Classification of penetration-aspiration versus healthy swallows using dual-axis swallowing accelerometry signals in dysphagic subjects

Ervin Sejdić, *Member, IEEE,* Catriona M. Steele and Tom Chau, *Senior Member, IEEE*

***Abstract*—Swallowing accelerometry is a promising non-invasive approach for the detection of swallowing difficulties. In this paper, we propose an approach for classification of swallowing acceleroemtry recordings containing either healthy swallows or penetration-aspiration (entry of material into the airway) in dysphagic patients. The proposed algorithm is based on the wavelet packet decomposition of swallowing accelerometry signals in combination with linear discriminant analysis as a feature reduction method and Bayes classification. The proposed algorithm was tested using swallowing accelerometry signals collected from 40 patients during the regularly scheduled videoflouroscopy exam. The participants were instructed to swallow several five milliliter sips of thin liquid barium in a head neutral position. The results of our numerical analysis showed that the proposed algorithm can differentiate healthy swallows from aspiration swallows with an accuracy greater than 90%. These results position swallowing accelerometry as a valid approach for the detection of swallowing difficulties, particularly penetration-aspiration in patients suspected of dysphagia.**

***Index Terms*—Dual-axis swallowing accelerometry signals, wavelet transformation, linear discriminant analysis, Bayes classification.**

I. Introduction

SWALLOWING is a well-defined process of transporting food or liquid from the mouth to the stomach [1]. Patients suffering from dysphagia (swallowing difficulty), usually deviate from this well-defined pattern of healthy swallowing. Dysphagia is a common problem encountered in the rehabilitation of stroke patients, head injured patients, and others with paralyzing neurological diseases [2]. Dysphagic patients are likely to choke or aspirate. Aspiration is defined as the entry of material into the airway below the true vocal folds [1]. Penetration is a related phenomenon in which material enters the space immediately above the true vocal folds (the supraglottic space) but is not observed to fall below the vocal folds during assessment. Dysphagia may have dire consequences including malnutrition and dehydration [3], [4], degradation in psychosocial well-being [5], [6], aspiration pneumonia [7], and even death [8].

The gold standard for the detection of penetration-aspiration is the videofluoroscopic swallowing study (VFSS) [1], [9]. However, VFSS requires expensive X-ray equipment as well as expertise from speech-language pathologists and radiologists, which limits the number of institutions that can offer VFSS [10]. As a result, VFSS has been associated with long waiting lists [11]. Also, day-to-day monitoring of dysphagia is crucial due to the fact that severity of dysphagia can change over time. Certainly, VFSS is not suitable for such day-to-day monitoring.

Swallowing accelerometry is a promising non-invasive tool for the assessment of swallowing difficulties, including penetration-aspiration, which refers to an approach employing an accelerometer as a sensor during cervical auscultation. Swallowing accelerometry has been used to detect dysphagia in several studies, which have described a shared pattern among healthy swallow signals, and verified that this pattern is either absent or delayed in dysphagic swallow signals [12]-[19]. Several studies have pointed to movement of the hyoid and larynx as the most likely sources of the vibratory signals detected in swallowing accelerometry, based on the close timing correspondence between spikes in the accelerometry signal and movement of these structures observed using videofluoroscopy. Additionally correlations have been reported between the extent of laryngeal elevation and the magnitude of the accelerometry signal [16], [20]. Proper hyolaryngeal movement with precise timing during bolus transit is vital for airway protection [1], [20]. Since the motion of the hyolaryngeal structure during swallowing occurs in both anterior-posterior (A-P) and superior-inferior (S-I) directions, the employment of dual-axis accelerometry seems well motivated. Since correlation has been reported between the extent of laryngeal elevation and the magnitude of the A-P swallowing accelerometry signal [16], it is hypothesized that vibrations in the S-I axis also capture useful information about laryngeal elevation. From a physiological stand point, the S-I axis appears to be as worthy of investigation as the A-P axis because the maximum excursion of the the hyolaryngeal structure during swallowing is of similar magnitude in both the anterior and superior directions [21], [22]. This has been confirmed in recent papers [20], [23].

Several recent contributions have proposed different classification approaches that differentiate between healthy swallows and swallows characterized by penetration-aspiration (e.g.,

Ervin Sejdić is with the Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA. E-mail: esejdic@ieee.org. Ervin Sejdić is the corresponding author.

Catriona M. Steele is with the Toronto Rehabilitation Institute and the Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, Canada. E-mail: catriona.steele@uhn.ca.

Tom Chau is with Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital and the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada. E-mail: tom.chau@utoronto.ca

Manuscript received September 24, 2012.

[19], [24], [25]). However, the main issue was that those approaches did not provide clinically sufficient accuracy. To alleviate this issue, we propose an algorithm based on the wavelet packed decomposition and Bayes classification. In this paper, the algorithm differentiates healthy swallows from those involving penetration-aspiration with an accuracy of 90%. The wavelet decomposition was chosen to capture nonstationarities present in swallowing accelerometry signals (e.g., [18]), while recognizing that other time-frequency approaches could possibly lead similar results. However, such investigations are beyond the scope of this manuscript.

The paper is organized as follows: Section II outlines the experimental approach and data acquisition procedures. In Section III, feature extraction and classification is considered in detail and a simple algorithm for differentiation between healthy swallows and penetration-aspiration swallows based on dual-axis swallowing accelerometry signals is described, while Section IV outlines the data processing steps. Section V presents the results of classification, and Section VI overviews the implications of these results. Conclusions are provided in Section VII followed by a list of references.

II. Methodology

Data were gathered from 40 consenting patients with dysphagia during routine videofluoroscopic swallowing study (VFSS) at the Toronto Rehabilitation Institute and Toronto East General, Toronto, Ontario, Canada. Four channels were used for data collection: a dual-axis accelerometer, a lapel microphone, a video camera and the videofluoroscopy recording. A dual-axis accelerometer (ADXL322, Analog Devices) was attached to the participant's neck (anterior to the cricoid cartilage) by using double-sided tape. The axes of acceleration were aligned to the anterior-posterior and superior-inferior directions. Signals were also band-pass filtered in hardware with a pass band of 0.1-3000 Hz and were saved for subsequent off-line analysis. A lapel microphone was attached around the patient's neck, so that voice recordings could be taken. A video camera was set up to capture an anterior view of the participant's face and neck, whose identity was concealed with dark glasses. The videofluoroscopy recording was digitally captured by inserting a splitter into the videofluoroscopy cable set-up to allow duplicate capture of the x-ray recording onto a laptop computer for research purposes without interfering with the recording of the x-ray for clinical purposes. A custom LabVIEW program running on a laptop computer was used to acquire acceleration signals, speech signals and images from the x-ray machine. The data were saved on the hard drive of the computer for subsequent off-line analyses.

