

Reply to "Comment on 'Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior'"

Edward Sazonov^{1,5}, Stephanie Schuckers¹, Paulo Lopez-Meyer¹, Oleksandr Makeyev¹, Nadezhda Sazonova², Edward L. Melanson³, and Michael Neuman⁴

¹ Department of Electrical and Computer Engineering, Clarkson University, 8 Clarkson Ave, Potsdam, NY 13699 USA

² Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26505, USA

³ University of Colorado, Denver, School of Medicine, Division of Endocrinology, Metabolism and Diabetes, and the Center for Human Nutrition, Denver, CO 80262 USA

⁴ Department of Biomedical Engineering, Michigan Technological University, Houghton, MI 49931 USA

E-mail: esazonov@cias.clarkson.edu

Non-invasive monitoring of ingestion is an exciting and intriguing area of research. The worldwide growth of obese population still puzzles the researchers who cannot provide a definite answer to the etiology of this "obesity epidemic". Our work in the area is focused on developing compliant, non-invasive and easy-to-use devices for monitoring of ingestive behaviors that may shed a light on eating behaviors associated with obesity and be used in behavioral therapies. We want to know when each episode of ingestions takes place, how much was consumed and, preferably, determine type and caloric density of the food being consumed. From the first journal publication of Amft and Troster in February, 2008 [1] we were glad to learn that somebody else shares our approach to non-invasive monitoring of ingestion. We are excited to see the growth of the research area and hope that it will lead to better methods of non-invasive monitoring and help in fighting the obesity epidemic.

We thank Dr. Amft for his comments and would like to use this opportunity to respond to some of the key criticisms expressed in his letter. The goal of any introduction and literature review is to highlight the state of the art, rather than to point out every existing publication in the field and this means that it generally will not be possible to reference every work in the area. For example, our early conference publications on swallowing sound recognition [2-3] were not referenced by Dr. Amft in [1] which proposes such a methodology. Time is also a factor which cannot be discounted. The journal publication [1] was published almost at the same time as our paper (February vs. April of 2008), and thus it was not possible to include that publication in the review of prior work at the time the manuscript was written. The same can be said about the publications that came out several months later [4] or are not yet published [5]. We appreciate the opportunity to alert readers to earlier less-accessible conference publications [6-10] of Amft et al.

⁵ Corresponding author: Edward Sazonov, PhD

Our publication makes the first, very crucial step into the design of automatic methods for monitoring of ingestion, that is, establishing reliability of the primary data. Automatic scoring is an active area of research which has yet to mature. Amft et al. [6] reported an accuracy of 99% in eating recognition and 80%-100% in food identification. However, when a certain food type (rice) was not properly detected or identified by the proposed algorithm it was omitted from the reporting to show artificially high accuracy of recognition and identification. As another example, Amft and Troster [1] reported per-sample recognition rates of 86% for chewing and 70% for swallowing. These rates translate into lower rates of detecting actual masticatory sequences and instances of deglutition and make the methods proposed in [1] of limited practical application. It seems reasonable to suggest that without establishing the reliability of the manual score it is impossible to quantify what factor was responsible for such low recognition accuracy reported in [1]: the proposed recognition algorithm or the quality of the manual score. Thus, we believe it is necessary to base our accuracy claims for any automatic algorithm by first establishing the reliability of manual scores. We are actively pursuing automatic recognition of chewing and swallowing in order to address this challenging area of research.

With no viable methods for automatic scoring available at this time, we employed manual raters to mark each event of ingestion and create a “gold standard”. Utilization of a multi-modal system including a video stream allowed the raters to better estimate crucial parameters such as event boundaries. High reliability of the manual scores was also achieved by a well-constructed formal protocol that defined how every known situation should be scored. While size constraints limit the publishing of our scoring protocol, it does contain guidelines on correct utilization of the scoring software, description of data streams to be used as aid during the scoring process and definitions of bite, chewing and swallowing events. Every rater used this protocol both in training on a sequence of test subjects (not reported) and in actual scores (reported) in our publication. All of this resulted in a score of bites, chews and swallows with known statistical reliability measures established by analyzing a subset of over 10 hours (579 bites, 9719 chews and 1866 swallows) of data scored by three different raters. It should also be noted that the sample size for the intra-class coefficients analysis is defined by the number of epochs (1256 to 3770 for epoch duration ranging from 30s to 10s) rather than bites, chews and swallows which represent the measure that is being rated. The described steps ensured high accuracy of the manual score which in the future will result in higher accuracy of the training and validation datasets used for design of automated computer algorithms and will establish the upper bound on reliability of the automatic scoring results.

