Understanding differences between healthy swallows and penetration-aspiration swallows via compressive sensing of tri-axial swallowing accelerometry signals

Ervin Sejdić[a], Joshua M. Dudik[a], Atsuko Kurosu[b], Iva Jestrović[a] and James L. Coyle[b]

[a]Department of Electrical and Computer Engineering, Swanson School of Engineering,
University of Pittsburgh, Pittsburgh, PA, 15261, USA
[b]Department of Communication Science and Disorders, School of Health and Rehabilitation
Sciences, University of Pittsburgh, PA, 15260, USA

ABSTRACT

Swallowing accelerometry is a promising tool for non-invasive assessment of swallowing difficulties. A recent contribution showed that swallowing accelerometry signals for healthy swallows and swallows indicating laryngeal penetration or tracheal aspiration have different time-frequency structures, which may be problematic for compressive sensing schemes based on time-frequency dictionaries. In this paper, we examined the effects of different swallows on the accuracy of a compressive sensing scheme based on modulated discrete prolate spheroidal sequences. We utilized tri-axial swallowing accelerometry signals recorded from four patients during routinely schedule videofluoroscopy exams. In particular, we considered 77 swallows approximately equally distributed between healthy swallows and swallows presenting with some penetration/aspiration. Our results indicated that the swallow type does not affect the accuracy of a considered compressive sensing scheme. Also, the results confirmed previous findings that each individual axis contributes different information. Our findings are important for further developments of a device which is to be used for long-term monitoring of swallowing difficulties.

Keywords: Swallowing difficulties, swallowing accelerometry signals, compressive sensing, modulated discrete prolate spheroidal sequences

1. 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 neurodegenerative diseases.[2] Patients suffering from dysphagia alongside certain other risk factors are prone to choking or other adverse events such as pneumonia, due to the increased chance of penetration/aspiration. Aspiration is defined as process when any food or fluids enter into the airway below the true vocal folds.[1] A related but less severe correlate to aspiration, laryngeal penetration is defined as the event when material enters the space of the upper airway that lies above the true vocal folds (the supraglottic space) but is not observed to fall below the vocal folds during assessment.[2]

In recent years, swallowing accelerometry has become a promising non-invasive tool for the screening of swallowing function, including prediction of the presence of penetration-aspiration. Swallowing accelerometry refers to the employment of an accelerometer as a sensor modality during cervical auscultation. A recent contribution showed that swallows in which penetration or aspiration occur, have different time-frequency structures from healthy swallows.[3] In this paper, we propose to examine whether the differences in the time-frequency structure of swallowing accelerometry signals will play a significant role in acquiring such signals using compressive sensing. In particular, we examined the compressive sensing approach based on modulated discrete prolate spheroidal sequences.[4,5]

Further author information: (Send correspondence to E.S.)
E.S.: E-mail: esejdic@ieee.org, Telephone: 1 412-624-0508
J.M.D.: E-mail: jmd151@pitt.edu
I. J.: E-mail: ivj2@pitt.edu
A. K.: E-mail: atk22@pitt.edu
J.L.C.: E-mail: jcoyle@pitt.edu

2. DATA COLLECTION

Four participants were recruited from the population of adult patients that had a history of swallowing difficulties and were undergoing a videofluoroscopic evaluation at the University of Pittsburgh Medical Center's Presbyterian University Hospital (Pittsburgh, PA). Those patients that had a history of head or neck surgery, were equipped assistive devices that obstructed the anterior neck, or were not in a condition to consent were not included in the study, but no other conditions were excluded. Data was recorded while the patient underwent the typical videofluoroscopic swallowing imaging studies under the guidance of a speech language pathologist. This procedure uses x-ray to observe the patient swallow liquids and solids of varying textures and volumes including thin and thick liquids, puree textures and masticated solids. This protocol was approved by the Institutional Review Board at the University of Pittsburgh.

