

## Research Article

# Time-Frequency Analysis and Hermite Projection Method Applied to Swallowing Accelerometry Signals

Irena Orović,<sup>1</sup> Srdjan Stanković,<sup>1</sup> Tom Chau,<sup>2,3</sup> Catriona M. Steele,<sup>4,5</sup> and Ervin Sejdić<sup>2,3</sup>

<sup>1</sup>Electrical Engineering Department, University of Montenegro, 81000 Podgorica, Montenegro

<sup>2</sup>Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada M4G 1R8

<sup>3</sup>Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada M5S 3G9

<sup>4</sup>Toronto Rehabilitation Institute, Toronto, Ontario, Canada M5G 2A2

<sup>5</sup>Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, Canada M5G 1V7

Correspondence should be addressed to Ervin Sejdić, [esejdic@ieee.org](mailto:esejdic@ieee.org)

Received 22 December 2009; Accepted 22 March 2010

Academic Editor: Aydin Akan

Copyright © 2010 Irena Orović et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fast Hermite projections have been often used in image-processing procedures such as image database retrieval, projection filtering, and texture analysis. In this paper, we propose an innovative approach for the analysis of one-dimensional biomedical signals that combines the Hermite projection method with time-frequency analysis. In particular, we propose a two-step approach to characterize vibrations of various origins in swallowing accelerometry signals. First, by using time-frequency analysis we obtain the energy distribution of signal frequency content in time. Second, by using fast Hermite projections we characterize whether the analyzed time-frequency regions are associated with swallowing or other phenomena (vocalization, noise, bursts, etc.). The numerical analysis of the proposed scheme clearly shows that by using a few Hermite functions, vibrations of various origins are distinguishable. These results will be the basis for further analysis of swallowing accelerometry to detect swallowing difficulties.

## 1. Introduction

Time-frequency analysis is of great interest when time- or frequency-based techniques provide insufficient information for signal analysis [1–3]. Time-frequency representations (TFRs) depict variations of the spectral characteristics of signals as a function of time, which is ideally suited for nonstationary signals [1–10], especially, nonstationary biomedical signals [4, 5].

Many biomedical signals (e.g., heart sounds [4], swallowing accelerometry signals [6]) are multicomponent, one-dimensional signals. The time-frequency analysis of these signals provides a two-dimensional representation of the signals' components, which is appropriate for diagnostic analyses [5]. The resolution, that is, the quality of a representation, depends on a specific time-frequency distribution [5]. Due to the multicomponent nature of the considered biomedical signals, distributions with reduced cross-terms are recommended.

In this paper, we propose an approach for the analysis of two-dimensional time-frequency regions by using a fast

Hermite projection method [11–13]. Based on this idea, the reconstruction of different time-frequency regions would require a different number of Hermite functions depending on the component complexity. Hence, we adapt the fast Hermite projection method to characterize time-frequency regions of different origins, namely, by using a small fixed number of Hermite functions, each region is reconstructed with a certain error. A numerical analysis of these errors has shown that vibrations associated with swallowing have distinctive mean square errors compared to the errors obtained for the vibrations associated with other phenomena (e.g., vocalization or noise). Note that, in this application, the spectrogram-based reconstructions produce satisfactory results [3].

The paper is organized as follows. Section 2 describes swallowing accelerometry and outlines the advantages of this approach for detecting swallowing difficulties. In Section 3, we describe the theory behind the time-frequency analysis and depict a few illustrative examples on how various time-frequency distributions can be used for the representation of swallowing accelerometry signals. Section 4 describes the fast

Hermite projections and their application for characterization of vibrations within swallowing accelerometry recordings. In Section 5, we describe the results of the application of the proposed approach to swallowing accelerometry signals.

