Automated Extraction of Swallowing Sounds Using a Wavelet-Based Filter

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

Abstract- **This paper presents an automated and objective method for extraction of swallowing sounds in a record of the tracheal breath and swallowing sounds. The proposed method takes advantage of the fact that swallowing sounds have more non-stationarity comparing with breath sounds and have large components in many wavelet scales whereas wavelet transform coefficients of breath sounds in higher wavelet scales are small. Therefore, a wavelet transform based filter was utilized in which a multiresolution decomposition-reconstruction process filters the signal. Swallowing sounds are detected in the filtered signal. The proposed method was applied to the tracheal sound recordings of 15 healthy and 11 dysphagic subjects. The results were validated manually by visual inspection using airflow measurement and spectrogram of the sounds and auditory means. Experimental results prove that the proposed method is more accurate, efficient, and objective than the methods proposed previously. Swallowing sound detection may be employed in a system for automated swallowing assessment and diagnosis of swallowing disorders (dysphagia) by acoustical means.**

Keywords- **Swallowing sounds, breath sounds, wavelet transform**

I. INTRODUCTION

Swallowing mechanism is a complicated process in which a number of events occur in a short period of time, i.e. less than 0.5 s. Coordination of swallowing mechanism events in relation to respiration is essential in a normal swallow. Lack of coordination may result in food or liquid entry into airway, which is called aspiration. A controlled and coordinated swallow does not consistently occur in patients with a variety of congenital abnormalities, structural damage, and/or medical conditions such as cerebral palsy, brain-stem stroke, head/neck and spinal cord injuries and structural damage after treatment for oral and pharyngeal cancer [1]. Swallowing disorder (dysphagia) can occur in all age groups from newborn to the elderly [1,2].

In recent years, acoustical analysis of swallowing mechanism has received considerable attention [2-7], in which respiratory and swallowing sounds are recorded by microphones and/or accelerometers and analyzed by digital signal processing techniques.

As the first step in the swallowing sound analysis, it must be extracted from the recorded signal. Currently, this is done in a time-consuming manual manner. Recently, there have been some attempts to develop automated algorithms for swallowing sound detection in the tracheal sound signal [3,4]. Once validated, such an automated method can be used as part of a system for swallowing assessment by acoustical means.

In [3], three tracheal sound signal features (autoregressive model coefficients, RMS of the signal in the time domain, and the average power of the signal over different frequency bands) were used to classify a tracheal sound segment as either swallowing or breath sound. That method yielded an average error of 21%.

In our previous work [4], an algorithm based on multilayer feed forward neural networks was proposed for decomposition of tracheal sounds into swallowing and respiratory sound segments. RMS of the signal in the time domain, the average power of the signal over 150-450 Hz and the waveform fractal dimension were applied to a neural network as input features. The algorithm was able to detect 91.7% of 253 swallows correctly.

In order to increase the accuracy of swallowing sound extraction, this study uses a discrete wavelet transform (DWT) based filter for automated separation of swallowing sounds from breath sounds. This filter has been used successfully for separation of discontinuous adventitious sounds from vesicular sounds [8], suppressing noise in partial discharge on-line monitoring [9], and enhancement of bowel sounds [10].

Swallowing sounds consist of discrete nonstationary click sounds whereas breath sounds are relatively stationary especially when the air flow plateaus. In the proposed method, we use wavelet transform to decompose the tracheal sound into multiscale details. As swallowing sounds have more non-stationarity compared with breath sounds, they have large components in many wavelet scales whereas wavelet coefficients of breath sounds in higher wavelet scales are small. In the proposed method, swallowing sounds are separated from breath sounds through a multiresolution decomposition-reconstruction process which is explained in section II.

