

## A wearable earpad sensor for chewing monitoring

**Citation for published version (APA):**

Amft, O. D. (2010). A wearable earpad sensor for chewing monitoring. In *Proceedings of IEEE Sensors Conference 2010, 1-4 November 2010, Waikoloa, Hawaii* (pp. 222-227). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ICSENS.2010.5690449>

**DOI:**

[10.1109/ICSENS.2010.5690449](https://doi.org/10.1109/ICSENS.2010.5690449)

**Document status and date:**

Published: 01/01/2010

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# A Wearable Earpad Sensor for Chewing Monitoring

Oliver Amft<sup>\*,†</sup>

<sup>\*</sup>ACTLab, Signal Processing Systems, TU Eindhoven, The Netherlands

<sup>†</sup>Wearable Computing Lab., ETH Zurich, Switzerland

Email: amft@ieee.org

WWW: <http://www.actlab.ele.tue.nl>

**Abstract**—Today dietary assessments require manual information sampling in paper or electronic questionnaires on food type and other diet-related details. Low accuracies of 50% are confirmed for self-reporting, which weakens diet coaching effectiveness and is a major limitation for today's diet programs. Automatic Dietary Monitoring (ADM) using ubiquitous sensors was proposed to alleviate this problem.

In this work, we present implementation and analysis results of a novel acoustic earpad sensor device to capture air-conducted vibrations of food chewing. In contrast to previous works, our new device reduces ear occlusion compared to laboratory setups by using wearable earpad headphones. We investigate the sensing principle, perform a spectral sound analysis, and compare food classification performance to a classic lab-based sensor setup. We present novel food texture clustering results for 19 foods, spurring the understanding of food texture structure. In addition, we detail findings of a recent exhibition installation, where 375 food samples were analysed using the new sensor prototype.

**Keywords**—Dietary monitoring, chewing sound, mastication, food intake.

## I. INTRODUCTION

Food selection and food consumption patterns are vital elements for describing an individual's dietary behaviour. Monitoring dietary behaviour is a prerequisite for effective diagnosis, prevention, and intervention [1]. As an example, most weight management programs use some form of dietary behaviour monitoring to support personal food selection and consumption behaviour coaching. These monitoring assessments have a long tradition in questionnaires, where respondents are asked to either recall consumptions, immediately annotate food intake, or describe typical consumption frequencies [2], [3]. Recalls often address the last 24 hours. The scope of frequency questionnaires is in weeks to months. In contrast to those, the method of immediate annotation intends to acquire actual intake timing and details on food type. Modern versions of this assessment use PDAs and mobile phones to let users record their intake information [4]. Such computerised questionnaires can relieve the diet counsellor from digitising and interpreting diaries. However, both paper and computerised diet questionnaires are known to achieve low accuracies in food intake reporting [3], [5]. Often accuracy was found to be hampered by the respondent's motivation, memorising, literate capabilities, or influenced by changing perceptions of desirability and increasing self-awareness due to the reporting itself [6]. Thus, approaches that could alleviate the respondent from the manual annotation task could provide

a substantial improvement over questionnaires.

Novel ubiquitous sensing and computing technologies could enable more accurate and robust assessments of dietary behaviour. Specifically, monitoring systems that could acquire user behaviour without manually recording of food intake, an approach termed Automatic Dietary Monitoring (ADM) could lead to novel assessments [7], [8]. One essential element of ADM-based systems are unobtrusive sensors that could continuously acquire dietary behaviour by observing users. Due to complexity and variability of dietary behaviour, food consumption can hardly be captured by a single sensing approach. However, individual sensing concepts can provide important information on dimensions of dietary behaviour [9], such as on the daily intake schedule and food type. In particular, measuring vibrations during food chewing could become a vital sensing concept for ADM systems. This sensing concept had been introduced previously [7], however its potential to identify different foods and food categories is not yet sufficiently understood.

This paper investigates chewing sound using a novel headphone-housed acoustic transducer to capture vibration and study signal patterns under minimal ear occlusion conditions. While a high ear occlusion helps to maintain good signal quality, a reduced occlusion is expected to result in a lower SNR and thus reduced recognition performance. However, this earpad design allows the sensor device to be used continuously and in natural non-stationary living environments.