Selection of the sensor modalities for capturing of ingestion is as much art as science and should rely on objective analysis and proven performance. Performance of sensors working in conjunction with an automatic computer algorithm for identification of an event or quality of interest has to be validated on real data with the results compared to the reference score. For example, registering the tilt of the head by the chewing sensor may give additional clues when the intake of liquids is taking place, however this hypothesis would need to be confirmed by a statistically valid approach. This is the topic of our ongoing research.

Overall, we completely agree with Dr. Amft on the existence of huge unexplored possibilities of further research in the field on non-invasive monitoring of ingestion. There are numerous sensor modalities to be explored, both in detecting ingestion and estimating properties of the food; there is a need for developing methods that can use sensor information and provide users of such techniques with informative feedback and insight into eating behaviors. We hope that such methodologies will lead to better understanding the etiology of obesity and new emerging therapies for modification of eating behaviors.

References

- [1] Oliver Amft and Gerhard Troster. Recognition of dietary activity events using on-body sensors. *Artificial Intelligence in Medicine*, 42(2):121–136, February 2008.
- [2] Oleksandr Makeyev, Edward Sazonov, Stephanie Schuckers, Paulo Lopez-Meyer, Ed Melanson, Michael Neuman. Limited receptive area neural classifier for recognition of swallowing sounds using continuous wavelet transform, in *Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, EMBC 2007, Lyon, France, August 23-26, 2007, pp. 3128-3131.
- [3] Oleksandr Makeyev, Edward Sazonov, Stephanie Schuckers, Ed Melanson and Michael Neuman. Limited receptive area neural classifier for recognition of swallowing sounds using short-time Fourier transform, in *Proc. International Joint Conference on Neural Networks IJCNN 2007*, Orlando, USA, August 12-17, 2007, pp. 1417.1-1417.6.
- [4] Holger Junker, Oliver Amft, Paul Lukowicz, and Gerhard Troster. Gesture spotting with bodyworn inertial sensors to detect user activities. *Pattern Recognition*, 41(6):2010–2024, June 2008.
- [5] O. Amft and G. Troster. Automatic dietary monitoring: on-body sensing solutions for diet coaching. *IEEE Perv Comput*, in press.
- [6] Oliver Amft, Mathias Stager, Paul Lukowicz, and Gerhard Troster. Analysis of chewing sounds for dietary monitoring. In Michael Beigl, Stephen Intille, Jun Rekimoto, and Hideyuki Tokuda, editors, *UbiComp 2005: Proceedings of the 7th International Conference on Ubiquitous Computing.*, volume 3660 of *Lecture Notes in Computer Science*, pages 56–72. Springer Berlin, Heidelberg, September 2005.
- [7] Oliver Amft, Holger Junker, and Gerhard Troster. Detection of eating and drinking arm gestures using inertial body-worn sensors. In Bradley Rhodes and Kenji Mase, editors, *ISWC 2005: IEEE Proceedings of the Ninth International Symposium on Wearable Computers.*, pages 160–163. IEEE Press, October 2005.
- [8] Oliver Amft and Gerhard Troster. Methods for detection and classification of normal swallowing from muscle activation and sound. In Emile Aarts, Ryuji Kohno, Paul Lukowicz, and Jorge Carlos Trainini, editors, *PHC 2006: Proceedings of the First International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–10. ICST, IEEE digital library, November 2006.
- [9] Oliver Amft, Martin Kusserow, and Gerhard Troster. Probabilistic parsing of dietary activity events. In *BSN 2007: Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks*, volume 13, pages 242–247. IFMBE Proceedings, Springer, March 2007.
- [10] Oliver Amft, Martin Kusserow, and Gerhard Troster. Automatic identification of temporal sequences in chewing sounds. In Tony Hu, Ion Mandoiu, and Zoran Obradovic, editors, *BIBM 2007: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*, pages 194–201, San Jose, CA, USA, November 2007. IEEE Press.