II. METHODOLOGY

A. Data

Data were adopted from a previous study [2]. Tracheal sound recordings of two groups of subjects were used in this
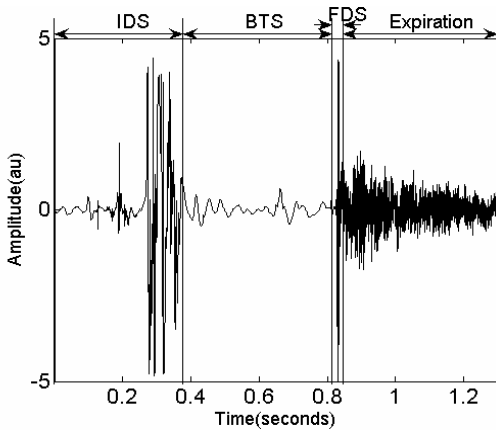
Fig. 1. A typical swallowing sound signal. Legend: 'au' arbitrary units for the normalized amplitude.

study. The first group consisted of 12 healthy children (3-16 years) and three healthy adults (ages 35, 38, and 54 years). All subjects of the first group were in good health without any history of swallowing disorder, eating or nutrition problems, or lower respiratory tract infection. The second group consisted of 11 young adult cerebral palsy patients (ages 16-25 years) at risk of dysphagia.

During the test, the participants were fed three textures: pre-packaged pudding (semisolid texture), diluted pudding (thick liquid texture) and fruit juice (thin liquid) in bolus size of 5 ml. Tracheal breath and swallowing sounds were recorded by a Siemens accelerometer (EMT25C) placed over suprasternal notch. The signals were bandpass filtered (50-2500 Hz) and digitized at a 10240 Hz sampling rate.

Airflow signal was also recorded simultaneously by nasal canulae attached to a pneumotachograph (Fleisch No 3) and Validyne differential pressure transducer. The air flow was used along with manual visual and auditory means for validation of the proposed method.

B. Multiresolution Analysis of Tracheal Sound

Swallowing sounds may be divided into three distinct segments: Initial Discrete Sounds (IDS), Bolus Transit Sounds (BTS) and Final Discrete Sounds (FDS). IDS and FDS precede and follow the BTS, respectively and are characterized by their time domain features such as relatively short duration. IDS may be repeated in a highly similar form [5]. 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.

In this study, multiresolution analysis of tracheal sound is the basis of wavelet transform based filter. The signal is decomposed into band limited components, which can be reassembled to reconstruct the original signal without error. Since the bandwidth of the resulting subbands is smaller than that of the original signal, the subbands can be downsampled without any loss of information. Reconstruction of the original signal is accomplished by upsampling, filtering and summing the individual subbands. The decomposition process can be iterated such that one
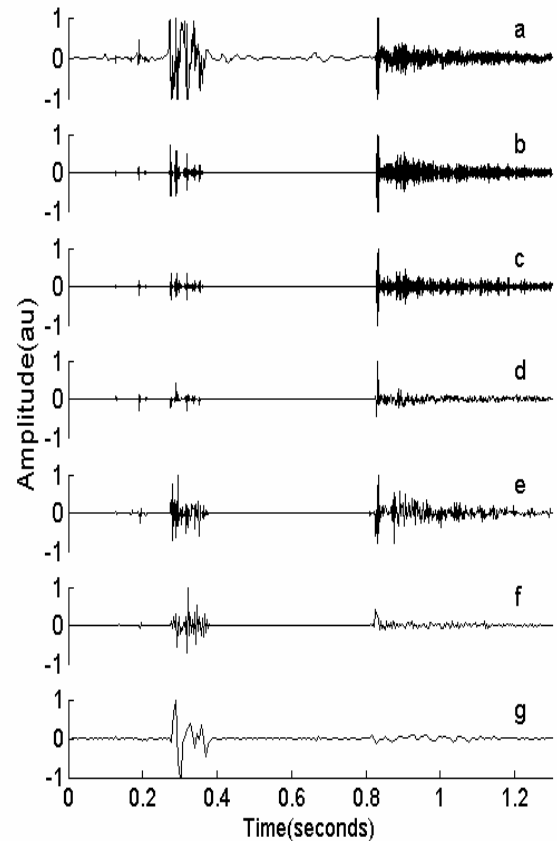

Fig. 2. A typical tracheal sound signal decomposed by multiresolution analysis. (a) Original signal (b-g) the wavelet coefficient of the tracheal signal at levels 1-6, respectively.

signal is broken into many lower resolution components. A complete description of multiresolution analysis can be found in [11].