This paper provides the following contributions to confirm viability of the chewing sensor prototype:

- 1) The sensing principle is discussed and analysed from spectral properties of different foods. This analysis is relevant for sensor design considerations and selection of features according to their spectral bandwidth.
- 2) A food type clustering approach is presented to analyse vibration pattern similarities among 19 foods and study acoustic relations of foods. This investigation provides insight into the feature-based grouping of food textures.
- 3) Using the headphone-housed sensor, a food recognition system was developed and evaluated in a museum. The prototype implementation and food classification results of this system are presented.

## II. ACOUSTIC FOOD INTAKE MONITORING

During chewing strokes (jaw opening and closing) food pieces (bites) are crushed between teeth as a first step in

food decomposition. During this breakdown process, vibrations are generated and conducted through teeth, mandible, and the skull [10]. This breakdown is partly audible since the crushing vibrations are emitted through air conduction as well. However, the transmission path through bones is more robust against environmental noises [7]. The ear canal forms a natural cavity where the propagated vibrations are eventually audible. Acoustic transducers can be used to capture the food breakdown sound in close proximity of the ear and ear canal [11].

Sound modulation and energy during the breakdown depend on the inner structure of the food material and its textural properties. Cell arrangement, impurities, and cracks were expected to influence sound production [12]. According to the cellular structure many foods can be categorised into dry and wet-crisp groups. Dry-crisp foods have air inclusions in their structure, such as in potato chips, whereas wet-crisp foods contain fluid compartments, as e.g. in apples [13]. However, many food products exist, where these categories do not fully apply. Thus, an important goal of this work is to determine the robustness in grouping different foods according to dry and wet-crisp categories, as well as to investigate the unsupervised grouping of foods.

Besides the vibration sensing concept, several further approaches to chewing monitoring exist. Jaw opening and closing during food breakdown in the mouth can be sensed from masseter and temporalis muscle activations using surface Electromyography (EMG) and jaw-attached movement sensors [14]. Since electrodes and sensors are exposed in facial regions, those measurement techniques are too obtrusive to be applied in natural environments. Although head and hyoid motion during chewing were observed to correlate with food piece size and food hardness [15], no robust relation and signal interpretation outside laboratories had been found yet. Dental implants were used to assess load during chewing with strain gauges [16], however these oral implants can be expected to alter oral sensation and may be infeasible for long-term food intake monitoring.

Initial recordings of acoustics during chewing were performed by Drake [11] for consuming crisp and hard foods. Subsequent studies focused on relating chewing sounds to sensory perception of foods and to food assessments, e.g. in [17]. Attempts to classify foods using pattern recognition techniques were initially performed by DeBelie et al. [18] and in the scope of ADM in [7]. These latter works showed that a low number of foods (below ten) can be classified in a laboratory setting using foam-based ear sensors, which result in high ear occlusion.

Moreover, these investigations found that chewing sound patterns changed during the breakdown process of several chewing cycles. This observation was later confirmed using an automatic unsupervised sequence searching technique to group chewing strokes [19]. In contrast, the current work aims to establish the viability of a reduced occlusion sensor prototype.

### A. Potential of the ADM approach

The development of intelligent monitoring technologies that support healthy dieting is a promising research topic in the field of Ambient Intelligence and Ubiquitous Computing. ADM systems could permanently accompany patients in dietary and weight management programs. They may be used to derive eating behaviour details during daily activities as well as providing personalised feedback and coaching assistance. Eventually, ADM systems could complement questionnaire-based assessments and replace them.

Designing effective dietary monitoring solutions and ubiquitously implementing intelligent monitoring assistants is nevertheless challenging, as eating behaviour patterns are very diverse and frequently changing even for one particular individual. Such modifying circumstances include different locations, e.g. eating in a restaurant, at home, or while commuting, food consumptions in social groups, as well as eating under different emotional conditions.

In the ADM concept, various sensing options are considered to acquire a rich set of relevant information details. These include intake motion, which could be used to distinguish food categories (such as using fork and knife or a spoon), chewing sounds, as discussed in this work, and swallowing, which relates to intake schedule and food consistency (as swallowing rate is increased during meals). While for these sensing concepts early research prototypes were realised they are not yet comfortable enough for long-term (months and years) of continuous use, which is needed to achieve sustained behavioural changes [20]. The approaches nevertheless highlight profound benefits of the ADM concept. Manual diaries are no viable option due to their required effort and errors. Hence any type of automatically derived information on eating behaviour can be valuable for diet coaching. Further discussions and a taxonomy of the various diet monitoring approaches can be found in [9].