A tri-axial accelerometer (ADXL 327, Analog Devices, Norwood, Massachusetts) was placed over the cricoid cartilage along the midline of the anterior neck and affixed with surgical tape. The main axes of the accelerometer were aligned approximately parallel to the cervical spine and coronal plane, respectively, which resulted in the third axis being perpendicular to those structures and approximately parallel to the patient's shoulders. The sensor was powered by a power supply (model 1504, BK Precision, Yorba Linda, California) with a 3V output, and the resulting signals were bandpass filtered from 0.1 to 3000 Hz with ten times amplification (model P55, Grass Technologies, Warwick, Rhode Island). The microphone (model C 411L, AKG, Vienna, Austria) was placed inferior to the accelerometer and slightly towards the left lateral side of the trachea to avoid contact between the two sensors during recording. The microphone's power supply (model B29L, AKG, Vienna, Austria) was set to set to match the output impedance via the 'line' setting and to amplify the signal with a volume setting of '9'. All four of these signals were input into a National Instruments 6210 DAQ and sampled at 20 kHz by a custom Labview program (National Instruments, Austin, Texas). Concurrent videofluoroscopy images output by the x-ray machine (Ultimax system, Toshiba, Tustin, CA) were obtained at 30 pulses per second and were recorded at 30 frames per second by a video capture card (AccuStream Express HD, Foresight Imaging, Chelmsford, MA) via the same Labview program.

3. COMPRESSIVE SENSING OF TRI-AXIAL SWALLOWING ACCELEROMETRY SIGNALS

Continuous monitoring of relatively low bandwidth signals such as swallowing accelerometry signals can produce a large number of redundant samples, which severely constraints our processing efforts. To reduce the number of available samples, we can implement a compressive sensing (CS) approach,[6–8] which is particularly suited for K-sparse signals. Such a K-sparse, discrete-time signal of dimension N is encoded by computing a measurement vector y that consists of $M << N$ linear projections of the vector s:

$$y = \Phi s \tag{1}$$

where Φ represents an $M \times N$ matrix and is often referred to as the sensing matrix.[7] A natural formulation of the recovery problem is within a norm minimization framework, which seeks a solution to the problem

$$\min ||s||_0 \text{ subject to } \|y - \Phi s\|_2 < \eta \tag{2}$$

where η is the expected noise of measurements, $||s||_0$ counts the number of nonzero entries of x and $\|\bullet\|_2$ is the Euclidian norm. Since eqn. (2) will not always yield an accurate representation of biomedical signals, it is desired to find a method that will allow for "precise" recovery of the signals (i.e., with a very small error). To accomplish this, an appropriate domain is chosen in which these biomedical signals are sparse. Rewriting eqn. (1) to accommodate this, we obtain:[9]

$$y = \Phi s = \Phi \Psi x \tag{3}$$

where $s = \Psi x$ represents a sparse representation of a biomedical signal in a domain given by Ψ and x represents expansion coefficients. There are a number of different choices for the matrix Ψ. We use a time-frequency dictionary based on modulated discrete prolate spheroidal sequences (MDPSS), which are based on discrete prolate spheroidal sequences (DPSS). Given N such that $n = 0, 1, ..., N-1$ and the normalized half-bandwidth, W such that $0 < W < 0.5$, the kth DPSS, $v_k(n, N, W)$, is defined as the real solution to the system of equations:[10]

$$\sum_{m=0}^{N-1} \frac{\sin[2\pi W(n-m)]}{\pi(n-m)} v_k(m, N, W) = \lambda_k(N, W) v_k(n, N, W) \quad k = 0, 1, ..., N-1 \tag{4}$$

with $\lambda_k(N, W)$ being the ordered non-zero eigenvalues of (4): $\lambda_0(N, W) > \lambda_1(N, W), ..., \lambda_{N-1}(N, W) > 0$. The DPSS are doubly orthogonal, that is, they are orthogonal on the infinite set $\{-\infty, ..., \infty\}$ and orthonormal on the finite set $\{0, 1, ..., N-1\}$. DPSS are well suited for signals that occupy the same band as these sequences. However, they do not necessarily yield a sparse representation when a signal is centered around some frequency $|\omega_o| > 0$ and occupies bandwidth smaller than $2W$. To resolve this isuee, MDPSS were proposed in,[4]:[11]

$$M_k(N, W, \omega_m; n) = \exp(j\omega_m n) v_k(N, W; n) \tag{5}$$

where $\omega_m = 2\pi f_m$ is a modulating frequency. MDPSS are also doubly orthogonal and are bandlimited to the frequency band $[-W + \omega_m : W + \omega_m]$.

MDPSS form a time-frequency dictionary with the first few bases in the dictionary being the actual DPSS with bandwidth W. Additional bases are obtained by partitioning the band $[-\omega; \omega]$ into K subbands with the boundaries of each subband given by $[\omega_k; \omega_{k+1}]$, where $0 \leq k \leq K-1$, $\omega_{k+1} > \omega_k$, and $\omega_0 = -\omega$, $\omega_{K-1} = \omega$. Hence, each set of MDPSS has a bandwidth equal to $\omega_{k+1} - \omega_k$ and a modulation frequency equal to $\omega_m = 0.5(\omega_k + \omega_{k+1})$. Obviously, a set of such function again forms a basis of functions limited to the bandwidth $[-\omega; \omega]$. Here, we partition the bandwidth in equal blocks to reduce amount of stored pre-computed DPSS.