Multiresolution decomposition of tracheal sound in 6 levels using Daubechies-5 wavelet is shown in Fig. 2. As the swallowing sound signal includes discrete nonstationary click sounds, it has large components in different wavelet scales whereas the breath sound is relatively stationary and its wavelet coefficients die out quickly in higher wavelet scales.

C. The DWT Based Filter

This study uses a similar DWT based filter as that was initially proposed in [8]. A schematic description of the filter is shown in Fig. 3. In this filtering method, iterative sequences of multiresolution decomposition (MRD) and multiresolution reconstruction (MRR) are employed.

In this study, swallowing sounds are detected in normalized N-sample tracheal sound signal denoted by $X(n)$, $n = 1,2,...,N$. During the kth iteration, the wavelet coefficients at levels $j = 1,..,6$ are calculated. The wavelet
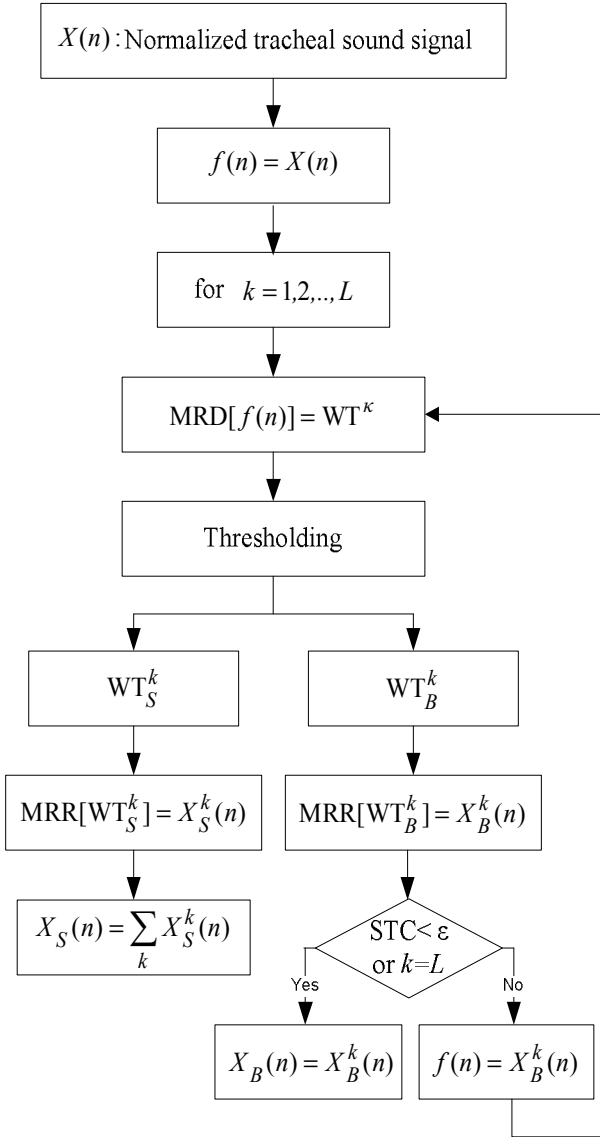
Fig. 3. Schematic description of the wavelet-based filtering method.