Data collection proceed following a standardized videofluoroscopy protocol at the two hospitals. The participants were cued to swallow 5ml sips of 100 grams of Liquid Polibar Suspension diluted with water to reach 250 ml, for a 22% w/v barium concentration. The procedure included a minimum of two and maximum of three 5ml sips of thin liquid with the head in neutral position. A fourth thin liquid barium swallowing task (i.e. a sequence of sips from a cup) was requested in some cases at the clinician's discretion. If the clinician detected serious swallowing difficulties upon the first two thin liquid sips, data collection for the research study ended. The videofluoroscopy continued with further tasks as determined appropriate by the attending clinician. The data that were extracted for the purposes of research analysis were restricted to the first 3 thin liquid sips (5 ml) performed in a head neutral position and the first cup drinking task performed with thin liquid.

III. Feature processing and classification

A. Wavelet transform

Wavelet transform is a mathematical technique which decomposes a signal into both time and scale using specific analyzing functions, often called wavelets. The continuous wavelet transform (CWT) of a signal $x(t) \in L^2(\mathbb{R})$ is given by [26]-[30]:

$$CWT(s,\tau) = \frac{1}{\sqrt{s}} \int x(t)\psi^* \left(\frac{t-\tau}{s} \right) dt \qquad (1)$$

where $\psi(t)$ is the analyzing wavelet, and an asterisk denotes the complex conjugate. The parameter τ indicates the translation in time, and the parameter s is the scale parameter. Certain conditions, such as zero mean and admissibility, are required in order for a function to be considered for an analyzing wavelet (e.g., [26]-[29]).

The calculation of the wavelet coefficients $CWT(s,\tau)$ at every possible scale and translation can be computationally demanding and may introduce redundant data. However, by limiting our analysis to specific scale and translation parameters, a signal can be described as

$$x(t) = \sum_k \sum_j a_{j,k} \psi_{j,k}(t) \qquad (2)$$

where $j, k \in \mathbb{Z}$. In particular, if it is assumed that $s = 2^j$, then

$$\psi_{j,k}(t) = 2^{j/2} \psi \left(2^{j/2} t - k \right) \qquad (3)$$

The set of expansion coefficients

$$a_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle \qquad (4)$$

is called the discrete wavelet transform (DWT) of $x(t)$ [29]. In comparison to Fourier series which map a one-dimensional function into a one-dimensional sequence of coefficients, the wavelet expansion maps it into a two-dimensional matrix of coefficients.

In practice, the DWT can be implemented using Mallat's algorithm [29]. To understand such an implementation, it should be pointed out that wavelets can be expressed in terms of so called scaling functions, $\varphi(t)$, as following [29]:

$$\psi(t) = \sqrt{2} \sum_n h_1(n) \varphi(2t-n) \quad n \in \mathbb{Z} \qquad (5)$$

where $h_1(n)$ represents a set of coefficients, and the scaling functions are defined as a weighted sum of shifted versions of themselves [29]:

$$\varphi(t) = \sqrt{2} \sum_n h(n) \varphi(2t-n) \quad n \in \mathbb{Z} \qquad (6)$$

with $h(n)$ representing a set of so called scaling coefficients. Then, using these wavelets and scaling functions, (2) can be expressed as

$$x(t) = \sum_k c_{j_0}(k)\varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{\infty} d_j(k)\psi_{j,k}(t) \quad (7)$$

where $j_0 \in \mathbb{Z}$, and the expansion coefficients can be calculated as

$$c_j(k) = \langle x(t), \varphi_{j,k}(t)\rangle \quad (8)$$

$$d_j(k) = \langle x(t), \psi_{j,k}(t)\rangle \quad (9)$$

The advantage of expanding $x(t)$ in terms of wavelets and scaling functions is that in many applications we can directly deal with $h_1(n)$ and $h(n)$ without using the wavelets or scaling functions. It is a well established fact in the wavelet literature that

$$c_j(k) = \sum_m h(m-2k)c_{j+1}(m) \quad (10)$$

and

$$d_j(k) = \sum_m h_1(m-2k)c_{j+1}(m) \quad (11)$$

and these equations can be implemented using a filter bank. Therefore, the two sets of coefficients are only needed to compute the wavelet transform of a signal.

The wavelet packet generalizes the structure of the wavelet transform to a full decomposition. At each level, DWT further decomposes the approximation coefficients to yield approximation and detail information. However, the wavelet packet analysis decomposes approximation and detail coefficients from each level into another set of approximation and detail coefficients. Hence, the wavelet packets achieve a better time-frequency localization of signal components.

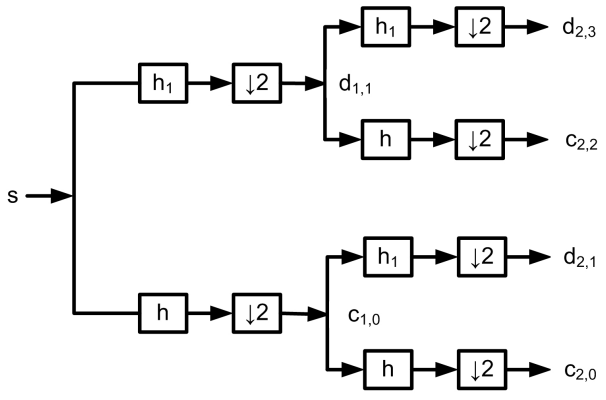

Fig. 1. A generic wavelet packet based decomposition.

