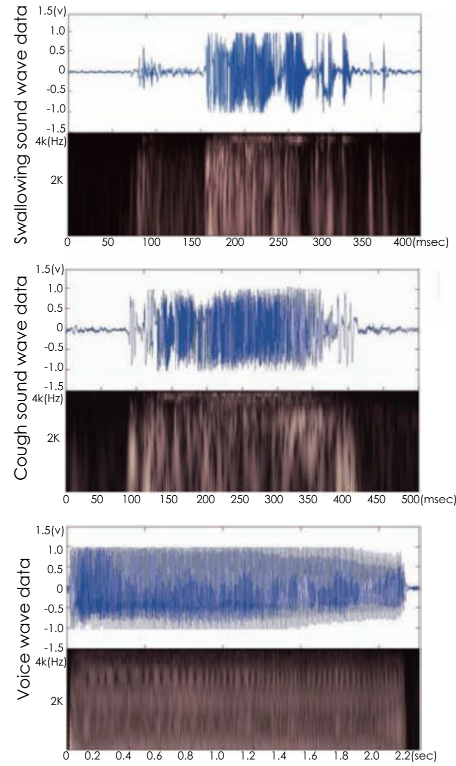

Fig. 4. Result of wavelet-transform analysis

C. Presentation of Swallowing Time

In the presentation unit, based on analyzed data, LEDs are used to give feedback about the condition of swallowing activity. This feedback information consists of three possible patterns that are displayed depending on the result: (1) a blue light indicates the start of measurement, (2) a green light indicates that swallowing was detected as normal, and (3) a red light indicates either that an abnormal sound was detected or that the swallowing sound was longer than usual. This information, given solely by the LEDs, can be intuitively understood.

D. Method of Abstraction of Swallowing

III. EXPERIMENTAL EVALUATION

A. Assessment of Measurement Position

For reasons related to the physiological activities involved in swallowing, the measurement device (to detect the swallowing sound) is recommended to be placed outside the trachea surrounding the annular cartilage [8]. Because of the differences in the body's shape and motion, however, it is difficult to perform measurements at the exact same location each time. In this section, we consider the influence of the measurement location on the frequency and intensity of the detected signal.

Fig.2 shows the locations at which measurements were taken during the experiment; Fig.3 shows the result of the frequency analysis performed at each location. Data loss can be observed at locations 3 and 6. This loss is due both to their

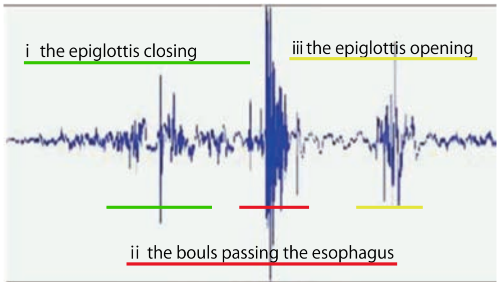

Fig. 5. Character of the swallowing sound wave

distance from the trachea surrounding the annular cartilage and to the influence of the muscles of the neck. However, as these losses correspond to low-frequency components only, they do not alter the final results. It can thus be considered that the extraction of swallowing sounds is both practicable and free of restrictions on the location of measurement.

Background noise such as conversations is included in any real life environment. Coughing is not only an additional event to be discriminated from the swallowing sound, but also an important event for the detection of aspiration and choking. We configured the proposed device to use both frequency and amplitude characteristics to discriminate among the sounds of swallowing, vocalizing, and coughing. "Signal peaks" refers to the number of times the input voltage exceeds a given threshold θ. We set θ=0.5[V] in this study.

Fig.4 shows the result of wavelet-transform analysis of the sounds of swallowing, vocalizing, and coughing. A swallowing sound is the result of three separate sounds: (1) the sound of the epiglottis closing, (2) the sound of the bolus passing the esophagus, and (3) the sound of the epiglottis opening. As a result, it exhibits a non-continuous wave, as shown in Fig.5. The number of peaks of length τ seconds, P[τ], at time t is estimated by the following equations:

$$P[\tau] = \sum_{i=t-\tau}^{t} s[i]Ts \tag{3}$$

where s[i] is the input voltage. Ts is sampling rate. For τ=100[ms], for the swallowing sound, P[τ]=100. For a cough, P[τ]=190. For vocalization, P[τ]$\geq$250. Further, it can been from the results of the wavelet-transform analysis that the voice wave has a weaker high-frequency component than the other waves and has peaks below 500[Hz]. Compared to the swallowing sound or a cough, vocalization waves show 500[Hz] peaks at their beginning and end. In addition, at approximately 1500[Hz], the vocalization wave exhibits a more powerful spectrum than the swallowing wave. These results show that swallowing sounds, coughing, and vocalizations have distinct characteristics, allowing them to be discriminated using acoustic analysis. We also conducted experiments to investigate possible differences in swallowing activity between elderly and young persons. The test subjects in the younger group were three males in their twenties. The older group consisted of four males and four females ranging in age from 50 to 80 years. The test subjects were instructed to swallow their saliva or a small volume of water on several different occasions.