## 2. Swallowing Accelerometry

**2.1. Background.** Deglutition, or swallowing, is a well-defined, complex process of transporting food or liquid from the mouth to the stomach. Swallowing consists of four distinct phases: oral preparatory, oral, pharyngeal, and esophageal [14]. Patients suffering from dysphagia (swallowing difficulty) usually deviate from the well-defined pattern of healthy swallowing. Dysphagia is a common problem encountered in the rehabilitation of stroke patients, head injured patients, and others patients with paralyzing neurological diseases [15]. In these patients, dysphagia can develop from lesions in certain areas of the cortex and brain stem, which control the swallowing mechanism, or from damage to the cranial nerves associated with swallowing function [16, 17]. Dysphagic patients are likely to aspirate. Aspiration is defined as the entry of material into the airway below the true vocal folds [14]. Aspiration may have dire consequences including malnutrition and dehydration, degradation in psychosocial well-being, aspiration pneumonia, and even death [14, 18, 19].

Today's dysphagia management relies heavily on the videofluoroscopic swallowing study (VFSS) [14]. Even though VFSS is accepted as the gold standard, it requires expensive X-ray equipment as well as expertise from speech-language pathologists and radiologists. Hence, only a limited number of institutions can offer VFSS [18]. As a result, VFSS has been associated with long waiting lists [19]. 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.

Cervical auscultation is a promising noninvasive tool for the assessment of swallowing disorders [20]. It is being adopted by dysphagia clinicians as a component of the clinical evaluation of swallowing [21]. The cervical auscultation approach involves the examination of swallowing signals acquired via a stethoscope or other acoustic and/or vibration sensors during deglutition [22]. One such approach is signal detection via swallowing accelerometry [6, 23], which refers to an approach employing an accelerometer as a sensor during cervical auscultation. Swallowing accelerometry, a technique that involves an accelerometer placed on the neck to monitor vibrations associated with swallowing activities, has been used to detect aspiration 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 (e.g., [16, 17, 24]).

Nevertheless, the presence of various vibrations not associated with swallowing can severely contaminate swallowing accelerometry signals (e.g., [6, 23]). In particular, vocalizations either voluntarily or involuntarily can have an adverse effect on these signals. Their presence masks the observed swallowing signals and renders the recordings

less useful for further analysis [23]. For example, automatic demarcation of swallowing signals becomes impossible in the presence of vocalizations because of their strong amplitudes [23]. In order to develop an approach for removal of these confounding vibrations, a complete understanding of cervical accelerometry is needed.

**2.2. Data.** The sample data considered in this paper were gathered over a three-month period from a public science centre in Toronto, Ontario, Canada. A dual-axis accelerometer (ADXL322, Analog Devices) was attached to the participant's neck (anterior to the cricoid cartilage) using double-sided tape in order to monitor vibrations associated with swallowing. The axes of acceleration were aligned to the anterior-posterior and superior-inferior directions.

All participants provided written consent. The study protocol was approved by the research ethics boards of the Toronto Rehabilitation Institute and Bloorview Kids Rehab, both located in Toronto, Ontario, Canada. Participants had no documented swallowing disorders and passed an oral mechanism exam prior to participation.

Data were band-pass filtered in hardware with a pass band of 0.1–3000 Hz and sampled at 10 kHz using a custom LabVIEW program running on a laptop computer. Data were saved for subsequent offline analysis.

During data collection, participants were cued to perform three types of swallows involving saliva and water swallows. The entire data collection session lasted 15 minutes per participant. The participants were instructed not to vocalize. Nonetheless, approximately one quarter of all recordings contained either voluntary or involuntary vocalizations.

## 3. Time-Frequency Analysis of Swallowing Accelerometry Signals

Time-frequency representations are often used for the analysis of biomedical and speech signals due to their non-stationary nature [1–4]. The simplest and most commonly used time-frequency distribution is the spectrogram [3], defined as a square modulus of the short-time Fourier transform (STFT):

$$\text{SPEC}(t, \omega) = |\text{STFT}(t, \omega)|^2 = \left| \int_{-\infty}^{\infty} x(t + \tau) w(\tau) e^{-j\omega\tau} d\tau \right|^2, \quad (1)$$

where  $x(t)$  is a signal, and  $w(t)$  is a window function. The spectrogram is the energetic version of STFT (which is linear). The spectrogram is suitable for multicomponent signals representation, since it does not produce cross-terms.