1) Wavelet packet coefficients as deature: The wavelet expansion provides a time-frequency localization of a signal, meaning that the coefficients obtained using (10) and (11) provide time-frequency dependent information. The next step is to somehow manipulate these coefficients in order to extract this information more consicely. Therefore, assume that we have a vector, w_x of length $N \in \mathbb{N}$, containing the L^{th} level wavelet expansion coefficients of $x(t)$:

$$w_x = \left[c_{L,0}\ d_{L,1} \dots c_{L,2^L-2} d_{L,2^L-1}\right] \quad (12)$$

with each coefficient being a vector of a varying length less than N. In essence, w_x is a vector containing 2^L vectors. Then we can define a mathematical operator $\Phi : \mathbb{C}^N \to \mathbb{C}^{2^L}$ to modify the existing vector of coefficients, such that:

$$W_x = \Phi\left[w_x\right]. \quad (13)$$

where the operator Φ acts on each individual coefficient, i.e., $W_x(1) = \Phi\left[w_x(1)\right] = \Phi\left[c_{L,0}\right]$, $W_x(2) = \Phi\left[w_x(2)\right] = \Phi\left[d_{L,1}\right] \dots W_x(2^L) = \Phi\left[w_x(2^L)\right] = \Phi\left[d_{L,2^L-1}\right]$. It should be noted that $L \ll N$. In particular, several such operators are considered in this paper:

- energy operator [31], [32]:

$$\Phi[w_x(j)] = \|w_x(j)\|_2^2 \quad \text{for } j = 1, 2, ..., 2^L \quad (14)$$

- log-energy operator:

$$\Phi[w_x(j)] = \log\|w_x(j)\|_2^2 \quad \text{for } j = 1, 2, ..., 2^L \quad (15)$$

- entropy operator [33]:

$$\Phi[w_x(j)] = \|w_x(j)\|_2^2 \log\|w_x(j)\|_2^2 \quad \text{for } j = 1, 2, ..., 2^L \quad (16)$$

B. Dimension reduction

As demonstrated in the previous subsection, by introducing various mathematical operators one can reduce dimensionality of the problem space by projecting wavelet coefficients into a subspace. Nevertheless, the subspace obtained by the mathematical transformation of the wavelet coefficients is not necessarily the smallest subspace which contains discriminant information and redundant information may still be present.

Linear discriminant analysis (LDA) projects samples to a lower dimensional space using a linear transformation [34]. In other words, for a number of independent features describing the data, LDA creates a linear combination of these features which yields the largest mean differences between the desired classes by utilization of within-class (Λ_w), between-class (Λ_b), and mixture scatter matrices [34], [35].

These scatter matrices can be used for a linear transformation of an $(L+1)$-dimensional feature vector $\mathbf{W}$ to an k-dimensional $\overline{\mathbf{W}}$ expressed by

$$\overline{\mathbf{W}} = \Xi^T \mathbf{W} \quad (17)$$

where Ξ is an $(L+1) \times k$ rectangular matrix and the column vectors are linearly independent with $(L+1) > k$. By transforming the feature vector W, the scatter matrices can be expressed in the transformed space as $\overline{\Lambda_b} = \Xi^T \Lambda_b \Xi$ and $\overline{\Lambda_w} = \Xi^T \Lambda_w \Xi$. Therefore, by searching for a subspace where the ratio of Λ_b and Λ_w is maximized the problem could be formulated to find a transformation $\overline{\Xi}$ such that

$$\overline{\Xi} = \arg\max_{\Xi} \frac{\Xi^T \Lambda_b \Xi}{\Xi^T \Lambda_w \Xi} \quad (18)$$

Finding a rectangular matrix Ξ that maximizes the above equation is not easy problem [36]. However, it has been shown that columns of an optimal Ξ correspond to the largest eigenvectors of the following eigenvalue problem [37]:

$$\Lambda_b \xi_i = \lambda_i \Lambda_w \xi_i \quad (19)$$

where ξ_i are generalized eigenvectors and λ_i are corresponding eigenvalues.

C. Feature classification

Bayes' decision theory, a fundamental statistical approach to the problem of pattern classification, is based on the assumption that the decision problem is posed in probabilistic terms with all of the relevant probability values being known [36]. For example, given a feature vector $\overline{W}$, one way of establishing the means for classification is through the posterior probability $p\left(\omega_q \left| \overline{W} \right.\right)$, where ω_q represents a q^{th} state of nature (i.e. a class). Bayes formula allows us to calculate such a probability given the prior probability $p\left(\omega_q\right)$ and the conditional density $p(\overline{\mathbf{W}} \left| \omega_q \right.)$ for different categories [38]:

$$p\left(\omega_q \left| \overline{\mathbf{W}} \right.\right) = \frac{p(\overline{\mathbf{W}} \left| \omega_q \right.) p\left(\omega_q\right)}{p(\overline{\mathbf{W}})}. \tag{20}$$

where

$$p(\overline{\mathbf{W}}) = \sum_{q=1}^{Q} p(\overline{\mathbf{W}} \left| \omega_q \right.) p\left(\omega_q\right) \tag{21}$$

The advantage of using Bayes approach is that it minimizes the average probability of error [36]. Furthermore, using these posterior probabilities for different classes, one can form discriminant functions

$$g_q\left(\overline{\mathbf{W}}\right) = \ln p(\overline{\mathbf{W}} \left| \omega_q \right.) + \ln p\left(\omega_q\right) \tag{22}$$

and if the conditional densities $p(\overline{\mathbf{W}} \left| \omega_i \right.)$ are multivariate normal Gaussian densities then (22) can be expressed as [36]:

$$\begin{aligned} g_q\left(\overline{\mathbf{W}}\right) = &-\frac{1}{2}\left(\overline{\mathbf{W}} - \mu_{\overline{\mathbf{W}}}\right)^T \Sigma_q^{-1} \left(\overline{\mathbf{W}} - \mu_{\overline{\mathbf{W}}}\right) \\ &-\frac{q}{2}\ln 2\pi - \frac{1}{2}\ln\left|\Sigma_q\right| + \ln p\left(\omega_q\right) \end{aligned} \tag{23}$$

where

$$\mu_{\overline{W}} = E\left\{\overline{W}\right\} \tag{24}$$

and

$$\Sigma_q = E\left\{\left(\overline{\mathbf{W}} - \mu_{\overline{W}}\right)^T \left(\overline{\mathbf{W}} - \mu_{\overline{W}}\right)\right\}. \tag{25}$$

Eqn. (23) can be simplified based on the structure of Σ_q and for more details one should refer to [36].