## III. EARPAD SENSOR DESIGN AND SIGNAL ANALYSIS

A wearable acoustic sensor was developed consisting of a commercial off-the-shelf headphone housing and an embedded transducer. The acoustic properties of different foods were analysed to establish a basis for optimising sensor design and subsequent signal processing steps.

### A. Sensor design

To implement the earpad sensor, an off-the-shelf headphone housing was used. The housing could be attached to the ear canal using soft foam cushions. Speaker and cabling were removed from the original headphone to obtain a housing. As acoustic transducer a Knowles omnidirectional miniature electret microphone (type: FG23329) was used. This model offered a constant sensitivity across the frequency range of 100 Hz to 10 kHz, where the relevant signal information was expected.

The transducer was embedded into the headphone housing using a modelling material compatible with human skin. Alignment of transducer and housing was made such that the

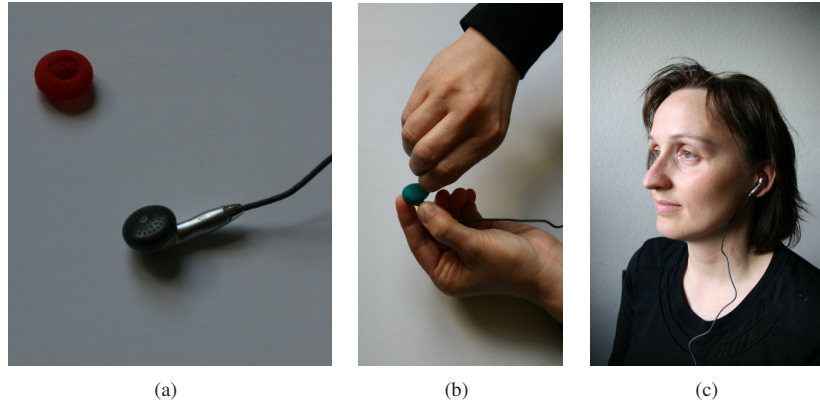


Fig. 1. Earpad sensor design to monitor chewing sounds. (a): Sensor prototype with earpad cushion. (b): Attachment of the foam cushion to the sensor. (c): Sensor prototype when worn by a test user.

active sensing area faced centrally outward of the housing. Thus, when worn this construction could capture sounds propagating in the ear canal.

For wearing comfort and hygienic reasons, a foam cushion could be attached to the final sensor design and replaced by users. The sensor was amplified and sampled at 44.1 kHz with a resolution of 16 bit for all investigations in this work. Figure 1 shows the sensor design, cushion attachment, and the worn prototype.

### B. Spectral signal analysis

A spectral analysis was performed for sounds from different food products. One subject consumed four foods while wearing an earpad sensor and a lab reference probe concurrently, one at each ear. The recordings were made in a laboratory environment with low environmental noise while the room was not damped. Lab reference comprised the same transducer integrated into a foam enclosure. Thus, this lab reference achieved a higher ear occlusion level compared to the earpad sensor.

For each food, one representative chewing stroke from the first 20% of chewing sequences was selected for this analysis. The first stroke of a sequence was not considered, since it may have involved biting instead of chewing. Selected strokes had a duration of  $\sim 500$  ms. Spectra were derived using a 512-point FFT and a Hanning window. Sound pressure was obtained in dB. Figure 2 shows spectra for all considered foods.

Overall signal power dropped below relevant pressure levels ( $< -90$  dB) at a frequency of  $\sim 8$  kHz, where a drop in power occurred for the earpad sensor. The effect of a higher occlusion was observed for the lab reference, which captured pressure levels above  $-90$  dB up to 16 kHz, in particular for chewing chocolate chunks. For apple slices and lettuce power differences between earpad sensor and lab reference were lower, which could be explained by an overall lower sound pressure in those recordings. Based on this analysis, it can be concluded that frequencies below 8 kHz should be primarily considered in chewing sound analyses. The band between 8 kHz and 16 kHz could be relevant for specific foods, however

it was not captured by the earpad sensor. Besides the lower occlusion, a foam cushion may have damped signals in the earpad sensor.