Signals considered in this paper mostly contain swallowing vibrations. However, some recordings also contain vibrations associated with vocalization (e.g., speech, cough, and laughter) and various burst components produced by the equipment and noise. The most dominant vibrations are those produced by swallowing and vocalization. Fortunately, vibrations associated with different phenomena provide

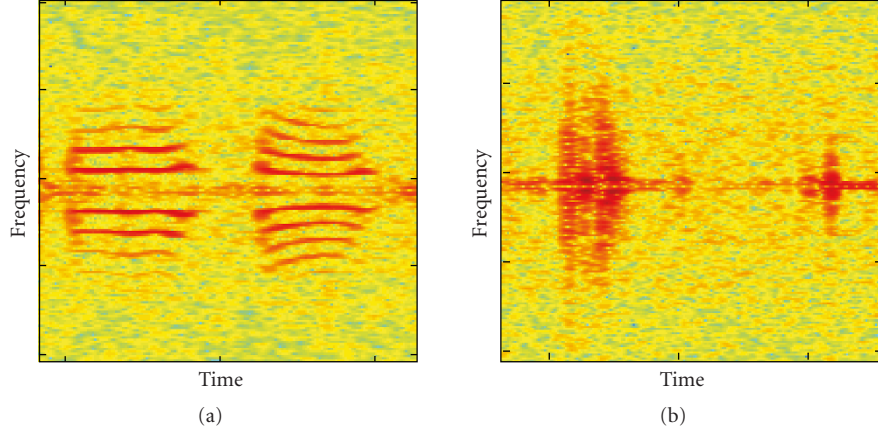


FIGURE 1: Time-frequency regions of spectrogram (a) for voiced sound (b) for swallowing vibrations.

unique time-frequency signatures (e.g., Figures 1(a) and 1(b)). Thus, time-frequency representations are crucial for the analysis and classification of these signals. Also, due to the multicomponent nature of these signals, time-frequency representations without cross-terms should be used. Hence, the spectrogram is employed in this paper.

The zoomed-in region of the spectrogram for sample vibrations associated with speech is given in Figure 1(a), while the region for the case of swallowing vibrations is given in Figure 1(b).

#### 4. Hermite Projection of Time-Frequency Regions

Two-dimensional regions in the time-frequency domain, as shown in Figure 1, can be considered to have specific “structures”. The complexity of these “structures” depends on the nature of vibrations while their refinement is related to the resolution achieved by the time-frequency distribution. Thus, the goal is to analyze these vibrations in the time-frequency domain in order to gain further insights into their complexities. To attain this goal, the fast Hermite projection method is adapted and applied to time-frequency regions.

**4.1. Fast Hermite Projection Method.** The fast Hermite projection method has been used in various image-processing applications, such as image database retrieval, projection filtering, and texture analysis [11–13]. It has been shown that, in comparison with the trigonometric functions, the expansion into Hermite functions provides better computational localization in both the signal domain and the transform domain [11, 25]. In addition, the expansion into Hermite functions allows for simultaneous analysis of the signal and its Fourier transform, since the Hermite functions are the eigenfunctions of the Fourier transform.