coefficients at level j are compared to a threshold, defined as:

$$T_j^k = \sigma_j^k . F_{adj} , \qquad (1)$$

where, σ_j^k is the standard deviation of the wavelet coefficients at level j during the kth iteration, and F_{adj} is a constant coefficient. Comparison to the threshold defined in (1) divides the Wavelet coefficients at each scale into big $(\geq T_j^k)$ and small $(< T_j^k)$ coefficients, WT_S^k and WT_B^k, respectively. Then, MRR is applied to WT_S^k and WT_B^k. The iterative procedure stops after a predefined number of iterations or when:

$$STC = | E\{X_{B,k-1}^2(n)\} - E\{X_{R,k}^2(n)\} | < \varepsilon , \qquad (2)$$

where, $E\{.\}$ denotes the expected value and ε is a very small positive number by which the desired accuracy in the separation procedure is determined. When convergence is achieved or after iteration L, $X_S(n)$ is determined by combining the coherent parts derived at iterations $k = 1,...,L$, i.e.,

$$X_S(n) = \sum_{k=1}^{L} X_{S,k}(n) . \qquad (3)$$

Although in [8]-[10], the filter described above has been reported to be successful in the separation of discontinuous adventitious sounds from vesicular sounds, suppressing noise in partial discharge on-line monitoring, and enhancement of bowel sounds, $X_S(n)$ and $X_B(n)$ obtained by this filtering method give only an approximation of swallowing sounds and breath sounds. In many cases, $X_B(n)$ contains some parts of swallowing sounds. Therefore, in order to detect boundaries of the swallowing and breath sounds more accurately, the RMS of 128 sample windows of $X_S(n)$ and $X_B(n)$ were compared. When for a segment

$$RMS(X_S(n)) > RMS(X_B(n)), \qquad (4)$$

the segment was considered to be part of the swallowing sound signal.

III. RESULTS AND DISCUSSION

The filtering method discussed in previous section was used to separate swallowing segments from breath sounds followed by the extra criterion in Eq. (4). Figure 4.a illustrates 5-second period of a typical tracheal sound signal which was processed by the proposed method. Fig. 4.b shows the refined swallowing sound signal $X_S(n)$. As it can be seen, the refined signal does not contain most parts of breath sounds. In Fig. 4.c, refined breath sound is shown. As it can be seen, some parts of swallowing sound are still in the signal. Therefore, the filtering method alone is not able to separate the two swallowing and breath sounds with accurate boundaries.

Applying the criterion in Eq. (4), the boundaries of the swallowing sounds were detected and the results are shown in Fig. 5. As it can be seen, during the detected swallowing period, the recorded airflow is zero. The onset of the swallow coincides with the end of the preceding respiratory phase and the end of the swallow is the onset of the respiratory phase following the swallow. The swallow shown in this figure is followed by a forceful expiration. The method was able to detect the boundaries of the swallows correctly. The results were validated manually by visual
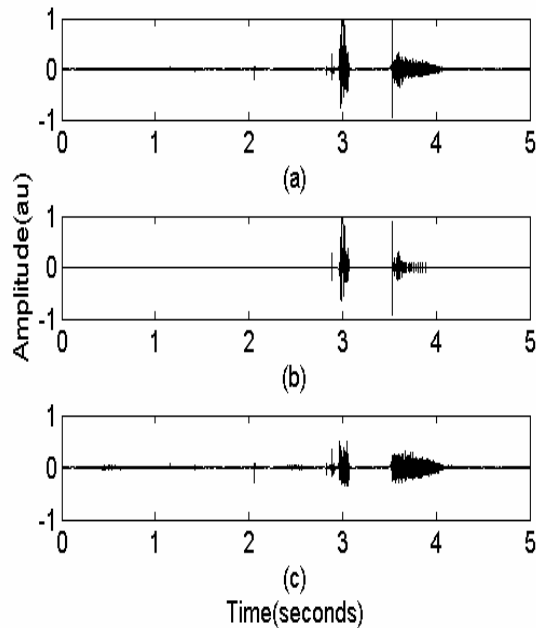
Fig. 4. A typical tracheal sound signal decomposed by wavelet-based filter: (a) original signal (b) extracted swallowing sound signal (c) extracted breath sound signal using wavelet-based filter.

inspection of airflow measurement and spectrogram of the sounds as well as by auditory means.

The proposed method identified 93% of the swallowing sound segments correctly. It yielded an average error of 7% which is significantly lower than 21% reported in [3]. In addition, false swallow detection is 3% which is much lower than 9.5% reported in [4].