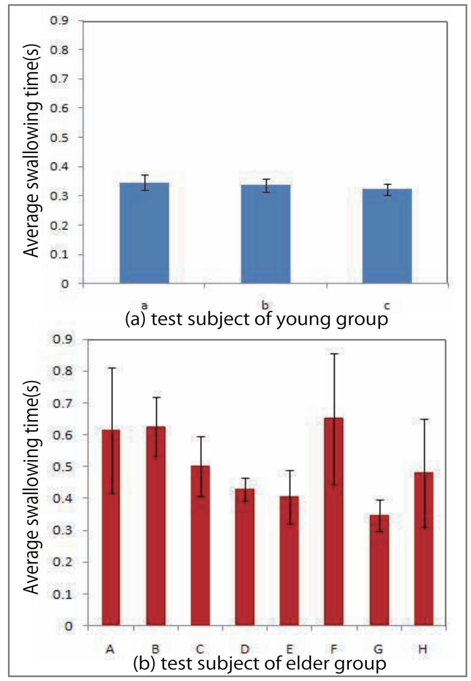

Fig. 6. Average swallowing time

Fig.6 shows the results for the younger group and the results for the older group. In the younger group, no significant differences among individuals were observed. However, in the older group, great differences among individuals were observed: swallowing action tends to last for a longer time. In addition, in the older group, variations in swallowing time were observed for the same subject. Furthermore, choking after a long swallowing action or a pattern consisting of consecutive short swallowing actions was observed. Such events occurred when a residue of the bolus failed in the oral cavity, causing the reflex compression of the bolus to the esophagus [9]. These features of the swallowing actions cause aspiration. From these results, we conclude that the duration of swallowing action has a tendency to decrease with age.

We define T_2/T_1 as the assessment function of swallowing time, where T_1 denotes the total length of the swallowing action and T_2 refers to the time from the start of the swallowing action until the bolus passes through the esophagus.

Table.I shows an example of an evaluation using the T_2/T_1 assessment function for the same young subject. Table.II shows an example of an evaluation using the T_2/T_1 assessment function for the same older subject. For the younger group, T_2/T_1 remains close to 0.5. However, for the older group, T_2/T_1 ranges from 0.3 to 0.77; that is, there is variation in the duration of the swallowing action. In addition, there are large differences within the test subject data because of choking or long swallowing time. As it allows discrimination between younger and older subjects, this assessment function can be considered applicable to the evaluation of swallowing activity.

TABLE I

T_2/T_1 OF TEST SUBJECT A (YOUNGER GROUP)

swallowing time (s)	0.31	0.29	0.32	0.30	0.30
T_2/T_1	0.51	0.51	0.52	0.51	0.47

TABLE II

T_2/T_1 OF TEST SUBJECT A (OLDER GROUP)

swallowing time (s)	0.45	0.81	0.39	0.82	0.80
T_2/T_1	0.37	0.40	0.56	0.66	0.32

B. Automatic Detection

We conducted experiments to verify the accuracy of the swallowing detection. We instructed three test subjects to swallow 15 times in a row and to perform swallowing, coughing, and vocalization freely. From 45 swallowing events, our interface could discriminate all 45 swallowing sounds correctly. However, several false positives were returned when the subject coughed or vocalized. A cough is a very important indicator of the swallowing condition because coughing after swallowing is a physiological activity allowing aspiration avoidance. False positive signals were observed when coughs or vocalization was performed in a row; our system required a 0.2-s period of silence to avoid false positives. However, if our proposed system cannot recognize sounds produced during the overlap of different actions, this does not represent an issue for evaluation of swallowing: a swallowing action is surrounded by a silent apneic time. Consequently, a cough after a swallowing action can be recognized accurately. The requirements for use of this tool for dysphagia rehabilitation are therefore met.

