The effects of compressive sensing on extracted features from tri-axial swallowing accelerometry signals

Ervin Sejdić[a], Faezeh Movahedi[a], Zhenwei Zhang[a], Atsuko Kurosu[b] 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

Acquiring swallowing accelerometry signals using a comprehensive sensing scheme may be a desirable approach for monitoring swallowing safety for longer periods of time. However, it needs to be insured that signal characteristics can be recovered accurately from compressed samples. In this paper, we considered this issue by examining the effects of the number of acquired compressed samples on the calculated swallowing accelerometry signal features. We used tri-axial swallowing accelerometry signals acquired from seventeen stroke patients (106 swallows in total). From acquired signals, we extracted typically considered signal features from time, frequency and time-frequency domains. Next, we compared these features from the original signals (sampled using traditional sampling schemes) and compressively sampled signals. Our results have shown we can obtain accurate estimates of signal features even by using only a third of original samples.

Keywords: Compressive sensing, tri-axial swallowing accelerometry signals, swallowing difficulties, modulated discrete prolate spheroidal sequences

1. INTRODUCTION

Dysphagia is referred to any difficulty in the swallowing stages[1] and often is caused by other medical conditions that can impair the structure, strength, and coordination of the muscles involved in the swallowing. Some common medical conditions related to dysphagia are stroke, head injury, and neurodegenerative diseases.[2] Dysphagia can result in serious health conditions, which could negatively affect the quality of life in patients suffering from swallowing difficulties. Aspiration and penetration of food into the airway are probable complications of dysphagia. In aspiration, a swallowed bolus passes into the airway and often into the lungs,[1] while in penetration, bolus remains in the upper airway and does not fall more deeply into the airway.[2]

Swallowing accelerometry has been recently considered as a new approach as a non-invasive screening test for swallowing difficulties and penetration-aspiration during swallowing. This approach is based on the transduction of vibrations recorded from the upper aerodigestive tract structures during swallowing the act of swallowing into a voltage signal. In this study, swallowing vibrations are recorded in the three axes of anterior-posterior (A-P), superior-interior (S-I), and medial-lateral (M-L). A recent study has captured discriminative time-frequency structures in swallows with penetration or aspiration in healthy subjects.[3] The aim of this study is to investigate the role of differences in the time-frequency structure of swallowing accelerometry signals in acquiring such signals by applying compressive sensing. We specifically administered modulated discrete prolate spheroidal sequences approach for the compressive sensing method.[4, 5]

2. DATA COLLECTION

In this study, we considered data from seventeen stroke patients that underwent videofluoroscopic evaluation at the University of Pittsburgh Medical Center's Presbyterian University Hospital (Pittsburgh, PA) due to their history of swallowing difficulties. In total, we considered 106 individual swallows from these patients. All considered data in this paper was obtained while participants were undergoing typical videofluoroscopic swallowing imaging studies under the guidance of a speech language pathologist. During this procedure, the structure and biomechanical activity of the upper aerodigestive tract are captured during ingestion of different consistencies of barium coated materials. All participants signed informed consent and the data collection protocol was approved by the University of Pittsburgh Institutional Review Board.

Our recording equipment consisted of a tri-axial accelerometer (ADXL 327, Analog Devices, Norwood, Massachusetts) attached to the participant's anterior neck with double-sided tape. The accelerometer was mounted in a custom plastic case, and affixed over the cricoid cartilage. The S-I and A-P axes of accelerometer were aligned approximately parallel to the cervical spine and perpendicular to the coronal plane, respectively, and the M-L axis was perpendicular to those axes 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 signals from the A-P, S-I, and M-L axes were bandpass filtered from 0.1 to 3000 Hz with ten times amplification (model P55, Grass Technologies, Warwick, Rhode Island). We also recorded the swallow sounds using a microphone (model C 411L, AKG, Vienna, Austria) however this data was not used for analysis in this study. The voltage signals for three axes of the accelerometer were fed into a National Instruments 6210 DAQ and recorded at 20 kHz by the LabView program Signal Express (National Instruments, Austin, Texas). Videofluoroscopy images were captured by the x-ray machine (Ultimax system, Toshiba, Tustin, CA) at 30 pulses per second and were obtained at 60 frames per second by a video card (AccuStream Express HD, Foresight Imaging, Chelmsford, MA), and information was recorded on hard drive by the LabView program Signal Express (National Instruments, Austin, Texas).