Although the computation of Hermite functions seems to be a demanding task, they could be easily obtained by

using recursive realization as follows [13]:

$$\begin{aligned}\Psi_0(x) &= \frac{1}{\sqrt[4]{\pi}} e^{-x^2/2}, & \Psi_1(x) &= \frac{\sqrt{2}x}{\sqrt[4]{\pi}} e^{-x^2/2}, \\ \Psi_p(x) &= x\sqrt{\frac{2}{p}}\Psi_{p-1}(x) - \sqrt{\frac{p-1}{p}}\Psi_{p-2}(x), & \forall p &\geq 2.\end{aligned}\quad (2)$$

The first step in the Hermite projection method is to remove the baseline, since [13]

$$\psi_p(x) \rightarrow 0, \quad |x| \rightarrow \infty. \quad (3)$$

For a two-dimensional signal  $I(x, y)$ , the baseline is defined as

$$b_x(y) = I(x, 0) + \frac{I(x, P) - I(x, 0)}{P} \cdot y, \quad (4)$$

where  $x = 0, \dots, P$  and  $y = 1, \dots, Q$ . Thus, the baseline is calculated for each  $y$  and then it is subtracted from the original values:

$$f(x, y) = I(x, y) - b(x, y). \quad (5)$$

The decomposition into  $N$  Hermite function is defined as [11]

$$f_y(x) = \sum_{p=0}^{N-1} c_p \Psi_p(x), \quad (6)$$

where  $f_y(x) = f(x, y)$  for a fixed  $y$  while the coefficients of Hermite expansion are obtained as

$$c_p(x) = \int_{-\infty}^{\infty} f_y(x) \Psi_p(x) dx. \quad (7)$$

The above equation can be approximated by Gauss-Hermite quadrature:

$$c_p(x) \approx \frac{1}{N} \sum_{m=1}^N \mu_{N-1}^p(x_m) f(x_m), \quad (8)$$

where  $x_m$  are zeros of Hermite polynomials:

$$H_p(x) = (-1)^p e^{x^2} \frac{d^p(e^{-x^2})}{dx^p}. \quad (9)$$

The constants  $\mu_{N-1}^p(x_m)$  are obtained by using the Hermite functions, as follows [13]:

$$\mu_{N-1}^p(x_m) = \frac{\psi_p(x_m)}{(\psi_{N-1}(x_m))^2}. \quad (10)$$

**4.2. Hermite Projection Method Applied to Time-Frequency Regions Associated with Various Vibrations.** To understand differences between swallowing vibrations and other vibrations (speech, laughter, cough, etc.) we observe the structures in the time-frequency regions.

A certain preprocessing during classification is applied, since the signal could be corrupted by noise. Thus, the time-frequency mask is defined to remove the noise influence and locate the significant signal components within the spectrogram:

$$L(t, \omega) = \begin{cases} 1 & \text{for } \text{SPEC}(t, \omega) \geq \xi, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where the threshold value  $\xi$ , that is, the energy floor, is obtained as

$$\xi = 10^{\sigma \log 10(\max_{t, \omega}(\text{SPEC}(t, \omega)))}. \quad (12)$$

Consequently, the filtered spectrogram is obtained as follows:

$$\text{SPEC}_{\text{filt}}(t, \omega) = L(t, \omega) \text{SPEC}(t, \omega). \quad (13)$$

Furthermore, the energy vector is calculated in order to determine the time intervals containing vibration activities:

$$E(t) = \sum_{\omega} |\text{SPEC}_{\text{filt}}(t, \omega)|. \quad (14)$$

Hence, the time support vector is obtained as

$$E_f(t) = \begin{cases} 1 & \text{for } E(t) \geq \beta, \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

where the threshold  $\beta$  is used to remove the remaining noise.

The time-frequency regions for classification are extracted from the  $\text{SPEC}_{\text{filt}}(t, \omega)$  for  $t = \arg\{E_f(t) = 1\}$ . The length of intervals containing vibrations is set to a constant value in order to have regions of the same size. In the next step, the fast Hermite projection method is applied to these regions.