Unfortunately, eqn. (2) is not suitable for many applications as it is NP-hard.[12] To avoid the computational burden, we focused on the matching pursuit[13] and MDPSS bases.[5, 14]

4. DATA ANALYSIS

Two judges, both speech language pathologists with published dysphagia research experience, visually inspected the fluoroscopic data to measure two parameters: the duration of the swallowing segments (in order to determine the portions of the recorded swallow events containing the acoustic and vibratory signals of interest), and the extent of airway penetration or aspiration during the swallowing segments.

Segment durations were defined as the duration between the first video frame at which time the leading edge of the swallowed material (bolus head) was visible within the pharynx, and the first video frame at which the hyoid bone had returned to a resting or stable position at the end of the swallow. The radiographic shadow of the posterior edge of the ramus of the mandible in the lateral plane image, is the anatomical landmark routinely used in dysphagia research to indicate the plane of the entrance to the pharynx.[15] Likewise, the return of the hyoid bone to rest after the swallow is routinely used in dysphagia research to indicate the physiologic end of the pharyngeal stage of the oropharyngeal swallow.[15]

One expert judge with previously established judgment validity and inter- and intra-rater reliability for these measures,[16] trained the second judge in methods of selection of frames for segment durations, and in rating of the extent of airway protection during the swallow, using the eight-point penetration-aspiration scale.[16, 17] After training, both judges evaluated a set of twenty-five video recordings of unfamiliar swallows, none of which were included in the participant data for the present study. Judgment reliability was evaluated using the intraclass correlation coefficient[18] and Cronbach's alpha, and both metrics were greater than 0.90. Following establishment of acceptable intra- and inter-rater reliability for segment durations and penetration-aspiration scores, each judge then evaluated approximately one-half of the swallows recorded from four subjects (77 swallows) and recorded segment onset, segment offset, and penetration-aspiration scale scores for each swallow.

Segmented swallows were then divided into two groups: (1) swallows with scores equal to 1 on the penetration-aspiration scale (i.e., completely healthy swallows); (2) swallows with some penetration-aspiration (i.e., swallows with scores from 2 to 8). Using these two groups, we divided each swallow into segments of 128 samples, and we attempted to accurately reconstruct each of the segments from sparse samples. Specifically, we assumed that only 30% of the original samples are available, while examining whether the uniform or non-uniform sub-Nyquist rates have significant effects on the overall accuracy. We use a 15-band MDPSS based dictionary with the normalized half-bandwidth equal to 0.25. To understand the effects of different type of swallows, we adopted performance metrics used in previous contributions.[5, 8, 19, 20] Percent root difference (PRD) quantifies distortion in reconstructed biomedical signals and is defined as:

$$PRD(\%) = \sqrt{\frac{\sum_{n=1}^{N}\left(x(n)-\widehat{x}(n)\right)^2}{\sum_{n=1}^{N}x^2(n)}} \times 100\% \tag{6}$$

Next, it is desired to minimize root mean square error (RMSE) when finding the optimal approximation of the signal. RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}\left(x(n)-\widehat{x}(n)\right)^2}{N}} \tag{7}$$

To understand the magnitude of local distortions, we utilized maximum error (MERR) defined as:

$$MERR = \max\left(x(n)-\widehat{x}(n)\right) \tag{8}$$

Cross-correlation (CC) defined as:

$$CC = \frac{\sum_{n=1}^{N}\left(x(n)-\mu_x\right)\left(\widehat{x}(n)-\mu_{\widehat{x}}\right)}{\sqrt{\sum_{n=1}^{N}\left(x(n)-\mu_x\right)^2}\sqrt{\sum_{n=1}^{N}\left(\widehat{x}(n)-\mu_{\widehat{x}}\right)^2}} \times 100\% \tag{9}$$

is used to evaluate the similarity between the original and the reconstructed signal. $x(n)$ is the original signal and $\widehat{x}(n)$ represents a reconstructed signal. In addition, μ_x and $\mu_{\widehat{x}}$ denote the mean values of $x(n)$ and $\widehat{x}(n)$, respectively.

In order to establish statistical significance of our results, a non-parametric inferential statistical methods known as the Mann-Whitney test[21] and the Kruskal-Wallis[22] were used. A 5% significance was used.