## IV. FOOD CATEGORY CLUSTERING

An unsupervised clustering approach was investigated to identify feature-based groupings among food products. This section details approach and grouping results.

### A. Unsupervised food grouping method

For the cluster analysis, chewing sound sequences from three male individuals (aged between 25 and 30) were recorded in a laboratory environment. In total 19 foods were considered in this analysis. Foods were selected such that the set included frequently consumed foods of participants, while at the same time maintain wide food texture diversity. Since participants were used to these foods, no aversion effects were expected.

From all continuous chewing sounds, 232 spectral features were derived in sliding windows of 512 samples size without overlap. The feature set included subsets of Linear Predictive Coefficients, Mel-frequency cepstral coefficients, auto-regressive coefficients, and further features derived from statistical models. Features were subsequently averaged over multiple sliding windows in 0.5 s intervals. From each chewing sequence, the first 30% were considered for this analysis to maintain stationary food and sound properties.

A Fisher discriminant analysis was applied on the raw feature set to condensed the set. For the Fisher discriminant filtering, categories were chosen according to individual foods. Subsequently, a hierarchical agglomerative clustering approach was used as unsupervised exploration approach. With this approach, food instances were grouped bottom-up, so that individual foods would fall together if they obtain similar feature values. To measure distances between food feature vectors, the Ward distance metric was used:

$$d_{Ward}(D_i, D_j) = \sqrt{\frac{n_i n_j}{n_i + n_j}} \|\mathbf{m}_i - \mathbf{m}_j\|_2, \quad (1)$$

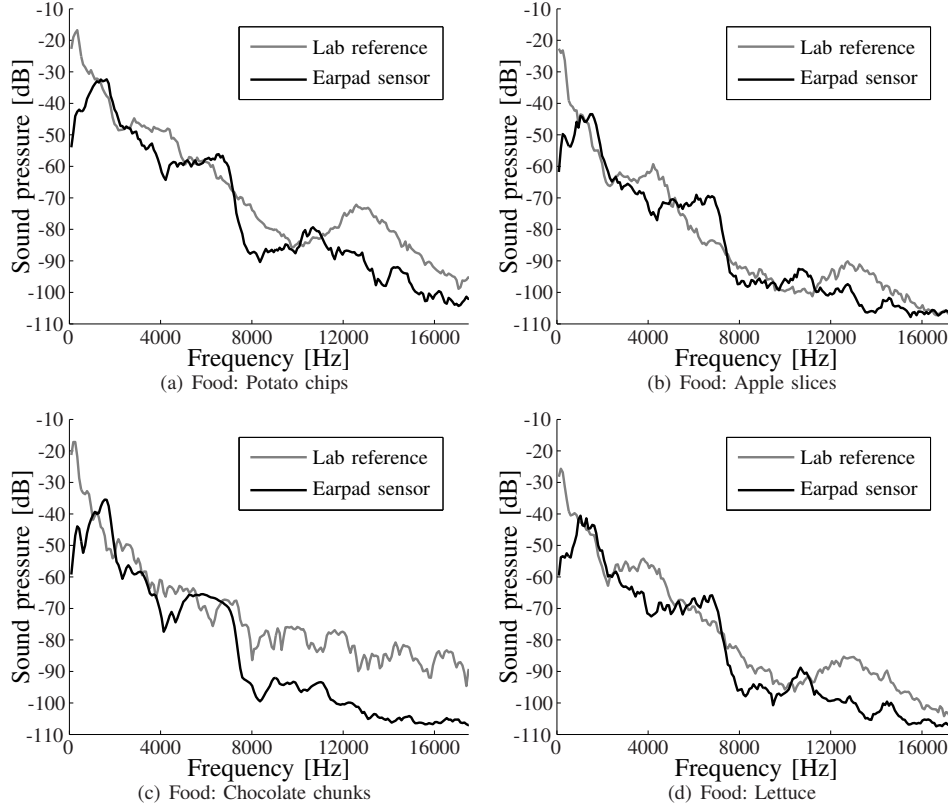


Fig. 2. Spectral analysis of the earpad sensor and a lab reference probe for four foods. While the earpad sensor is embedded in a headphone housing, the lab reference is using a foam enclosure that achieved a higher ear occlusion level.