The results are two types of regions: one with simple structures and the other with complex structures. The simple structure regions can be accurately reconstructed even with a small number of Hermite functions. On the other hand, complex structure regions require a significantly higher number of functions for successful reconstruction. Therefore, the goal in this paper is to use a small number

of functions for reconstruction. Such an approach allows us to reconstruct simple structure regions with a small error and complex structure regions with a significantly larger error. As a result, the difference between the original and reconstructed regions (i.e., the error, which depends on the structure of a region) is used for characterization. This difference is measured by the mean squared error as follows:

$$\text{MSE} = \frac{1}{T\Omega} \sum_t \sum_{\omega} (R(t, \omega) - R'(t, \omega))^2, \quad (16)$$

where  $R(t, \omega)$  denotes the original region from  $\text{SPEC}_{\text{filt}}(t, \omega)$ ,  $R'(t, \omega)$  is the reconstructed region, and  $T$  and  $\Omega$  are the region dimensions. Note that the dimensions  $T$  and  $\Omega$  in (16) should be the same for all regions. In other words, values for these dimensions are empirically derived in order for any region to include most of the vibration components.

It has been experimentally shown that in comparison to the regions containing swallowing vibrations, the vocalization regions (speech, cough, laughter, etc.) are more complex and thus have higher MSE values. Moreover, there is a significant difference between the values of MSE for vocalization and swallowing vibrations. As well, the MSE for regions with swallowing vibrations is much higher than for regions containing noise. Consequently, the difference between the original and reconstructed regions, given in terms of MSE, can be used as a parameter to characterize a region (e.g., swallowing, noise, or vocalizations).

As a remark, it should be noted that even though it is labelled as a projection method, the Hermite projection method actually expands a signal into a finite series of Hermite functions (multiplied by Hermite coefficients). In this paper, we apply the Hermite projection method to time-frequency regions by expanding each row (i.e., the instantaneous spectrum) into a series of Hermite functions.

## 5. Experimental Results

The window size, used for the spectrogram calculation is 512 samples. The parameter  $\sigma = 0.25$  is used in (12) to define the energy floor  $\xi$ . The size of the regions is 200 samples in time and 256 samples in frequency. Both the positive and negative frequency ranges are used to provide more components for the analysis. The illustration of filtered time-frequency regions for a complex signal is given in Figure 2(a). Note that the tested signals were disturbed by a strong component around zero frequency that has been pre-filtered. The time support vector is denoted by white color in Figure 2(b). The fast Hermite projection method is completed by using ten Hermite functions.

The results for different regions and vibrations are given in Table 1. The following regions are observed: regions containing noise, regions with bursts, regions with swallowing vibrations, and regions containing voiced vibrations (speech).

The regions with noise and regions with bursts have the smallest MSE, since they do not contain any significant structures. On the other hand, the regions containing vibrations associated with speech are very complex for reconstruction,



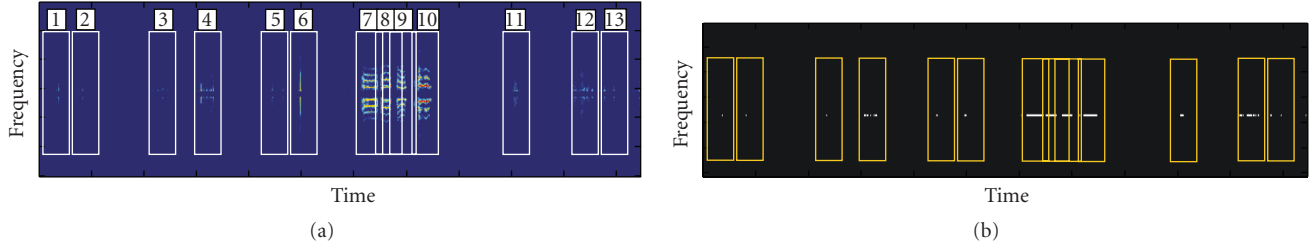


FIGURE 2: (a) Filtered time-frequency representation and assigned regions, (b) time support vector.

