A novel approach for chewing detection based on a wearable PPG sensor

Vasileios Papapanagiotou[1], Christos Diou[1], Lingchuan Zhou[2]
Janet van den Boer[3], Monica Mars[3], Anastasios Delopoulos[1]

***Abstract*— Monitoring of human eating behaviour has been attracting interest over the last few years, as a means to a healthy lifestyle, but also due to its association with serious health conditions, such as eating disorders and obesity. Use of self-reports and other non-automated means of monitoring have been found to be unreliable, compared to the use of wearable sensors. Various modalities have been reported, such as acoustic signal from ear-worn microphones, or signal from wearable strain sensors. In this work, we introduce a new sensor for the task of chewing detection, based on a novel photoplethysmography (PPG) sensor placed on the outer earlobe to perform the task. We also present a processing pipeline that includes two chewing detection algorithms from literature and one new algorithm, to process the captured PPG signal, and present their effectiveness. Experiments are performed on an annotated dataset recorded from 21 individuals, including more than 10 hours of eating and non-eating activities. Results show that the PPG sensor can be successfully used to support dietary monitoring.**

I. Introduction

Obesity (OB) is a major public health problem globally. The latest pharmacological interventions are failing [1] and surgical procedures, while more successful [2], are highly invasive, riddled with undesirable side effects and not as relevant as prevention strategies [3]. As an alternative, monitoring and modification of dietary behaviour has shown to be a significantly more promising approach for the treatment of OB [4], while behavioural interventions have been successful in clinical environments [5]. In this context, a critical indicator of risk is the frequent occurrence of snacking events or, in more extreme cases, "continuous eating".

To objectively measure this behaviour, various approaches exist; Amft et al. introduce a method for detecting chewing sounds in [6], with promising results. In their work, a common air microphone is housed in commercial earphones, since this position is found to yield the best results. A comparison of seven chewing detection algorithms is performed in [7]. The algorithms are evaluated on a large dataset of 51 subjects, where a total of 6 different food types were consumed. It is important to note however that microphones require high sampling rates (44.1 kHz in [6] and 11 kHz in [7]); this significantly increases both power and processing requirements. More recently, in [8] the Fractal Dimension of chewing sounds is examined, and found to be a highly discriminative attribute, compared to other non-chewing related sounds captured by such microphones. Experiments with resampled versions of the acoustic signal showed that a low sampling frequency of 2 kHz is sufficient.

Alternative systems for chewing detection have also been explored. In [9] a jaw motion sensor is realised using a piezoelectric strain gauge sensor. Feature selection is performed and inter and intra subject classification experiments using Support Vector Machines yield high accuracy of 81%. Further use of strain sensors for chewing detection are presented in [10].

In this work, we introduce a photoplethysmography (PPG) sensor housed in a prototype chewing sensor that mounts on the subject's outer ear. The sensor sampling frequency is approximately 20 Hz, significantly lower than other systems; furthermore, contrary to audio sensors, it is not affected by talking, ambient noise, etc, making it a more robust means to detect chewing. The sensor hardware and signal processing methods are presented in Section II. The experimental setup and evaluation metrics are described in Section III. Section IV presents the results, and Section V concludes the paper.

II. The PPG Chewing Sensor

Food intake involves the mastication (chewing) of small pieces of food. During chewing, four muscles are used to move the jaw and crush the food, progressively transforming it into a bolus. These muscle movements trigger various activations, including blood flow variations in areas close to temporalis and ear, which have long been reported (e.g. in [11]). Furthermore, the close relationship between the ear and the jaw during embryonic development results in the same nerve, the trigeminal nerve, controlling both the tensor tympani and the chewing muscles. Consequently, signals which are sent through the trigeminal nerve can affect both the jaw and the ear muscles. The two joints which attach the jaw to the skull are located just in front of the ears. These muscle activations can be captured using EMG placed at appropriate locations in order to detect chewing; we have used this technique to generate ground truth as described in Section III, however it is hard to build cheap and inconspicuous EMG sensors for everyday dietary monitoring.

In this paper, we study and propose the use of PPG to capture the jaw and ear muscle motion during chewing. PPG is an optical measurement method, which is widely used to measure perfusion via pulse oximetry (SpO_2) [12]. It has lately been applied in applications such as heart-rate monitoring in wearables [13], where the change in volume caused by the pressure pulse is detected by two

*The work leading to these results has received funding from the European Community's ICT Programme SPLENDID under Grant Agreement No. 610746, 01/10/2013 - 30/09/2016 (`splendid-program.eu`).

