Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Analysis of Temporal Pattern of Swallowing Mechanism

Mohammad Aboofazeli, *Student Member, IEEE*, and Zahra Moussavi, *Senior Member, IEEE*

Abstract- **This paper presents an objective method for analysis of temporal pattern of swallowing mechanism based on analysis of swallowing sounds and submental surface electromyogram (EMG). In this study, swallowing sound signal and submental EMG of 12 healthy subjects were recorded. Swallowing sound signals were divided into 25 millisecond segments each of which was represented by waveform fractal dimension (WFD). Temporal pattern of swallowing sound signal was identified based on hidden Markov modeling (HMM) of the WFD sequences. Submental muscle contraction was marked by thresholding the RMS values of the EMG signals. Duration of the swallowing sound phases, duration of the submental muscle contraction, and time difference between the onset of submental muscle contraction and the opening of cricopharynx were calculated. Experimental results suggest that the proposed method is efficient in the study of temporal pattern of swallowing mechanism and can provide an objective and accurate approach for swallowing mechanism analysis.**

Keywords- **Hidden Markov model, submental surface electromyogram, swallowing sounds, waveform fractal dimension**

I. INTRODUCTION

Swallowing mechanism is a complex process of several events in which the proper timing of the events is a key factor. Lack of coordination in swallowing mechanism events may result in aspiration (food or liquid entry into the airway). Aspiration is very likely in patients with swallowing disorder (dysphagia). Dysphagia occurs as a result of congenital abnormalities, structural damage, and/or medical conditions [1, 2].

The process of normal swallowing may be divided into three distinct phases [1]: oral, pharyngeal, and esophageal phases. In the oral phase, the tongue propels the bolus toward the pharynx. During pharyngeal phase, closure of the velopharyngeal port in addition to the elevation of the hyoid and larynx and closure of the larynx and epiglottis prevent the bolus from entering the nasal cavity and the airway. Also, Opening of the cricopharyngeal sphincter allows the bolus to pass from the pharynx into the esophagus. In the final phase, esophageal peristalsis moves the bolus into the stomach.

One of the primary tools for swallowing assessment is cervical auscultation (listening to the swallowing sounds using a stethoscope). However, Information obtained by this method depends on the skills of the examiner [3]. In recent years, swallowing sound analysis has drawn more attention in which swallowing sounds are recorded by microphones and/or accelerometers and analyzed by digital signal processing techniques. Some researches have shown promises on the use of swallowing sound analysis for diagnostic purposes [4-6]. Also, Temporal and durational patterns associating respiration and swallowing [7], assessment of the maturity and competence of swallowing mechanisms [8] have been studied based on swallowing sound analysis.

Swallowing sound is a non-stationary signal generated by contribution of several events in swallowing mechanism. The opening of the cricopharynx in pharyngeal phase is speculated to contribute to the most distinct characteristics of the swallowing sounds, so called the initial discrete sounds (IDS) [9]. IDS is followed by a gurgle sound, so called the bolus transmission sounds (BTS), which is due to bolus transmission into the esophagus in the pharyngeal phase. In some swallows, BTS is followed by the final discrete sounds (FDS) [9]; however, our experience shows that FDS is not necessarily present in all swallows. These three segments of a typical swallow are shown in Fig.1.

Swallowing sounds recorded from a subject, being fed the same bolus texture and size, do not have a unique temporal and durational pattern. Additionally, the three phases of swallowing sounds, i.e. IDS, BTS, and FDS, are not easily distinguishable from each other. Therefore, detection of the swallowing sounds stages is a challenging task. Fig. 2.a. presents a typical time-domain swallowing sound signal including IDS, BTS, and FDS, whereas Fig. 2.b. presents another normal swallowing sound signal from the same subject and the same texture with non-distinguishable IDS and BTS and without FDS.

Swallowing sound segmentation was addressed in [5, 10] using variance dimension and waveform fractal dimension algorithms. However, none of these two algorithms were able to detect the swallowing phases in all subjects with a high accuracy; hence the manual identification of the swallowing phases were used prior to classification of healthy signals from dysphagic ones [10]. On the other hand, HMM has offered some promising results in adaptive segmentation of the swallowing sounds [11].