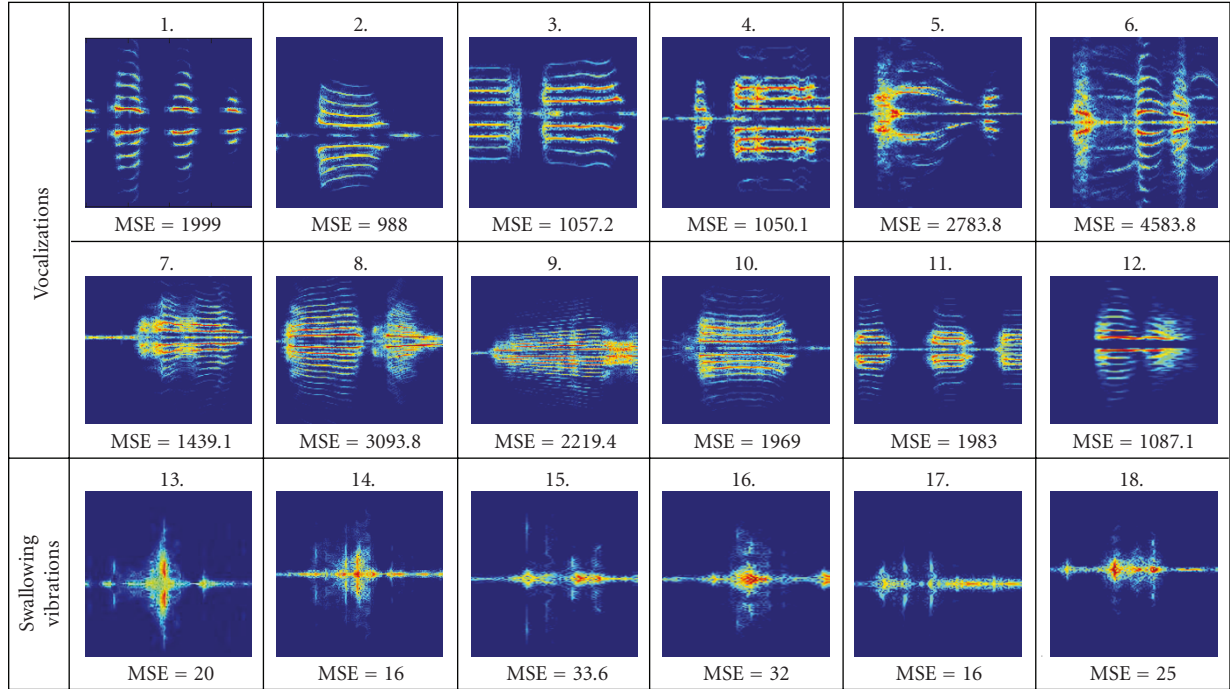


FIGURE 3: The illustrations of several zoomed time-frequency regions and corresponding MSEs.

TABLE 1: MSEs for the selected regions.

No.	Region description	MSE
1	Noise	0.086
2	Noise	0.01
3	Noise	0.22
4	Swallowing vibrations	36.29
5	Burst	0.13
6	Burst	3.4
7	Vocalization	669
8	Vocalization	1057.2
9	Vocalization	762
10	Vocalization	505
11	Noise	0.88
12	Swallowing vibrations	18.6
13	Noise	3

and thus we observed the highest MSEs in these regions. Finally, the regions with swallowing vibrations have MSEs

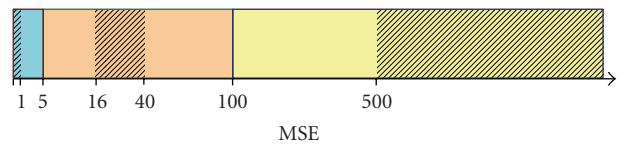


FIGURE 4: The illustration of classification intervals.

which are significantly smaller than those of vocalization regions and significantly higher than the MSEs for noisy regions.