Then, if we use the Mahalanobis distance $d_i(x, \mu_i)$ defined as

$$d_i(x, \mu_i) = \sqrt{(x - \mu_i)^T \Sigma_i \left(x - \mu_i\right)} \tag{26}$$

for $Q = 2$ considered in this paper, a single discriminant function can be used

$$\begin{aligned} g(\overline{\mathbf{W}}) = -\frac{1}{2}\Big[&\left(\overline{\mathbf{W}} - \mu_1\right)^T \Sigma_2^{-1} \left(\overline{\mathbf{W}} - \mu_1\right) \\ &- \left(\overline{\mathbf{W}} - \mu_2\right)^T \Sigma_2^{-1} \left(\overline{\mathbf{W}} - \mu_2\right)\Big] - \frac{1}{2}\ln\frac{|\Sigma_1|}{|\Sigma_2|} + \ln\frac{p\left(\omega_1\right)}{p\left(\omega_2\right)} \end{aligned} \tag{27}$$

which simplifies the decision process to

$$\text{Decide } \overline{\mathbf{W}} \in \omega_1 \text{ if } G(\overline{\mathbf{W}}) > 0 \tag{28}$$

$$\text{Decide } \overline{\mathbf{W}} \in \omega_2 \text{ if } G(\overline{\mathbf{W}}) < 0 \tag{29}$$

where

$$G(\overline{\mathbf{W}}) = g(\overline{\mathbf{W}}) - \left[\ln\left(p\left(\omega_1\right)/p\left(\omega_2\right)\right) - 0.5\ln\left(\left|\Sigma_1\right|/\left|\Sigma_2\right|\right)\right]. \tag{30}$$

From the previous two equations it is clear that a very simple classifier can be formed for a two-class case. Furthermore, if the feature space is reduced to a very low dimension, such a classifier becomes computationally very efficient.

D. Dual-axis swallowing accelerometry signals classification process

The previous sections introduced a general framework used for classification of dual-axis swallowing accelerometry signals. Here, we describe a step-by-step procedure to use such a framework in this particular application. The classification process for dual-axis swallowing accelerometry signals based on the afore mentioned approaches consists of the following steps:

1) Decompose signals from A-P and S-I directions using L^{th} level discrete wavelet decomposition in order to obtain sets of wavelet coefficients as follows:

$$\mathbf{w}_{AP} = \left[c_{AP_{1,0}}\ d_{AP_{1,1}} \ldots d_{AP_{2^L-1}}\right] \tag{31}$$

$$\mathbf{w}_{SI} = \left[c_{SI_{1,0}}\ d_{SI_{1,1}} \ldots d_{SI_{2^L-1}}\right] \tag{32}$$

where $\mathbf{w}_{AP}$ and $\mathbf{w}_{SI}$ are the wavelet coefficients in A-P and S-I directions, respectively. It should be mentioned that we did not use $c_{L,0}$ in either direction in order alleviate the effects of any very-low frequency components associated with other physiological phenomena.

2) Convert the obtained wavelet coefficients to desired features according to using mathematical operators Φ:

$$\mathbf{W} = \left[\Phi_{AP}[\mathbf{w}_{AP}] \quad \Phi_{SI}[\mathbf{w}_{SI}]\right] \tag{33}$$

3) By using LDA, reduce the feature space to a smaller subspace with transformation matrices $\overline{\Xi}$ found according to (18):

$$\overline{\mathbf{W}} = \overline{\Xi}^T \mathbf{W} \tag{34}$$

4) Using Bayes' classifer, that is, using the discriminant function as shown in (27), form a two-element decision vector θ, whose elements are assigned as following:

$$\theta = \begin{cases} 1 & \text{if } G\left(\overline{\mathbf{W}}\right) > 0 \\ 0 & \text{if } G\left(\overline{\mathbf{W}}\right) < 0 \end{cases} \tag{35}$$

If θ is true, then the swallow represents aspiration/penetration, otherwise assume it is assumed to be a healthy swallow.

5) Repeat steps 1-4 for each unknown swallow.

As it can be seen, the algorithm applies wavelet decomposition on signals from A-P and S-I directions for each unknown swallow. Then using mathematical operators, the wavelet coefficients are transformed into a smaller subspace as described in Section III-B. One should note that the operators do not have to be the same in the A-P and S-I direction as shown in (33). The further feature space reduction is achieved by using LDA as shown in (34). Using the discriminant function shown in (27), which is developed based on the training samples, one would then classify signals as either healthy swallows or as swallows involving penetration-aspiration, based on objective evaluations of the videofluoroscopy data that are used to divide the data set for algorithm training purposes.

IV. Data processing

The data for each participant contained recordings that were several minutes long. For this paper, we were only interested in the actual recorded swallows. Hence, to extract swallows, the videofluoroscopy recordings were spliced into individual swallow clips capturing the interval between the arrival of the bolus head at the mandibular ramus, and the minimum hyoid position following each swallow. Spontaneous clean-up swallows, following the initial swallow of each bolus, were spliced into separate clips, beginning at the lowest hyoid position before each new swallow event. The cropped recordings were then arranged in random order and reviewed by two speech-language pathologists, blinded to patient identity. The 8-point Penetration-Aspiration Scale (e.g., [39]) was used to rate the occurrence of bolus airway invasion. Penetration-aspiration scale ratings were subsequently collapsed to a binary scale (≤ 2 vs. ≥ 3), distinguishing the transient entry of material into the laryngeal vestibule with subsequent clearance from the deeper entry of material without clearance, i.e., penetration-aspiration [39]. These ratings were used as the gold standard against which we examined the accuracy of the proposed algorithm.