5. RESULTS AND DISCUSSION

Figures 1 and 2 summarize the results of our analysis. Figure 1 is devoted to the results obtained while sampling the tri-axial accelerometry signals at regular intervals, while Figure 2 summarizes the results for irregular sampling intervals.

Statistical differences were only observed between the two groups when considering the PRD metric for swallows in the A-P direction ($p < 0.01$). No other statistical differences were found. Hence, even though healthy swallows and penetration-aspiration swallows may have different time-frequency structures, these differences are not important while implementing a compressive sensing approach based on a time-frequency dictionary considered here.

Interestingly enough, most of the metrics are statistically different amongst three anatomical directions ($p < 0.05$), which confirms our earlier findings that these direction carry mutually exclusive information.

6. CONCLUSIONS

In this paper, we demonstrated that the type of swallows (healthy swallows versus swallows indicating penetration-aspiration) will not affect the accuracy of the considered compressive sensing approach. We also confirmed earlier findings that the three anatomical directions carry different information as demonstrated by different magnitudes of the considered metrics.

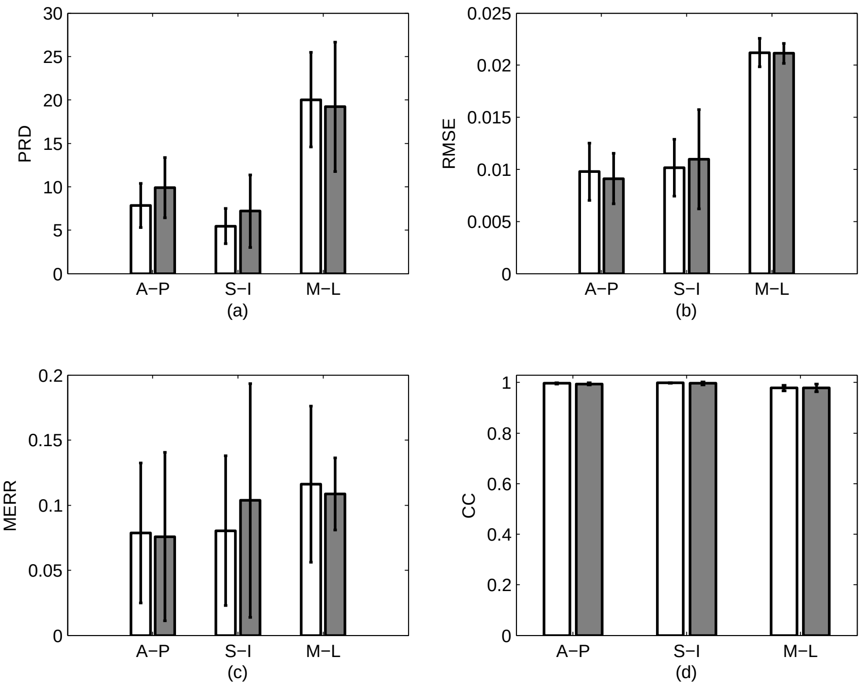

Figure 1. The assessment metrics for the case of reduced regular sampling.

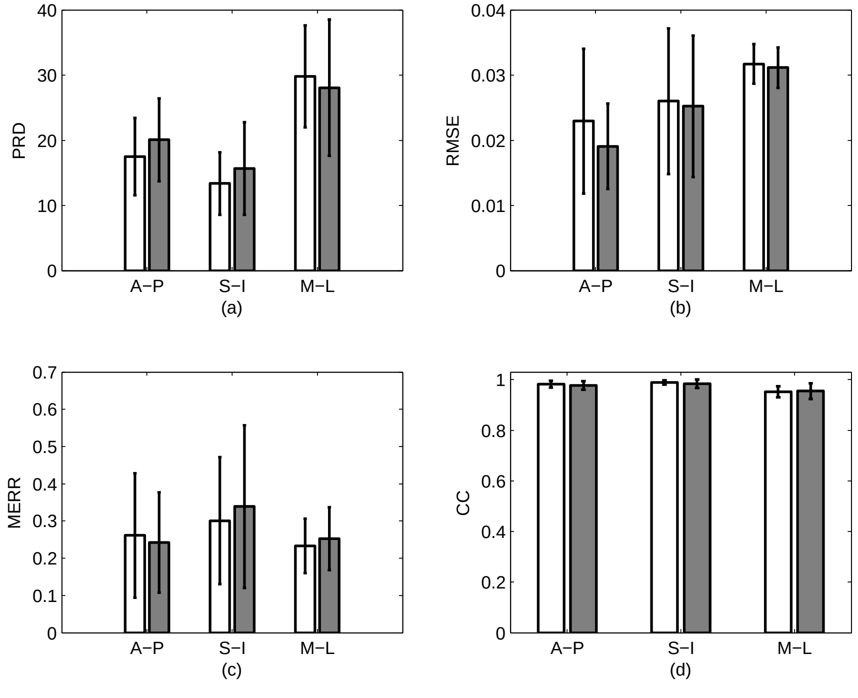

Figure 2. The assessment metrics for the case of reduced irregular sampling.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute Of Child Health and Human Development of the National Institutes of Health under Award Number R01HD074819. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