where  $D_i, D_j$  corresponds to instances of clusters  $i$  and  $j$ ,  $\mathbf{m}_i$  and  $\mathbf{m}_j$  are cluster centroids, and  $n_i, n_j$  are instance counts of each cluster.  $\|\cdot\|_2$  denotes the Euclidean distance. Cluster centroids  $\mathbf{m}_i$  and  $\mathbf{m}_j$  were derived according to  $\mathbf{m}_i = \frac{1}{n_i} \sum_{x \in C_i} x$ , where  $x$  is an individual instance of the cluster set  $C_i$ .

### B. Hierarchical food clustering results

The food grouping analysis revealed three most prominent clusters. These clusters could be interpreted based on food properties perceived by subjects as “wet, loud”, “dry, loud”, and “soft, quiet”. Figure 3 shows a dendrogram of the hierarchical clustering result. By using the Fisher discriminant feature filtering pure initialisation clusters were obtained with one food dominating each cluster. Purity is indicated in percent at the vertical dendrogram axis.

The vertical line marks a subjectively selected grouping that offers an interpretable result. At this Euclidean distance level, foods can be attributed according to their origin or manufacturing process. For example, apple and lettuce are grouped, relating to their similar wet-crisp cell structure. The global clustering result into “wet, loud”, “dry, loud”, and “soft, quiet” shows that the initial assumption of a texture-based

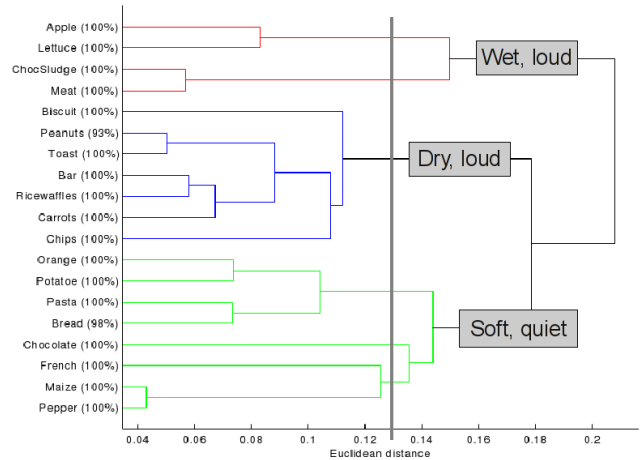


Fig. 3. Dendrogram of the hierarchical food clustering approach. Purity of clusters is shown at the vertical axis. Vertical line marks a subjectively selected interpretable grouping (see main text for more details). Overall, three groups can be identified, of “wet, loud”, “dry, loud”, and “soft, quiet” foods.

structure was confirmed. The “soft, quiet” cluster includes a mixture of foods that could not be further resolved with this approach.

## V. EARPAD SENSOR VALIDATION

The earpad sensor design was validated in a exhibition installation with museum visitors. In this section, system deployment, and food classification performance results are detailed that illustrate the exhibition system’s functionality.

### A. Exhibition installation

A food recognition system was developed with the aim to showcase the earpad sensor function in a public experiment. Participants could chew samples of different self-selected foods, while a computer-based classification algorithm identified the food type from acoustic patterns. Food recognition results were subsequently shown to the participant. The system was initially installed in the museum Alimentarium at Vevey, Switzerland.

The food recognition system comprised an Eeebox computer, a sensor acquisition board, a table with an integrated touchscreen, and the earpad sensor. The computer ran a specifically developed wizard software that guided participants through the sensor attachment, food selection, recognition results, and presented additional explanations of the system’s purpose. Figure 4 shows the exhibition installation.