HMM was used in [11] to model swallowing sound signals recorded from healthy children and showed that the

M. Aboofazeli is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada; e-mail: mohammad@ee.umanitoba.ca

Z. Moussavi is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada; e-mail: mousavi@ee.umanitoba.ca

This research was supported by Telecommunications Research Laboratories (TRLabs) and Natural Sciences and Engineering Research Council (NSERC) of Canada.

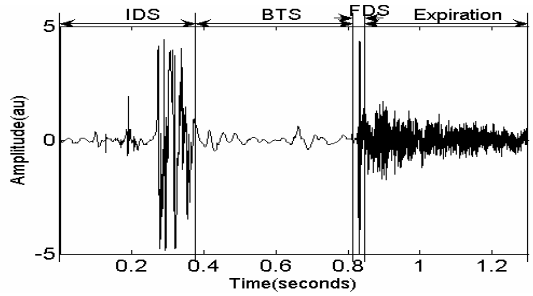

Fig.1. A typical swallowing sound signal. Legend: 'au' arbitrary units for the normalized amplitude

swallowing sounds have three states (phases) regardless of the textures.

On the other hand, electromyography (EMG) of the muscles involved in swallowing mechanism may provide additional information on the temporal pattern of muscle activities. The use of submental surface EMG signals monitors the activity of mylohyoid, geniohyoid and the anterior belly of the digastric muscles [7]. The time of the onset of submental muscle contraction was found relative to the onset of swallowing [7].

In this study, swallowing sound signal along with the surface EMG recorded from midsagittal line of the submental area were used to analyze the temporal and durational pattern of the swallowing mechanism. The RMS of the EMG signal was used for timing analysis of submental muscle activities in relation to swallowing process.

II. Methodology

A. Data

12 healthy subjects participated in this study. All subjects were young adults (ages 21-35, 9 males) and in good health without any history of respiratory or swallowing disorder, eating or nutrition problem or neurological disorder. Three recordings were conducted. In the first and second recordings, the subjects were fed pre-packed pudding

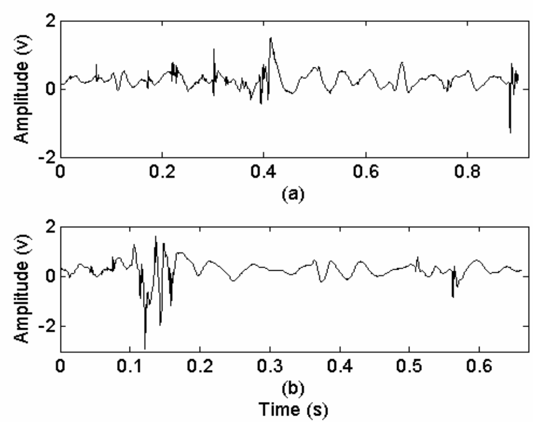

Fig.2. Two different swallowing sound signals of one subject while eating pudding.

(semisolid texture), and orange juice (thin liquid) with a bolus size of 5 ml teaspoon, respectively. In the third recording, the subjects drank orange juice by straw. In each recording trial, 5-8 swallows were recorded. For this study, in total 96 swallows with different textures and feeding patterns were analyzed.

Swallowing sounds along with tracheal breath sounds were recorded by a Siemens accelerometer (EMT25C) placed over suprasternal notch. In order to find the respiration phases, i.e. inspiration/expiration, another accelerometer was placed on the left or right second intracostal space (whereever the difference between inspiration and expiration was greater), in midclavicular line to record lung sounds. This lung sound was used only for respiratory phase information using the method introduced in [12] for manual confirmation and validation of swallowing localization in a breath cycle.

EMG signal was recorded with one bipolar channel using Ag-AgCl disposable electrodes placed over midsagittal line of the submental area. The sounds and EMG signals were amplified and bandpass filtered in the range of 30-2500 Hz and 15-500 Hz, respectively. The signals were digitized simultaneously at a 10240 Hz sampling rate. The swallowing sounds were extracted from the breath and swallowing record (the signal recorded over suprasternal notch) manually by both auditory means (listening to the sounds repeatedly) and visual inspection (watching the signal in time domain as well as its spectrogram in time-frequency domain).