Since accelerometry signals were recorded simultaneously with the videofluoroscopy images, we used the swallow timing identified through these images, to segment the accelerometry signals into individual swallows. As the first step, we filtered the signals using filters identified in [40]:

$$\begin{aligned} H_{A-P}(z) = 1 \diagup \big(&1 - 0.8850z^{-1} + 0.2983z^{-2} - 0.0445z^{-3} \\ &- 0.0018z^{-4} - 0.0095z^{-5} + 0.0205z^{-6} \\ &-0.0220z^{-7} + 0.0156z^{-8} - 0.0071z^{-9}\big) \end{aligned} \tag{36}$$

$$H_{S-I}(z) = \frac{1}{1 - 0.8798z^{-1} + 0.2939z^{-2} - 0.0461z^{-3}} \tag{37}$$

Using the identified transfer functions, an inverse filtering approach (e.g., [41]) was implemented to annul the effects of the data collection system. Second, any motion artifacts were removed using the algorithm based on splines [42]. Specifically, to remove signal components associated with motion artifacts, $\zeta(n)$, we assumed $M \ll N$ knots, where N is the length of the signal, and determined a least square approximation of the expansion coefficients using

$$c_{app}(k) = \left(\left[b_m^p * b_m^p\right]\right)^{-1} * \left[b_m^p * s\right]_{\uparrow m}(k) \tag{38}$$

where $s(n)$ was a filtered swallowing vibration signal in either the A-P or S-I directions and $b_m^p(n)$ represented the discrete B-splines with upsampling integer m [43]. Next, using the approximated expansion coefficients, we reconstructed $\zeta(n)$ using the B-spline indirect transformations:

$$\widehat{\zeta}(k) = b_m^p * \left[c_{app}\right]_{\uparrow m}(k) \tag{39}$$

Lastly, we removed the low frequency components associated with head movements using $\widehat{\zeta}(n)$:

$$z(n) = s(n) - \widehat{\zeta}(n) \tag{40}$$

As the last preprocessing step, the obtained accelerometry signals, $z(n)$, were denoised using a wavelet-based algorithm [44]. First, the variance of the noise ε was estimated from the median, MED_x, of $N/2$ wavelet coefficients at the finest scale [28], [45]:

$$\widehat{\sigma_\varepsilon} = \frac{MED_x}{0.6745} \tag{41}$$

Based on the estimated noise variance and for each τ selected from a set $0 < \tau \leq \widehat{\sigma_\varepsilon}\sqrt{2\log(N)}$, the upper bound was calculated [46]:

$$\begin{aligned} r_{eub}\left(\widehat{m}(\tau), \sigma^2, \alpha, \beta\right) = &\frac{\sigma_\varepsilon^2\sqrt{2\widehat{m}(\tau)}}{N}\left(\sqrt{2\widehat{m}(\tau)} + \beta\right)^2 \\ &+ \frac{2\alpha\sigma_\varepsilon}{\sqrt{N}}\sqrt{\frac{\alpha^2\sigma_\varepsilon^2}{N} + d_e - \left(1 - \frac{\widehat{m}(\tau)}{N}\right)\frac{\sigma_\varepsilon^2}{2}} \\ &+ \frac{2\alpha^2\sigma_\varepsilon^2}{N} + d_e - \sigma_\varepsilon \end{aligned} \tag{42}$$

where α and β represent the parameters for validation probability ($p_v = Q(\alpha)$) and confidence probability ($p_c = Q(\beta)$), with $Q(\cdot)$ for an argument λ being defined as $Q(\lambda) = \int_{-\lambda}^{+\lambda}\left(1/\sqrt{2\pi}\right)e^{-x^2/2}dx$. $\widehat{m}(\tau)$ denotes the number of bases whose expansion coefficients are greater than τ in some subspace of $\mathcal{B}_N$. We used the the soft thresholding procedure to compute the data error required for the evaluation of the upper bound. Next, the optimal threshold for wavelet denoising was determined as:

$$\tau_{opt} = \mathrm{argmin}_\tau r_{eub}\left(\widehat{m}(\tau), \widehat{\sigma_\varepsilon}^2, \alpha, \beta\right) \tag{43}$$

Lastly, we denoised $z(n)$ using the optimal value of threshold, τ_{opt}, and the soft thresholding procedure.

Using the processed accelerometry signals, we analyzed the effects of various wavelet functions on the accuracy of the proposed scheme. In particular, we examined the Morlet wavelet function and the family of Coiflets wavelets (so-called Coiflets 1, Coiflets 2, Coiflets 3, Coiflets 4 and Coiflets 5). These mother wavelets were chosen as their shapes resemble dual-axis swallowing accelerometry signals shown in Figure 2. More detailed optimization of our classification approach with respect to the choice of the mother wavelet is beyond the scope of the current manuscript. Using these six wavelets functions, we varied the decomposition level between 2 and 9 for A-P and S-I signals. By varying these parameters, classifier accuracy was estimated using a leave-one-out cross-validation [47]. We also evaluated the following performance metrics:

- True positives (TP) - the number of aspiration swallows correctly identified as penetration-aspiration swallows;
- False positives (FP) - the number of healthy swallows incorrectly identified as penetration-aspiration swallows;
- True negatives (TN) - the number of healthy swallows correctly identified as healthy swallows;
- False negatives (FN) - the number of penetration-aspiration swallows incorrectly identified as healthy swallows.

Using these metrics, we calculated the sensitivity ($= TP/(TP+FN)$), the specificity ($= TN/(TN+FP)$), and the accuracy ($= (TP+TN)/(TP+FP+FN+TN)$).

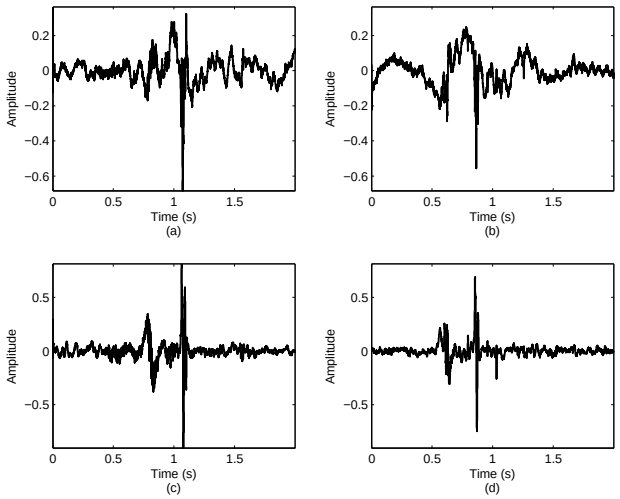

Fig. 2. A sample healthy swallow in A-P and S-I directions is shown in (a) and (c), respectively. A sample penetration-aspiration swallow in A-P and S-I directions is shown in (b) and (d), respectively.