While the experiment yielded over a thousand recordings, we tested our scheme on one hundred signals for illustrative purposes. The results for 18 sample regions from these recordings are presented in Figure 4. The corresponding MSEs are also included within Figure 3. The analysis of these recordings showed that in most cases, the regions containing noise and bursts have  $MSE < 1$ , though there are a few cases whereby  $MSE > 1$ . The MSE was never greater than 5 for these regions. The MSEs for the swallowing vibrations

are usually between 16 and 40 but always less than 100. For the vocalization vibrations, the majority of MSEs are higher than 500. The MSE intervals for various vibrations are illustrated in Figure 4: blue color denotes noise and bursts, red denotes swallowing vibrations, while vocalizations are denoted by yellow. Textures are added to illustrate three regions of MSE (noisy, swallowing, and vocalization) with the highest probability.

## 6. Discussion

It has been empirically shown that MSEs significantly differ between the considered vibration structures due to their different complexities. In other words, swallowing vibrations can be clearly differentiated from the vibrations associated with vocalizations and noise.

One of the advantages of the proposed approach is its simplicity. Specifically, it does not require pattern extraction or matching, and it is simpler than other texture analysis-based algorithms. For example, the texture analysis has been used for the time-frequency-based audio classification [26], where the spectrogram of musical sounds was composed of specific repetitive patterns and was treated as a visual texture. The feature extraction was performed by using time-frequency blocks where various blocks sizes were considered. Then, a block matching algorithm was used for each block to capture time-frequency structures at different orientations and scales. In comparison, our approach does not require preprocessing steps.

Additionally, even if we disregard a higher computational complexity, it is difficult to use texture analysis for the characterization of time-frequency regions associated with the considered vibrations (cough, laughter, speech, swallowing, etc.), since they do not follow typical image textures. Textures are usually defined as a regular repetition of an element or pattern on a surface [26, 27]. This is not the case with the considered time-frequency structures (e.g., Figure 3). Particularly, the swallowing sounds have a noise-like nature without edges and repetitive elements. Moreover, even for the regions involving coughing or laughing, it is difficult to define 2D patterns and rules that properly describe their directionality and periodicity.

## 7. Conclusion

In this paper, we proposed a procedure involving the Hermite projection of time-frequency regions as a way to characterize vibrations of different origins in swallowing accelerometry signals. In particular, the goal of this procedure was to detect regions associated with swallowing and vocalization vibrations (speech, cough, and laughter). The procedure involved two steps. First, we performed the time-frequency analysis of a signal yielding two-dimensional time-frequency regions containing different features depending on the nature of vibrations. Second, the fast Hermite projection method using a small number of Hermite functions was implemented for the time-frequency region reconstruction. Using the proposed approach, we successfully detected regions

associated with swallowing and vocalization by observing the MSE between the original and reconstructed regions. More precisely, the regions containing swallowing vibrations were reconstructed with significantly lower MSE in comparison to the regions containing vocalization vibrations. Furthermore, the MSE for the swallowing vibrations was sufficiently high to differentiate them from noise or bursts resulting from the signal recording equipment.

## Acknowledgments

I. Orović and S. Stanković's work is supported by the Ministry of Education and Science of Montenegro. T. Chau, C. M. Steele, and E. Sejdić's work is funded in part by the Ontario Centres of Excellence, the Toronto Rehabilitation Institute, the Health Technology Exchange, Bloorview Kids Rehab, and the Canada Research Chairs Program.