C. Development of a Neck Mounted Interface

We developed a neck-mounted interface and conducted an experiment to visualize the swallowing activity and coughs using LEDs. Fig.7 shows the information collected when swallowing. It was prepared as follows. First, the sound data extracted from the measurement unit was converted into digital data and fed to the computer, which stored it. The data was then analyzed entry-by-entry. If the input voltage was higher than a given threshold level, the input data was considered "valid." Valid data is accumulated on the computer (up to a maximum of 1000 measurements). If there is no valid measurement, the data is considered "silent data." If silent data is registered t_0 seconds after valid data has been registered, the sound segment is considered to have been completed. The frequency of the resulting segment is analyzed and determined to be swallowing, coughing, or vocalization. As apneic periods of several milliseconds surround a swallowing event [6], we set t_0= 0.2[s]. Once discrimination has been performed, the result is recorded and displayed on the LED presentation unit and on the computer. Fig.8 shows the LED states during the visualization of swallowing and coughing.

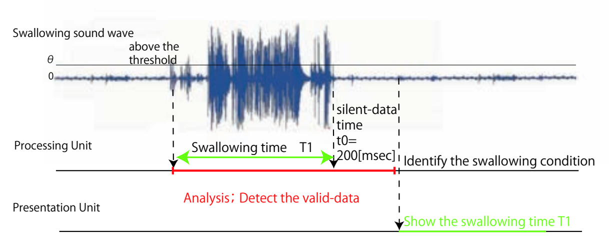

Fig. 7. Presentation of swallowing

Fig. 8. Neck mounted interface (Swallowscope)

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a novel method for the evaluation of swallowing activity by means of an acoustic analysis. We developed an interface (to be mounted at the back of the neck) that evaluates swallowing activities that are normally hidden, physiological activities. Experiments were conducted to measure the reliability of the proposed device. Contrary to the results of already existing methods, the proposed interface is noninvasive and can be used in day-to-day life. By displaying real-time information about swallowing actions, the device also aims at being an efficient rehabilitation tool. In the future, work aimed at providing a more detailed abstraction of the swallowing sound characteristics will be conducted, so as to give more precise information about the conditions of the swallowing action and to predict possible aspiration. We also plan to develop an interface that uses functional electric stimulation to reestablish normal eating ability in dysphagia patients.

REFERENCES

[1] H. Taniguchi, S. Ootaki, Y. Kajii, Y. Yamada, and M. Inoue, "Relationship between muscular atrophy of the tongue and impairment of bolus transport during swallowing in patients with amyotrophic lateral sclerosis", *The Journal of Japanese Society of Stomatognathic Function*, vol. 15, no. 1, pp. 30-37, 2008.

[2] K. Oguchi, E. Saitoh, M. Baba, S. Kusudo, T. Tanaka, and K. Onogi, "The repetitive saliva swallowing test (RSST) as a screening test of functional dysphagia: (2) Validity of RSST," *The Japanese Journal of Rehabilitation Medicine*, vol. 37, no. 6, pp. 383-388, 2000.

[3] S. S. Shirazi and Z. Moussavi, "Acoustical modeling of swallowing mechanism," *32nd IEEE EMBS Conf.*, pp.13-16, 2010.

[4] C. S. S. Greco, L. G. M. Q. Nunes, and P. L. Melo, "Instrumentation for bedside analysis of swallowing disorders," *32nd IEEE EMBS Conf.*, pp. 923-926, 2010.

[5] T. Igasaki, H. Koga, H. Ninomiya, and N. Murayama, "Is the active electrode for surface EMGs applicable in evaluation of swallowing?" *IEICE technical report. ME and Bio Cybernetics*, vol. 108, no. 405, pp. 1-5, 2009.

[6] S. Oyado, "Rehabilitation of dysphagia with cervical stethoscopy," *Nisoken*, 2009.

[7] T. Sato, K. Komai, K. Kato, N. Kawashima, T. Agishi, M. Akamatsu, and T. Ando, "Diagnosis of swallowing difficulties using time frequency analysis of swallowing sound signals based on wavelet transformation," *ASIO Journal*, vol. 51, no. 2, p. 15A, 2005.

[8] R. L. Drake, W. Vogl, and A. W. M. Mitchell, *GRAY'S Anatomy for Students*, Elsevier: Churchill Livingstone, 2005.

[9] M. Saito, "The study of dysphagia (in Japanese)," *Journal of Cookery Science of Japan*, vol. 43, no. 5, pp. 281-285, 2010.