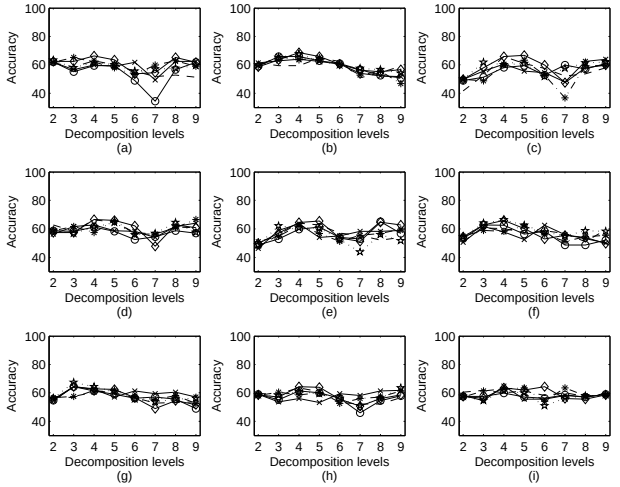

Fig. 3. The accuracy of the proposed scheme when considering the same decomposition levels and the same wavelets functions in both directions.

V. Results

Figure 2 depicts sample healthy and penetration-aspiration swallows. Figure 3 summarizes the results of the first investigation where we examined the accuracy of the proposed classification scheme when we use the same decomposition levels and the same wavelets for both A-P and S-I signals. The x-marked line represents the Meyer wavelet; the dashed line represents the Coiflet 1 wavelet; the dashdot line with stars represents the Coiflet 2 wavelet; the solid line with diamonds represents the Coiflet 3 wavelet; the dotted line with pentagrams represents the Coiflet 4 wavelet; and the solid line with circles represents the Coiflet 5 wavelet. The presented results differ in the features used in the A-P and S-I directions. Figure 3 (a) summarizes the results when the energy operator is used as a feature in both directions. The average accuracy is around 60%, except for the Coiflet 5 wavelet when using 7th decomposition level in both directions. In this case, the accuracy drops to around 30%. Figure 3 (b) depicts when the

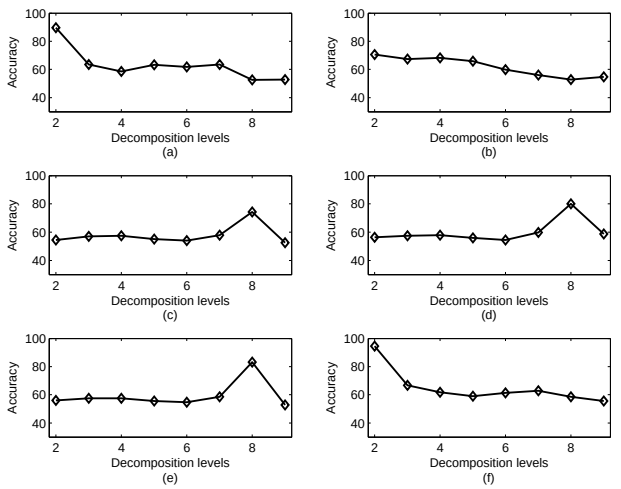

Fig. 4. The accuracy of the proposed scheme when keeping the decomposition level in the A-P direction at a constant value, while varying the decomposition level in the S-I direction.

log-energy operator is used as a feature in both directions. A slight increase in the accuracy is observed when the fourth decomposition level is used in both direction. The results obtained when the entropy operator is used as a feature in both directions are shown in Figure 3 (c). We observed a decrease in the accuracy around the seventh decomposition level as in Figure 3 (a), but generally the accuracy hovers around 60%-70%. The same trends can be observed while considering other combinations of the features: the energy operator in the A-P direction and the log-energy operator in the S-I direction as shown in Figure 3 (d); the energy operator in the A-P direction and the entropy operator in the S-I direction as shown in Figure 3 (e); the log-energy operator in the A-P direction and the energy operator in the S-I direction as shown in Figure 3 (f); the log-energy operator in the A-P direction and the entropy operator in the S-I direction as shown in Figure 3 (g); the entropy operator in the A-P direction and the energy operator in the S-I direction as shown in Figure 3 (h); and the energy operator in the A-P direction and the log-energy operator in the S-I direction as shown in Figure 3 (i). Overall, the classification results were around 60%.

Figure 4 summarizes the results when considering different decomposition levels and different wavelet functions in two anatomical directions. The decomposition level was constant in the A-P direction, while the decomposition level was varied between second and ninth decomposition levels in the S-I direction. Figure 4(a) shows the results when the Coiflet 5 wavelet is used in the A-P direction at the eight decomposition level, while in the S-I direction we used the Coiflet 5 wavelet. The log-energy feature is used in both directions. Using again the log-energy feature in both directions, the results for the Coiflet 1 wavelet in the A-P direction with ninth decomposition level and the Coiflet 3 wavelet in the S-I direction are shown in Figure 4(b). Figure 4(c) shows the results when the Meyer wavelet is used in the A-P direction at the second decomposition level, while in the S-I direction using the Coiflet 1 wavelet. The entropy feature is used

TABLE I
A SUMMARY OF CASES WHEN THE SENSITIVITY AND SPECIFICITY ARE ABOVE 70%.

Wavelets		Features		Dec. level		Accuracy		
A-P	S-I	A-P	S-I	A-P	S-I	SEN	SP	ACC
Coiflet 5	Coiflet 5	log-energy	log-energy	8	2	86.3	91.2	89.7
Meyer	Coiflet 1	entropy	log-energy	2	8	73.8	74.6	74.3
Coiflet 3	Coiflet 1	entropy	log-energy	2	8	85.0	77.9	80.1
Coiflet 5	Coiflet 1	entropy	log-energy	2	8	83.8	82.8	83.1
Coiflet 5	Coiflet 3	log-energy	log-energy	8	2	92.5	95.6	94.6

in the A-P direction, while the log-energy was used as a feature in the S-I direction. Figure 4(d) shows the results when the Coiflet 3 wavelet is used in the A-P direction at the second decomposition level, while in the S-I direction using the Coiflet 1 wavelet. The entropy feature is used in the A-P direction, while the log-energy was used as a feature in the S-I direction. Figure 4(e) shows the results when the Coiflet 5 wavelet is used in the A-P direction at the second decomposition level, while in the S-I direction using the Coiflet 1 wavelet. The entropy feature was used in the A-P direction, while the log-energy feature was used in the S-I direction. Figure 4(f) shows the results when the Coiflet 5 wavelet is used in the A-P direction at the eight decomposition level, while in the S-I direction using the Coiflet 3 wavelet. The log-energy feature was used in both directions. The presented results show that good accuracy is achieved when considering low decomposition levels (e.g., second decomposition) or high decomposition levels (e.g. eight decomposition level). These results reinforce the results shown in Figure 3 in the sense that the choice of the mother wavelet, the decomposition level and the used features greatly influenced the accuracy.