In addition, data of 3 healthy adults, who had gone through videofluoroscopic (VFS) swallowing study (observing the swallowing process with a continuous stream of X-ray while they swallow barium-mixed boluses) were adopted from a previous study [5] in which the participants were fed three textures mixed with barium: pre-packaged pudding, diluted pudding and fruit juice in 5 ml bolus size. The bolus movement patterns on the videofluoroscopy screen were recorded by a VCR simultaneously with the audio signals using the same signal conditioning as the rest of the data used in this study. VFS data allows visual assessment of swallowing mechanism. Those data were used for validation of the proposed method's capability in accurately detecting the boundaries of different phases of the swallowing sound signal.

B. Hidden Markov Model of swallowing sounds

In order to detect swallowing sound signal phases, the signal was modeled by a first-order discrete hidden Markov model as described in [11]. Theory of hidden Markov model may be found in [13, 14]. To model swallowing sound signal with HMM, the signal was divided into a sequence of 25 ms (i.e. 256 samples) segments with 50% overlap between the successive segments. Each segment was then represented by its waveform fractal dimension (WFD) trajectory. WFD is a measure of the degree of complexity of a signal. In this study, WFD was calculated based on Katz algorithm [15] which is less sensitive to noise compared to other methods [16]. WFD of a waveform is calculated as:

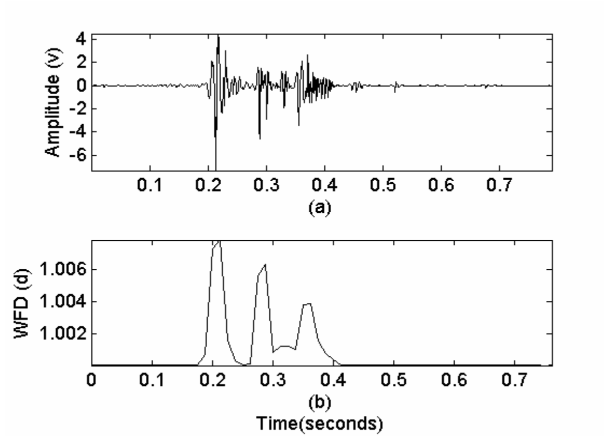

Fig. 3. (a) The waveform of swallowing sound signal in time-domain and (b) its WFD.

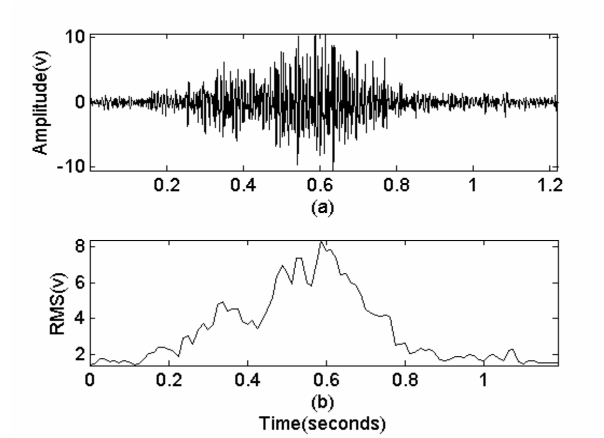

Fig. 4. (a) The waveform of EMG signal in time-domain and (b) its RMS.

$$D = \frac{\log n}{\log n + \log \dfrac{d}{L}}, \tag{2}$$

where L is the total length of the curve (the sum of distances between successive points), d is the diameter estimated as the distance between the first point of the each segment and the point of the segment that provides the farthest distance and n is the number of steps in the curve defined as $n = L / a$, where a is the average step or average distance between successive points [15]. A typical example of the swallowing sound signal in time-domain and its WFD are shown in Fig. 3.

Although the swallowing sounds heard from each subject being fed particular bolus texture are widely variable, the number of hidden states representing the swallowing mechanism was found to be fixed for most swallowing sound signals [11]. Therefore, in this study the number of states was set to 3.

In order to find model parameters (The state transition matrix and the observation probability distribution matrix) the method proposed in [11] was employed and the most likely states were determined using the Viterbi algorithm [17]. The boundaries of the detected states were evaluated by the available VFS data of the three healthy subjects.