IV. CONCLUSION

This paper proposes an automated and objective method to identify swallowing sounds from breath sounds in healthy and dysphagic subjects. Evaluation of the performance through experimental results on data of 15 healthy subjects and 11 patients at risk of dysphagia are encouraging to establish the method as an automated and objective method with low error (7% overall error and 3% false positive error). The outcomes of this study may be utilized in evaluation of the performance of oropharyngeal swallowing mechanism based on its acoustical signature.

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] L. J. Lazareck, and Z.K. Moussavi, "Automated algorithm for swallowing sound detection," in *Proc. Canadian Med. and Biol. Eng. Conf.*, Oct. 2002.
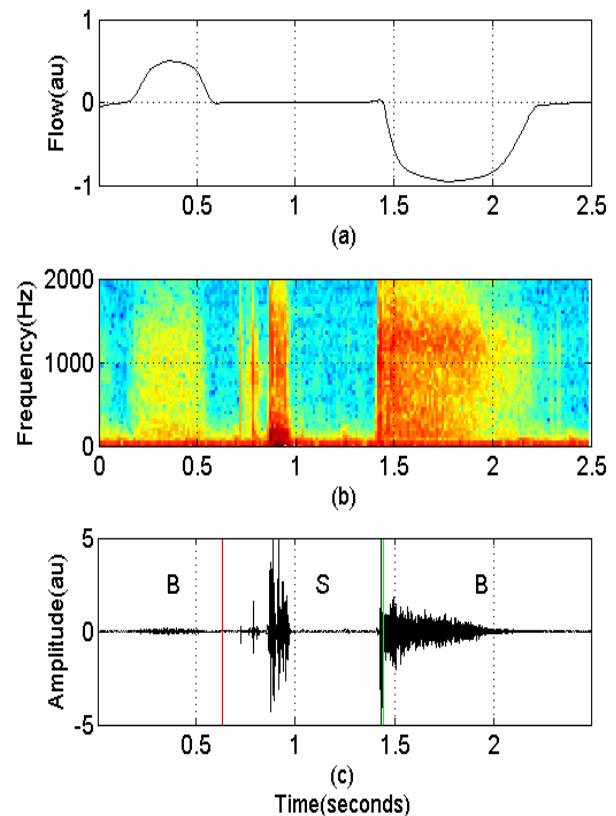
Fig. 5. (a) Airflow signal, (b) spectrogram and (c) tracheal sound waveform as a function of time. Legend: B: Breath Sound, S: Swallowing sound, au: arbitrary units for normalized amplitude, red vertical lines show onset of swallows and green vertical lines show the end of swallowing segments.

[4] M. Aboofazeli, Z. Moussavi, "Automated classification of swallowing and breath sounds," in *Proc. 26th Conf. IEEE Eng. In Med. And Biol. Society (EMBS)*, pp. 3816-3819, Sept. 2004

[5] 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.

[6] M.S. Klahn, and A.L. Perlman, "Temporal and durational patterns associating respiration and swallowing," *Dysphagia*, vol. 13,pp. 131-138, 1999.

[7] W.G. Selley, F.C. Flack, and R.E. Ellis, "The Exeter dysphagia assessment technique," *Dysphagia*, vol. 4, pp. 227-235, 1990.

[8] L. J. Hadjileontiadis, and S.V. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 1269-1283, December 1997

[9] H. Mingyou, H. Xie, T.B. Tiong, and X. Wu, "Study on a spatially selective noise filtration technique for suppressing noises in partial discharge on-line monitoring," in *proceeding of the 6th Conference on properties and Application of Dielectric Materials*, pp. 689-692, June 2000

[10] L. J. Hadjileontiadis, C.N. Liastos, C.C. Mavrogiannis, T.A. Rokkas, and S.M. Panas, "Enhancement of bowel sounds by wavelet-based filter," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 876-886, July 2000

[11] R. Gonzalez, and R. Woods, *Digital Image processing*, 2nd Edition, Prentice Hall Inc., 2002