## References

- [1] P. J. Loughlin, "Scanning the special issue on time-frequency analysis," *Proceedings of the IEEE*, vol. 84, no. 9, p. 1195, 1996.
- [2] L. Cohen, *Time-Frequency Analysis: Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ, USA, 1995.
- [3] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, Elsevier, Amsterdam, The Netherlands, 1993.
- [4] E. Sejdić and J. Jiang, "Comparative study of three time-frequency representations with applications to a novel correlation method," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 2, pp. 633–636, Montreal, Canada, May 2004.
- [5] 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, 2009.
- [6] 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, 2008.
- [7] H. Choi and W. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 6, pp. 862–871, 1989.
- [8] LJ. Stanković, "Auto-term representation by the reduced interference distributions: a procedure for kernel design," *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1557–1563, 1996.
- [9] LJ. Stanković, "A method for time-frequency analysis," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 225–229, 1994.
- [10] I. Orović, S. Stanković, LJ. Stanković, and T. Thayaparan, "Multiwindow S-method for instantaneous frequency estimation and its application in radar signal analysis," *IET Signal Processing*, In press, 2010.
- [11] A. Krylov and D. Korchagin, "Fast Hermite projection method," in *Proceedings of the 3rd International Conference on Image Analysis and Recognition (ICIAR '06)*, vol. 1, pp. 329–338, Povia de Varzim, Portugal, September 2006.
- [12] D. Kortchagine and A. Krylov, "Image database retrieval by fast Hermite projection method," in *Proceedings of the 15th*

- International Conference on Computer Graphics and Applications (GraphiCon '05)*, pp. 308–311, Novosibirsk, Russia, June 2005.
- [13] D. Kortchagine and A. Krylov, "Projection filtering in image processing," in *Proceedings of the 10th International Conference on Computer Graphics and Applications (GraphiCon'00)*, pp. 42–45, Moscow, Russia, August–September 2000.
  - [14] J. A. Logemann, *Evaluation and Treatment of Swallowing Disorders*, PRO-ED, Austin, Texas, USA, 2nd edition, 1998.
  - [15] A. J. Miller, *The Neuroscientific Principles of Swallowing and Dysphagia*, Singular Publishing, San Diego, Calif, USA, 1999.
  - [16] N. P. Reddy, E. P. Canilang, R. C. Grotz, M. B. Rane, J. Castlerline, and B. R. Costarella, "Biomechanical quantification for assessment and diagnosis of dysphagia," *IEEE Engineering in Medicine and Biology Magazine*, vol. 7, no. 3, pp. 16–20, 1988.
  - [17] 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, 1990.
  - [18] C. M. Steele, C. Allen, J. Barker, et al., "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.
  - [19] 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, 2006.
  - [20] J. A. Y. Cichero and B. E. Murdoch, "The physiologic cause of swallowing sounds: answers from heart sounds and vocal tract acoustics," *Dysphagia*, vol. 13, no. 1, pp. 39–52, 1998.
  - [21] S. Hamlet, D. G. Penney, and J. Formolo, "Stethoscope acoustics and cervical auscultation of swallowing," *Dysphagia*, vol. 9, no. 1, pp. 63–68, 1994.
  - [22] S. R. Youmans and J. A. G. Stierwalt, "An acoustic profile of normal swallowing," *Dysphagia*, vol. 20, no. 3, pp. 195–209, 2005.
  - [23] E. Sejdić, C. M. Steele, and T. Chau, "Segmentation of dual-axis swallowing accelerometry signals in healthy subjects with analysis of anthropometric effects on duration of swallowing activities," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1090–1097, 2009.
  - [24] 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, article 14, pp. 1–17, 2006.
  - [25] W. F. Eberlein, "A new method for numerical evaluation of the Fourier transform," *Journal of Mathematical Analysis and Applications*, vol. 65, no. 1, pp. 80–84, 1978.
  - [26] G. Yu and J.-J. Slotine, "Audio classification from time-frequency texture," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '09)*, pp. 1677–1680, Taipei, Taiwan, 2009.
  - [27] M. Tuceryan and A. K. Jain, "Texture analysis," in *The Handbook of Pattern and Computer Vision*, C. H. Chen, L. F. Pau, and P. S. P. Wang, Eds., pp. 207–248, World Scientific, River Edge, NJ, USA, 2nd edition, 1998.