C. Analysis of EMG timing

During swallow, the floor of the mouth muscles contract and pull the larynx and hyoid bone. This contraction results in elevation and forward movement of the larynx and hyoid bone. The elevation contributes to closure of the airway entrance and the forward movement contributes to opening of the upper esophageal sphincter. When the submental muscles start contracting, the amplitude of submental surface EMG increases drastically. Accordingly, the RMS of the EMG activities during muscle contraction is much larger than the RMS of the EMG signal when submental muscles are at rest. In this study, EMG signal was divided into a sequence of 25 ms segments with 50% overlap between the successive segments. Each segment was then represented by its RMS. A typical example of the swallowing EMG signal in time-domain and its RMS sequence are shown in Fig. 4. RMS of each segment was compared with a threshold (15% of the maximum value of the RMS sequence). This threshold level was empirically selected to ensure that strong noises were not considered as muscle contraction. When RMS was larger than the threshold, submental muscles were considered to be active.

III. Results And Discussion

The HMM method was used to detect the boundaries of the states (phases) of the swallowing sound signals and the results were evaluated with the available VFS data. The results show that the first state of the swallowing sound model determined by HMM coincides with the quiet oral phase of the swallow, in which tongue moves the bolus. The onset of the second state determined by HMM coincides with the onset of pharyngeal phase, in which IDS are generated. Additionally, the onset of the third state coincides with the beginning of bolus transmission into esophagus. A typical example of the swallowing phases detected by HMM is shown in Fig. 5. b. The advantage of using HMM in swallowing phases' boundary detection is that when there are several local maxima of the sequence of features, HMM can still detect boundaries of the swallowing sound signals correctly. As the number of frames in VFS data was 30 frames/s, time difference between the successive frames was 33.3 ms. This limited the time resolution of the evaluation. The time difference between the detected boundaries by HMM and those by manual inspection of VFS data in all swallowing sounds were less than 33.3 ms. However, if VFS data is available with higher resolution than 30 frames/s, this difference might be even less.

The results of detected swallowing sound phases by HMM and the corresponded submental muscle activities for a typical swallow when the subject was fed juice are shown in Fig. 5. As can be seen the RMS activity proceeds well before the onset of esophageal phase of swallow. Temporal

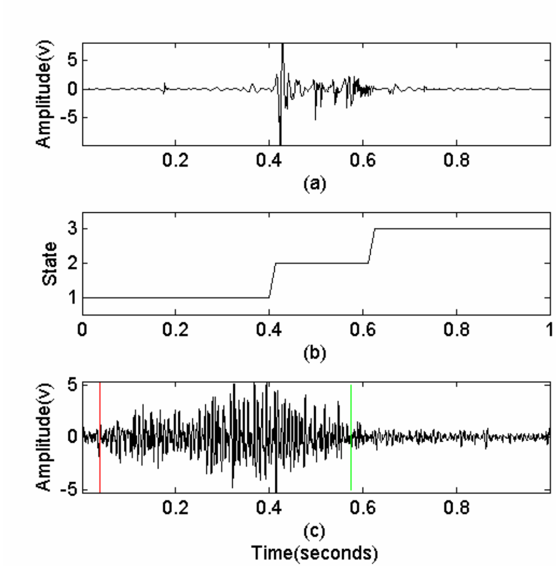

Fig. 5. (a) The waveform of swallowing sound signal, (b) the most likely states of the swallowing sound, and (c) EMG signal. Red and green vertical lines show beginning and end of muscle contraction, respectively.

pattern showed in Fig. 5 was common in all swallowing sounds and EMG signals.

Table 1 shows the mean and standard deviation values of the duration of each swallowing state (phase), the duration of submental muscle contraction, and time difference between the onset of submental muscle contraction and the opening of cricopharynx when subjects were fed thin liquid and semisolid textures. The values were averaged among all subjects. Durational patterns of swallowing mechanism are functions of the subject's interaction and the viscosity and volume of the bolus [1, 7, 9]. The results shown in Table 1 are congruent with the duration of deglutition apnea (total duration of the 3 states) reported in [9] and the duration of opening and transmission (states 2 and 3) reported in [7].

IV. Conclusion

The goal of this study was to investigate temporal and durational pattern of swallowing mechanism of healthy individuals based on analysis of swallowing sounds and submental surface EMG in relation to swallowing phases. The HMM was applied to detect the boundaries of the swallowing sounds phases and its accuracy was validated by the use of the available VFS data. The time difference between the phases' boundaries detected by HMM method and those detected manually by VFS data was less than 33 ms. Studying the EMG activities of the submental muscles showed that the muscles contraction starts well before the onset of opening of the esophagus.