1. J. A. Logemann, *Evaluation and treatment of swallowing disorders*, PRO-ED, Austin, Texas, USA, second ed., 1998.
2. A. J. Miller, *The neuroscientific principles of swallowing and dysphagia*, Singular Publishing Group, San Diego, USA, 1999.
3. E. Sejdić, C. M. Steele, and T. Chau, "Classification of penetration-aspiration versus healthy swallows using dual-axis swallowing accelerometry signals in dysphagic subjects," *IEEE Transactions on Biomedical Engineering* **60**, pp. 1859–1866, July 2013.
4. E. Sejdić, M. Luccini, S. Primak, K. Baddour, and T. Willink, "Channel estimation using DPSS based frames," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, pp. 2849–2852, (Las Vegas, Nevada, USA), Mar./Apr. 31–4, 2008.
5. E. Sejdić, A. Can, L. F. Chaparro, C. M. Steele, and T. Chau, "Compressive sampling of swallowing accelerometry signals using time-frequency dictionaries based on modulated discrete prolate spheroidal sequences," *EURASIP Journal on Advances in Signal Processing* **2012**, pp. 101–1–14, May 2012.
6. D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory* **52**, pp. 1289–1306, Apr. 2006.
7. W. Dai and O. Milenković, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transaction on Information Theory* **55**, pp. 2230–2249, May 2009.
8. K.-K. Poh and P. Marziliano, "Compressive sampling of EEG signals with finite rate of innovation," *EURASIP Journal on Advances in Signal Processing* **2010**, pp. 183105–1–12, 2010.
9. S. Stanković, I. Orović, and E. Sejdić, *Multimedia Signals and Systems*, Springer US, New York, NY, 2012.
10. D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case," *The Bell System Technical Journal* **57**, pp. 1371–1430, May/June 1978.
11. J. Oh, S. Senay, and L. F. Chaparro, "Signal reconstruction from nonuniformly spaced samples using evolutionary Slepian transform-based POCS," *EURASIP Journal on Advances in Signal Processing* **2010**, pp. 367317–1–12, 2010.
12. H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Transactions on Information Theory* **54**, pp. 2210 –2219, May 2008.
13. S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing* **41**, pp. 3397–3415, Dec. 1993.
14. E. Sejdić and L. F. Chaparro, "Time-frequency representations based on compressive samples," in *Proc. of 21st European Signal Processing Conference (EUSIPCO'13)*, pp. 1569742057–1–4, (Marrakech, Morocco), Sept. 9–13, 2013.
15. G. Lof and J. Robbins, "Test-retest variability in normal swallowing," *Dysphagia* **4**(4), pp. 236–242, 1990.
16. J. C. Rosenbek, J. A. Robbins, E. B. Roecker, J. L. Coyle, and J. L. Wood, "A penetration-aspiration scale," *Dysphagia* **11**, pp. 93–98, Mar. 1996.
17. J. Robbins, J. Coyle, J. Rosenbek, E. Roecker, and J. Wood, "Differentiation of normal and abnormal airway protection during swallowing using the penetration-aspiration scale," *Dysphagia* **14**(4), pp. 228–232, 1999.
18. P. E. Shrout and J. K. Fleiss, "Intraclass correlations: Uses in assessing rater reliability," *Psychological Bulletin* **86**, pp. 420–428, Mar. 1979.
19. B. Boucheham, Y. Ferdi, and M. C. Batouche, "Recursive versus sequential multiple error measures reduction: A curve simplification approach to ECG data compression," *Computer Methods and Programs in Biomedicine* **81**, pp. 162–173, Feb. 2006.
20. F. Scholkmann, S. Spichtig, T. Muehlemann, and M. Wolf, "How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation," *Physiological Measurement* **31**, pp. 649–662, May 2010.
21. H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics* **18**, pp. 50–60, Mar. 1947.
22. W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion analysis of variance," *Journal of the American Statistical Association* **47**, pp. 583–621, Dec. 1952.