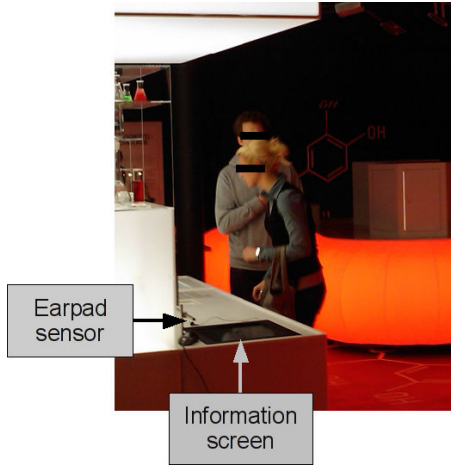


Fig. 4. Exhibition installation using the earpad sensor design to monitor chewing sounds. Exhibition visitors were invited to participate in an experiment where they could freely chose and consume a food piece from a given food selection. A recognition system captured chewing sounds using the earpad sensor and presented the system’s food classification result.

Food samples from four different foods were offered to the users: apple slices, chocolate pieces, potato chips, and carrot chunks. Pattern models for each food type were developed before the system deployment.

During several exhibition months, the recognition system needed several refinement steps to cope with varying environmental noise and to improve mechanical robustness due to the numerous users. Nevertheless, the exhibition confirmed that the earpad sensor design is viable to be used continuously.

### B. Food classification performance

The food recognition system shown in the exhibition was developed based on a dedicated dataset of chewing sound

recordings of two subjects and varying environmental conditions. In total 375 chewing sequences were recorded using a recording setup that matches the exhibition system. About 10% of these recordings were made in the exhibition environment and the rest was acquired in the lab with different background noises.

From these continuous chewing sounds features were derived in sliding windows and averaged in 0.5s intervals as detailed in Section IV before. Nevertheless, the complete chewing sequence was considered in this classification analysis. A combination of a Fisher discriminant filter and a naïve Bayes classifier was used to perform feature reduction and food classification respectively. For each instance (chewing sequence) a majority vote from all classifications was derived as final class decision.

The performance of this system was evaluated using a leave-one-out cross-validation scheme. In each fold all but one chewing sequence were used as instances for acquiring food models and the left out sample was used for testing. For all derived pattern models the dataset was balanced between classes such that no class skew persisted. For classes that had more instances available, a random selection was made.

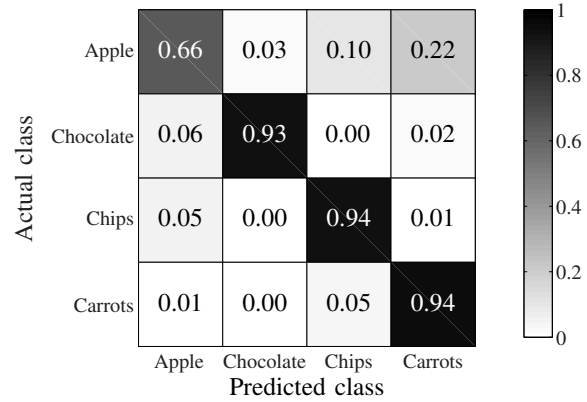


Fig. 5. Recognition performance analysis of the exhibition system showing classifier accuracies and the confusion matrix. The evaluation was performed using a leave-one-out cross-validation scheme. Overall a recognition accuracy of 86.6% was obtained.

Figure 5 shows the classifier confusion matrix for the above described setup and methodology. Overall, the system achieved a recognition performance of 86.6%. This result confirms that the system performed sufficiently well to be used in the exhibition. Moreover, this result encourages further investigations on real-life implementations of the earpad sensor design.

## VI. CONCLUSION AND FURTHER WORK

The new earpad sensor design and implementation showed promising results, both in laboratory evaluation as well as in a practical deployment. Although the construction using headphone housing and foam cushion reduce SNR, recognition results indicate that the concept could function sufficiently accurate. Further work will be performed to study additional

real-life deployment scenarios that could reveal long-term system robustness.

The food clustering analysis showed that dry and wet food categories are prominently represented in the considered feature set. Moreover, it became clear that signal energy is a critical element in clustering foods. A separate cluster appeared, which contained foods that had low acoustic energy content when being chewed. In the presence of loud foods, it is challenging to discern food properties in this cluster. Further analysis steps will be directed to isolate acoustic properties to confirm this grouping. In the current analysis the inclusion of complete chewing sequences and an arbitrary selection of the first 30% of each sequence could influence clustering results.