Next, we considered only the cases when both the sensitivity and the specificity are above 70%. These results are summarized in Table I. These results showed that most accurate results were obtained when considering the Coiflet 5 wavelet in the A-P direction, and Coiflet 3 wavelet in the S-I direction. Furthermore, the decomposition levels in both directions were either the at second or eight decomposition levels. When considering the choice of a feature in the A-P direction, these results demonstrated that entropy or log-energy features can yield accurate results. However, the log-energy feature in the S-I direction was the only feature that yielded accurate results.

VI. DISCUSSION

In this study, we proposed a wavelet-based classifier to differentiate penetration-aspiration versus healthy swallows in patients suspected of having dysphagia. The presented results showed that the proposed approach is a viable approach for classification of swallowing vibrations yielding accuracy over 90%. These are more accurate results then any other previously proposed algorithms (e.g., [19], [24], [25]), while being of approximately the same computational complexity. It should be mentioned that these results are even more accurate than clinical judgement made by nurses or speech-language pathologists. Specifically, in our recent paper [48], nurses and speech-language pathologists, not aware of patient's clinical history, were asked to rate concurrently captured movies showing the faces of these same participants performing the swallow screening tasks. These clinicians were supposed to document any observed clinical signs of swallowing difficulty. The results of this study showed that nurses and speech-language pathologists achieved 54-75% sensitivity and specificities of 25-44% while detecting penetration-aspiration swallows.

The results also showed that in order to obtain very high accuracy the swallowing vibrations in two different directions should be decomposed using different decomposition levels, different wavelet functions and different features. This is an expected result as it has been already shown that swallowing vibrations in the A-P and S-I directions have different time-frequency characteristics (e.g., [23]). Our results re-enforce the previous contributions in the sense that vibrations in A-P and S-I directions provide complimentary information about swallowing.

The presented results depict the classification accuracy on a swallow-by-swallow basis. However, patients are asked to perform swallow several boluses in clinical settings and a patient may have multiple swallows from a single bolus. Therefore, obtaining very high accuracy on a swallow-by-swallow basis inherently means that we can identify aspirators with a very high accuracy. In other words, if we consider these results at the patient level, we would observe that the accuracy of aspiration detection is even greater.

It should be pointed out that the participants in the current study were patients with suspected swallowing difficulties. Therefore, the occurrence of penetration-aspiration is higher than in a general population. Therefore, future studies should involve healthy participants in addition to dysphagic participants in order the test the robustness of the algorithm.

VII. CONCLUSION

In this paper, we proposed a wavelet-based classification scheme for differentiating healthy swallows against aspiration swallows in patients suspected of dysphagia. The algorithm combined the wavelet packet analysis and Bayes classification along with linear discriminant analysis for feature reduction. The accuracy of the algorithm was examined using swallowing accelerometry recordings from forty patients during the regularly scheduled videouflouroscopy exams. The results of our numerical analysis showed that the proposed scheme achieved the classification accuracy over 90%.

REFERENCES

[1] J. A. Logemann, *Evaluation and treatment of swallowing disorders*, 2nd ed. Austin, Texas, USA: PRO-ED, 1998.
[2] A. J. Miller, *The neuroscientific principles of swallowing and dysphagia*. San Diego, USA: Singular Publishing Group, 1999.

[3] J. E. Curran, *Dysphagia : diagnosis and management*. Boston, USA: Butterworth-Heinemann, 1992, ch. Nutritional Considerations, pp. 255–266.

[4] D. G. Smithard, P. A. O'Neill, C. Park, J. Morris, R. Wyatt, R. England, and D. F. Martin, "Complications and outcome after acute stroke: Does dysphagia matter?" *Stroke*, vol. 27, no. 7, pp. 1200–1204, Jul. 1996.

[5] L. L. Riensche and K. Lang, "Treatment of swallowing disorders through a multidisciplinary team approach," *Educational Gerontology*, vol. 18, no. 3, pp. 277–284, 1992.

[6] O. Ekberg, S. Hamdy, V. Woisard, A. Wuttge-Hannig, and P. Ortega, "Social and psychological burden of dysphagia: Its impact on diagnosis and treatment," *Dysphagia*, vol. 17, no. 2, pp. 139–146, Apr. 2002.

[7] R. M. Miller, *Dysphagia : diagnosis and management*. Boston, USA: Butterworth-Heinemann, 1992, ch. Clinical Examination for Dysphagia, pp. 143–162.

[8] R. Ding and J. A. Logemann, "Pneumonia in stroke patients: A retrospective study," *Dysphagia*, vol. 15, no. 2, pp. 51–57, Mar. 2000.

[9] A. Tabaee, P. Johnson, C. J. Gartner, K. Kalwerisky, R. B. Desloge, and M. Stewart, "Patient-controlled comparison of flexible endoscopic evaluation of swallowing with sensory testing (FEESST) and videofluoroscopy," *The Laryngoscope*, vol. 116, no. 5, pp. 821–825, May 2006.

[10] D. J. C. Ramsey, D. G. Smithard, and L. Kalra, "Can pulse oximetry or a bedside swallowing assessment be used to detect aspiration after stroke?" *Stroke*, vol. 37, no. 12, pp. 2984–2988, Dec. 2006.

[11] C. Steele, C. Allen, J. Barker, P. Buen, R. French, A. Fedorak, S. Day, J. Lapointe, L. Lewis, C. MacKnight, S. McNeil, J. Valentine, and L. Walsh, "Dysphagia service delivery by speech-language pathologists in Canada: results of a national survey," *Canadian Journal of Speech-Language Pathology and Audiology*, vol. 31, no. 4, pp. 166–177, 2007.