3. COMPRESSIVE SENSING OF TRI-AXIAL SWALLOWING ACCELEROMETRY SIGNALS

In our previous contributions, we have demonstrated that continuous monitoring of swallowing accelerometry signals can produce a large number of redundant samples.[5, 6] As these signals are K-sparse signals, we proposed a compressive sensing (CS) approach such that a 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:

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

where $s = \Psi x$ represents a sparse representation of a biomedical signal in a domain given by Ψ and x represents expansion coefficients. We use a time-frequency dictionary based on modulated discrete prolate spheroidal sequences (MDPSS), which are based on discrete prolate spheroidal sequences. To resolve this isuee, MDPSS were proposed in,[4][7]

$$M_k(N, W, \omega_m; n) = \exp(j\omega_m n) v_k(N, W; n) \tag{2}$$

where $\omega_m = 2\pi f_m$ is a modulating frequency, $v_k(n, N, W)$ is the kth discrete prolate spheroidal sequence as defined in.[8]

MDPSS form a time-frequency dictionary with the first few bases in the dictionary being the actual DPSS with bandwidth W. To avoid the computational burden, we focused on the matching pursuit[9] and MDPSS bases.[5, 10]

4. DATA ANALYSIS

Two judges who are both speech language pathologists with published dysphagia research experience measure two parameters by inspecting the fluoroscopic data visually using image processing software (ImageJ, NIH): the duration of the swallowing segments and the extent of airway penetration or aspiration during the swallowing

segments.[12] The measure of duration is to determine the portions of the recorded swallowing events containing the acoustic and vibratory signals of interest.

We define the duration between the first video frame when the leading edge of the swallowed material (bolus head) was visible within the pharynx, and until the first video frame when hyoid bone returned to a resting or stable position after swallow as segment duration. The radiographic shadow of the posterior edge of the ramus of the mandible in the lateral plane image indicates the plane of the entrance to the pharynx, which is the anatomical landmark commonly used in dysphagia research.[11] Likewise, the return of the hyoid bone to rest after the swallow commonly indicates the physiologic end of the pharyngeal stage of the oropharyngeal swallow.[11]

One judge was trained by another expert judge who previously established judgment validity and inter- and intra-rater reliability for these measures[12] The train contained methods of selection of frames for segment durations, and the eight-point penetration-aspiration scale for rating of the extent of airway protection during the swallow.[12, 13] Twenty five unfamiliar video recording, which will not be used in present study, were evaluated by both judges after training. The intraclass correlation coefficient[14] and Cronbach's alpha were used to evaluate judgment reliability, and both metrics were greater than 0.90. After establishing acceptable intra- and inter-rater reliability for segment durations and penetration-aspiration scores, each judge then recorded segment onset, segment offset, and penetration-aspiration scale scores for each swallow.

Next, we extracted a number of features that were previously considered for swallowing accelerometry signals.[15–17] Specifically, we considered the standard deviation (σ), skewness (γ) and kurtosis (ς) in the time domain. In the frequency domain, we considered the peak frequency (f_p), centroid frequency ($\hat{f}$) and bandwidth (BW) values. Lastly, we consider the wavelet entropy (θ) as a sample time-frequency (time-scale) feature.

In order to establish statistical significance of our results, a non-parametric inferential statistical methods known as the Mann-Whitney test[18] and the Kruskal-Wallis[19] were used. A 5% significance was used.

5. RESULTS AND DISCUSSION

Tables 1-3 summarize the results of our analysis. When considering the time domain features, except a single value for ς_{ML} at $N/4$, no other statistical differences were noted. Even when only 20% of the original samples were used, we were still able to accurately extract time domain features.

Table 1. The effects of compressive sampling on typical time-domain features. $\sharp$ denotes statistical differences between the original values and $N/4$ values.

	Original	$N/2$	$N/3$	$N/4$	$N/5$
σ_{AP}	0.12 ± 0.06	0.12 ± 0.06	0.12 ± 0.06	0.12 ± 0.06	0.12 ± 0.06
σ_{SI}	0.16 ± 0.09	0.16 ± 0.09	0.16 ± 0.09	0.16 ± 0.09	0.16 ± 0.09
σ_{ML}	0.11 ± 0.09	0.11 ± 0.09	0.11 ± 0.09	0.12 ± 0.09	0.12 ± 0.09
γ_{AP}	0.29 ± 1.32	0.29 ± 1.32	0.24 ± 1.20	0.13 ± 1.27	0.09 ± 1.17
γ_{SI}	-0.44 ± 1.21	-0.44 ± 1.21	-0.43 ± 1.15	-0.40 ± 0.96	-0.39 ± 0.98
γ_{ML}	0.06 ± 0.62	0.06 ± 0.62	0.06 ± 0.61	0.06 ± 0.64	0.03 ± 0.62
ς_{AP}	14.0 ± 18.0	13.8 ± 17.3	12.6 ± 15.5	11.9 ± 12.7	11.5 ± 12.9
ς_{SI}	7.82 ± 10.1	7.82 ± 10.2	7.45 ± 9.42	7.87 ± 8.73	7.26 ± 7.87
ς_{ML}	4.70 ± 3.82	4.70 ± 3.80	4.59 ± 3.51	$5.00 \pm 3.25^{\sharp}$	4.63 ± 2.74

While considering the frequency features, we can notice slightly different patterns than for the time domain features. In this case, most robust results are for $N/3$ samples (i.e, we have the least amount of statistical differences). However, it is interesting to note that the estimate of the peak frequency value is robust regardless of the number of samples used, while bandwidth values are the least sensitive to the number of samples used.

Lastly, we can observe that the wavelet entropy values are less robust for $N/4$ and $N/5$ number of samples. However, it is interesting to note that even for $N/3$ samples, we have very accurate estimates of the wavelet entropy.