#### ACKNOWLEDGEMENT

The author would like to thank the Alimentaryum museum coordinator, Dr. Esther Schärler-Züblin and her staff for their support in setting up an early version for the food recognition prototype in the museum. This work was supported in part by project grants of Nestle Foundation, Switzerland, and the 3TU federation NIRICT, The Netherlands.

#### REFERENCES

- [1] F. Bellisle, "Why should we study human food intake behaviour?" *Nutr Metab Cardiovasc Dis*, vol. 13, no. 4, pp. 189–193, 2003.
- [2] B. S. Burke, "The dietary history as a tool in research," *J Am Diet Assoc*, vol. 23, pp. 1041–1046, 1947.
- [3] J. C. Witschi, "Short-term dietary recall and recording methods," in *Nutritional Epidemiology*. Oxford University Press, 1990, vol. 4, pp. 52–68.
- [4] L. E. Burke, M. Warziski, T. Starrett, J. Choo, E. Music, S. Sereika, S. Stark, and M. A. Sevick, "Self-monitoring dietary intake: current and future practices." *J Ren Nutr*, vol. 15, no. 3, pp. 281–290, 2005.
- [5] J. Beasley, "The pros and cons of using pdas for dietary self-monitoring." *J Am Diet Assoc*, vol. 107, no. 5, p. 739, 2007.
- [6] R. J. Hill and P. S. Davies, "The validity of self-reported energy intake as determined using the doubly labelled water technique." *Br J Nutr*, vol. 85, no. 4, pp. 415–430, 2001.
- [7] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster, "Analysis of chewing sounds for dietary monitoring," in *UbiComp 2005: Proceedings of the 7th International Conference on Ubiquitous Computing*, ser. Lecture Notes in Computer Science, vol. 3660. Springer Berlin, Heidelberg, 2005, pp. 56–72.
- [8] O. Amft and G. Tröster, "On-body sensing solutions for automatic dietary monitoring," *IEEE Perv Comput*, vol. 8, no. 2, pp. 62–70, 2009.
- [9] O. Amft, "Ambient, on-body, and implantable monitoring technologies to assess dietary behaviour," in *International Handbook of Behavior, Diet and Nutrition*. Springer, 2011, p. in press.
- [10] P. J. Lillford, "The materials science of eating and food breakdown," *MRS Bulletin*, vol. 25, no. 12, pp. 38–43, 2000.
- [11] B. Drake, "Food crushing sounds. an introductory study," *J Food Sci*, vol. 28, no. 2, pp. 233–241, 1963.
- [12] W. AlChakra, K. Allaf, and A. Jemai, "Characterization of brittle food products: Application of the acoustical emission method," *J Tex Stud*, vol. 27, no. 3, pp. 327–348, 1996.
- [13] J. Edmister and Z. Vickers, "Instrumental acoustical measures of crispness in foods," *J Tex Stud*, vol. 16, no. 2, pp. 153–167, 1985.
- [14] J. Dahan and C. Boitte, "Comparison of the reproducibility of emg signals recorded from human masseter and lateral pterygoid muscles." *J Dent Res*, vol. 65, no. 3, pp. 441–447, 1986.
- [15] B. Häggman-Henrikson and P.-O. Eriksson, "Head movements during chewing: relation to size and texture of bolus." *J Dent Res*, vol. 83, no. 11, pp. 864–868, 2004.
- [16] E. Stellar and E. E. Shrager, "Chews and swallows and the microstructure of eating." *Am J Clin Nutr*, vol. 42, no. 5 Suppl, pp. 973–982, 1985.
- [17] Z. M. Vickers, "Food sounds: How much information do they contain?" *J Food Sci*, vol. 45, no. 6, pp. 1494–1496, 1980.
- [18] N. DeBelie, V. De Smedt, and D. B. J., "Principal component analysis of chewing sounds to detect differences in apple crispness," *Postharvest Biol Technol*, vol. 18, pp. 109–119, 2000.
- [19] O. Amft, M. Kusserow, and G. Tröster, "Automatic identification of temporal sequences in chewing sounds," in *BIBM 2007: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*. IEEE Press, 2007, pp. 194–201.
- [20] R. R. Wing and S. Phelan, "Long-term weight loss maintenance." *Am J Clin Nutr*, vol. 82, no. 1 Suppl, pp. 222S–225S, 2005.