This paper proposes an objective method to analyze the timing of swallowing mechanism events. The proposed method may prove useful in dynamic behavior assessment of swallowing mechanisms. Furthermore, it may be utilized in assessment of abnormalities in swallowing mechanism.

Table 1
Duration of swallowing sound phases and submental muscle contraction in different textures in seconds
(Mean ± standard deviation)

	Semisolid	Liquid
1st state of swallowing sounds	0.24±0.11	0.23±0.07
2nd state of swallowing sounds	0.20±0.12	0.18±0.10
3rd state of swallowing sounds	0.47±0.18	0.48±0.19
Duration of muscle contraction	0.48±0.09	0.58±0.08
Time difference between the onset of muscle contraction and the opening of cricopharynx	0.25±0.12	0.19±0.09

References

[1] J.A. Logemann, *Evaluation and Treatment of Swallowing Disorders*, Austin, TX, PROED, 1998.
[2] G. Rempel, Z. Moussavi, "The effect of viscosity on the breath-swallow pattern of young people with cerebral palsy," *Dysphagia*, vol. 20, no. 2, pp. 108-112, 2005.
[3] S. Hamlet, D. Penney, and J. Formolo, "Stethoscope acoustics and cervical auscultation of swallowing," *Dysphagia*, vol. 9, pp. 63-68, 1994
[4] W.G. Selley, F.C. Flack, and R.E. Ellis, "The Exeter dysphagia assessment technique," *Dysphagia*, vol. 4, pp. 227-235, 1990.
[5] L. Lazareck and Z. Moussavi, "Classification of Normal and Dysphagic Swallows by Acoustical Means," *IEEE Trans. Biomedical Engineering*, vol. 51, pp. 2103-2112, Dec. 2004
[6] M. Aboofazeli and Z. Moussavi, "Analysis and Classification of Swallowing Sounds Using Reconstructed Phase Space Features," in *Proc. Of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, pp. 421-424, March 2005.
[7] M.S. Klahn, and A.L. Perlman, "Temporal and durational patterns associating respiration and swallowing," *Dysphagia*, vol. 13, pp. 131-138, 1999.
[8] I. H. Gewolb, J. F. Bosma, V. L. Taciak, and F. L. Vice, "Abnormal developmental patterns of suck and swallow rhythms during feeding in preterm infants with bronchopulmonary dysphagia," *Developmental Med. Child Neurol.*, vol. 43, no. 7, pp. 454–459, July 2001.
[9] F. L. Vice, J. M. Heinz, G. Giuriati, M. Hood, and J.F. Bosma, "Cervical auscultation of suckle feeding in newborn infants," *Developmental Medicine and Child Neurology*, vol. 32, pp.760-768, 1990.
[10] L. Lazareck and Z. Moussavi, "Adaptive swallowing sound segmentation by variance dimension," *Proc European Med Biol Eng Conf (EMBES)*, Oct. 2002.
[11] Z. Moussavi, "Assessment of swallowing sounds' stages with hidden Markov model," in *Proc. of 27th Annual International Conference of the IEEE Eng. in Med. and Biol. Society (EMBS)*, September 2005
[12] Z.K. Moussavi, M.T. Leopando, H. Pasterkamp, and G. Rempel, "Computerized acoustical respiratory phase detection without airflow measurement," *Journal of Medical & Biolog. Eng. & Comp.*, Vol. 38, no. 2, pp. 198-203, March 2000
[13] L. R. Rabiner, "A tutorial on hidden Markov models and selected application in speech recognition," Proc. IEEE, vol. 77, Feb. 1989.
[14] B. Gold and N. Morgan, *Speech and audio signal processing*, John Wiley & Sons, New York, NY, 1999.
[15] M.J. Katz, "Fractals and the analysis of the waveforms," comput. Biol. Med., vol.18, no.3, pp.145-156, 1988.
[16] R. Esteller, G. Vachetsevanos, J. Echauz, and Brian Litt, "A comparison of waveform fractal dimension algorithms," *IEEE Trans. Circuits and systems*, vol. 48, no.2, pp. 177-183, 2001.
[17] A.J. Viterbi, "Error bounds for deconvolutional codes and an asymptotically optimal decoding algorithm," IEEE Trans Info Theory, vol. IT-13, pp. 260-269, Apr. 1967.