Table 2. The effects of compressive sampling on typical frequency-domain features. ‡ denotes statistical differences between the original values and $N/2$ values. † denotes statistical differences between the original values and $N/3$ values. ♯ denotes statistical differences between the original values and $N/4$ values. ♮ denotes statistical differences between the original values and $N/5$ values.

	Original	N/2	N/3	N/4	N/5
f_{pAP}	6.26 ± 18.3	6.26 ± 18.3	6.26 ± 18.3	5.80 ± 12.2	7.85 ± 36.2
f_{pSI}	4.07 ± 5.91	3.82 ± 5.46	3.84 ± 5.38	4.72 ± 8.94	4.38 ± 8.09
f_{pML}	3.50 ± 5.55	3.25 ± 5.05	3.25 ± 5.05	5.19 ± 13.4	3.79 ± 5.88
$\hat{f}_{AP}$	103 ± 152	89.0 ± 101	81.3 ± 81.8	$187 \pm 95.0^{\sharp}$	$179 \pm 105^{\natural}$
$\hat{f}_{SI}$	72.5 ± 159	56.7 ± 92.3	51.0 ± 70.3	163 ± 101	$130 \pm 86.4^{\natural}$
$\hat{f}_{ML}$	76.5 ± 154	60.9 ± 86.0	55.7 ± 62.5	$175 \pm 80.0^{\sharp}$	$143 \pm 77.2^{\natural}$
BW_{AP}	155 ± 109	$131 \pm 70.5^{\ddagger}$	130 ± 66.7	$293 \pm 66.1^{\sharp}$	$261 \pm 60.9^{\natural}$
BW_{SI}	127 ± 101	$101 \pm 57.3^{\ddagger}$	$287 \pm 73.8^{\dagger}$	$287 \pm 73.8^{\sharp}$	$236 \pm 68.3^{\natural}$
BW_{ML}	141 ± 98.6	$118 \pm 52.5^{\ddagger}$	121 ± 49.4	$308 \pm 58.0^{\sharp}$	$255 \pm 59.6^{\natural}$

Table 3. The effects of compressive sampling on typical wavelet-domain features. † denotes statistical differences between the original values and N/3 values. ♯ denotes statistical differences between the original values and N/4 values. ♮ denotes statistical differences between the original values and N/5 values.

	Original	N/2	N/3	N/4	N/5
θ_{AP}	0.54 ± 0.51	0.54 ± 0.51	0.54 ± 0.51	$0.79 \pm 0.78^{\sharp}$	$0.81 \pm 0.79^{\natural}$
θ_{SI}	0.44 ± 0.49	0.44 ± 0.49	0.44 ± 0.49	$0.64 \pm 0.70^{\sharp}$	0.61 ± 0.71
θ_{ML}	0.36 ± 0.44	0.36 ± 0.44	0.36 ± 0.44	$0.62 \pm 0.67^{\sharp}$	0.51 ± 0.63

6. CONCLUSIONS

In this paper, we demonstrated that we can estimate signal characteristics accurately even with a reduced number of samples. We have shown that with one third of the original samples, almost all considered features are identical to features extracted from original (fully sampled) signals. We have also determined that time domain features are very robust to the number of considered samples, and we can obtain very accurate estimates of these features even with 20% of the original samples.

ACKNOWLEDGMENTS

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. E. Sejdić, J. M. Dudik, A. Kurosu, I. Jestrović, and J. L. Coyle, "Understanding differences between healthy swallows and penetration-aspiration swallows via compressive sensing of tri-axial swallowing accelerometry signals," in *Proc. of SPIE Compressive Sensing III*, **9109**, pp. 91090M–1–91090M–6, (Baltimore, MD, USA), May 2014.

7. 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.

8. 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.

9. S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing* **41**, pp. 3397–3415, Dec. 1993.

10. 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.

11. G. Lof and J. Robbins, "Test-retest variability in normal swallowing," *Dysphagia* **4**(4), pp. 236–242, 1990.

12. 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.

13. 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.

14. P. E. Shrout and J. K. Fleiss, "Intraclass correlations: Uses in assessing rater reliability," *Psychological Bulletin* **86**, pp. 420–428, Mar. 1979.

15. J. M. Dudik, A. Kurosu, J. L. Coyle, and E. Sejdić, "A statistical analysis of cervical auscultation signals from adults with unsafe airway protection," *Journal of NeuroEngineering and Rehabilitation* **13**, pp. 7–1–10, 2016.

16. J. M. Dudik, I. Jestrović, B. Luan, J. L. Coyle, and E. Sejdić, "Characteristics of dry chin-tuck swallowing vibrations and sounds," *IEEE Transactions on Biomedical Engineering* **62**, pp. 2456–2464, Oct. 2015.

17. J. M. Dudik, J. L. Coyle, and E. Sejdić, "Dysphagia screening: Contributions of cervical auscultation signals and modern signal processing techniques," *IEEE Transactions on Human-Machine Systems* **45**, pp. 465–477, Aug. 2015.

18. 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.

19. 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.