[12] N. P. Reddy, B. R. Costarella, R. C. Grotz, and E. P. Canilang, "Biomechanical measurements to characterize the oral phase of dysphagia," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 4, pp. 392–397, Apr. 1990.

[13] N. P. Reddy, E. P. Canilang, J. Casterline, M. B. Rane, A. M. Joshi, R. Thomas, and R. Candadai, "Noninvasive accelaration measurements to characterize the pharyngeal phase of swallowing," *Journal of Biomedical Engineering*, vol. 13, pp. 379–383, Sep. 1991.

[14] S. Suryanarayanana, N. P. Reddy, and E. P. Canilang, "A fuzzy logic diagnosis system for classification of pharyngeal dysphagia," *International Journal of Bio-Medical Computing*, vol. 38, no. 3, pp. 207–215, Mar. 1995.

[15] N. P. Reddy, R. Thomas, E. P. Canilang, and J. Casterline, "Toward classification of dysphagic patients using biomechanical measurements," *Journal of rehabilitation research and development*, vol. 31, no. 4, pp. 335–344, Nov. 1994.

[16] N. P. Reddy, A. Katakam, V. Gupta, R. Unnikrishnan, J. Narayanan, and E. P. Canilang, "Measurements of acceleration during videofluorographic evaluation of dysphagic patients," *Medical Engineering and Physics*, vol. 22, no. 6, pp. 405–412, Jul. 2000.

[17] A. Das, N. P. Reddy, and J. Narayanan, "Hybrid fuzzy logic committee neural networks for recognition of swallow acceleration signals," *Computer Methods and Programs in Biomedicine*, vol. 64, no. 2, pp. 87–99, Feb. 2001.

[18] T. Chau, D. Chau, M. Casas, G. Berall, and D. J. Kenny, "Investigating the stationarity of paediatric aspiration signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 1, pp. 99–105, Mar. 2005.

[19] J. Lee, S. Blain, M. Casas, D. Kenny, G. Berall, and T. Chau, "A radial basis classifier for the automatic detection of aspiration in children with dysphagia," *Journal of NeuroEngineering and Rehabilitation*, vol. 3, no. 14, Jul. 2006, 17 pages.

[20] D. Zoratto, T. Chau, and C. M. Steele, "Hyolaryngeal excursion as the physiological source of swallowing accelerometry signals," *Physiological Measurement*, vol. 31, no. 6, pp. 843–855, Jun. 2010.

[21] Y. Kim and G. H. McCullough, "Maximum hyoid displacement in normal swallowing," *Dysphagia*, vol. 23, no. 3, pp. 274–279, Sep. 2008.

[22] R. Ishida, J. B. Palmer, and K. M. Hiiemae, "Hyoid motion during swallowing: Factors affecting forward and upward displacement," *Dysphagia*, vol. 17, no. 4, pp. 262–272, Dec. 2002.

[23] J. Lee, C. M. Steele, and T. Chau, "Time and time-frequency characterization of dual-axis swallowing accelerometry signals," *Physiological Measurement*, vol. 29, no. 9, pp. 1105–1120, Sep. 2008.

[24] M. Nikjoo, C. M. Steele, E. Sejdić, and T. Chau, "Automatic discrimination between safe and unsafe swallowing using a reputation-based classifier," *BioMedical Engineering OnLine*, vol. 10, no. 1, pp. 100–1–17, Nov. 2011.

[25] C. Merey, E. Sejdić, A. Kushki, G. Berall, and T. Chau, "Quantitative classification of pediatric swallowing through accelerometry," *Journal of Neuroengineering and Rehabilitation*, 2012, accepted.

[26] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics, 1992.

[27] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[28] S. G. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. San Diego: Academic Press, 1999.

[29] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Upper Saddle River, N. J., USA: Prentice Hall, 1998.

[30] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, Jan. 2009.

[31] S. Pittner and S. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 83–88, Jan. 1999.

[32] G. G. Yen and K.-C. Lin, "Wavelet packet feature extraction for vibration monitoring," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 3, pp. 650–667, Jun. 2000.

[33] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1186–1191, Nov. 1993.

[34] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[35] K. Fukunaga, *Introduction to statistical pattern recognition*, 2nd ed. Boston, USA: Academic Press, 1990.

[36] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis*. New York, USA: Wiley, 1973.

[37] S. S. Wilks, *Mathematical statistics*. New York, USA: Wiley, 1962.

[38] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: WCB/McGraw-Hill, 1991.

[39] J. C. Rosenbek, J. A. Robbins, E. B. Roecker, J. L. Coyle, and J. L. Wood, "A penetration-aspiration scale," *Dysphagia*, vol. 11, no. 2, pp. 93–98, Mar. 1996.

[40] E. Sejdić, V. Komisar, C. M. Steele, and T. Chau, "Baseline characteristics of dual-axis swallowing accelerometry signals," *Annals of Biomedical Engineering*, vol. 38, no. 3, pp. 1048–1059, Mar. 2010.

[41] E. A. Clancy and N. Hogan, "Single site electromyograph amplitude estimation," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 2, pp. 159–167, Feb. 1994.

[42] E. Sejdić, C. M. Steele, and T. Chau, "A method for removal of low frequency components associated with head movements from dual-axis swallowing accelerometry signals," *PLoS ONE*, vol. 7, no. 3, pp. e33 464–1–8, Mar. 2012.

[43] M. Unser, "Splines: a perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[44] E. Sejdić, C. M. Steele, and T. Chau, "A procedure for denoising dual-axis swallowing accelerometry signals," *Physiological Measurement*, vol. 31, no. 1, pp. N1–N9, Jan. 2010.

[45] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994.

[46] S. Beheshti and M. A. Dahleh, "A new information-theoretic approach to signal denoising and best basis selection," *IEEE Transactions on Signal Processing*, vol. 53, no. 10 Part 1, pp. 3613–3624, Oct. 2005.

[47] A. K. Jain., R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4 –37, Jan. 2000.

[48] C. M. Steele, S. M. Molfenter, G. L. Bailey, R. C. Polacco, A. A. Waito, D. C. B. H. Zoratto, and T. Chau, "Exploration of the utility of a brief swallow screening protocol with comparison to concurrent videofluoroscopy," *Canadian Journal of Speech-Language Pathology and Audiology*, vol. 35, no. 3, pp. 228–243, 2011.