[1]Multimedia Understanding Group, Information Processing Laboratory, Aristotle University of Thessaloniki, Greece
[2]CSEM SA, Switzerland
[3]Wageningen University, Netherlands

Light Emitting Diodes (LED) with different wave lengths (in general a red and an infrared).

Based on the physiology of chewing described above, placing the PPG sensor in the ear concha seems like a natural choice; it is close to the temporalis and can effectively capture blood variations caused by activations of chewing muscles, it is non-intrusive, and it is also possible to combine a PPG sensor with a microphone, to further increase the effectiveness of the monitoring system. Traditionally, measurement of heart rate or blood oxygenation by PPG is highly sensitive to movement artefacts; in our application, the signal of interest is in fact the signal captured due to movement caused by mastication.

A. *Hardware of the PPG sensor*

For chewing detection with PPG, we are not interested in measuring the SpO_2. Therefore only one LED is used instead of two. The sensor operates with the ear skin separating the LED and the photodiode (see Fig. 1); thus, longer wave lengths increase the transmission capability given a certain current exerting the LED. We have chosen an infrared (about 950 nm) instead of a green (525 nm).

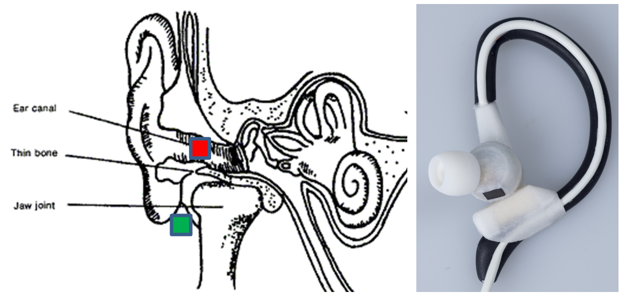
The proposed PPG chewing sensor (Fig. 1b) therefore includes only one LED which lightens the skin and a photodiode which detects the tiny modulations of light intensity caused by pulsating blood flow in the tissue. As shown in Fig. 1a, the LED is placed at the bottom of the ear and the photo-diode is inserted in the ear canal. The PPG sensor has been successfully embedded into an earphone-like housing. Our aim is to use the ear hook of the commercial off-the-shelf earphone NB439B from New Balance and design a customised housing in which the optical PPG sensor is integrated.

The signal acquired by the photo-detector is pre-amplified, filtered, and digitised at a fixed sampling frequency of $64/3$ Hz (the AD convert is clocked at 64 Hz and alternatively samples 3 channels). In order to avoid signal saturation due to high density of ambient light, a proprietary compensation technique is applied. Fig. 2 shows the system block diagram.

The manufacturing by rapid prototyping has been carried out by the Swiss company Von Allemen. The main criterion of selection for the manufacturer is the bio-compatibility aspect. Indeed, since the sensor has to be in contact with the skin, it is essential to use a material that does not cause any irritation or other adverse effect on the skin. A material, called DuraForm PA Powder which is FDA and USP class VI certified was used. The fabrication process was Selective Laser Sintering, an additive manufacturing technique that uses a laser as the power source to sinter powdered material.

B. *PPG signal processing*

We incorporate a three-stage processing pipeline for the PPG signal. A sequence of raw signal samples of the PPG chewing sensor is presented in Fig. 3. The very low frequency components are removed at the first, pre-processing stage using a high-pass FIR filter with cut-off frequency around 0.5 Hz. A segment of the filtered signal is shown



(a) (b)

Fig. 1: The prototype PPG chewing sensor 1b and its LED transmitter and receiver placement 1a.

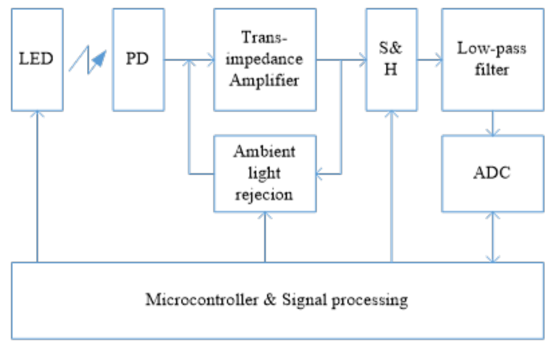


Fig. 2: The PPG chewing sensor block diagram

in Fig. 4. Additionally, signal amplitude is normalised based on the control signals received from the amplifier, as the amplifier automatically adjusts the level of amplification based on external lighting conditions.

In the second stage, chewing activity detection is performed. Three algorithms are tested for the detection task. The first one, called Maximum Sound/Signal Energy (MSEA), is presented in [7] where it is originally used on audio signals. The signal is divided into consecutive, non-overlapping frames and the signal energy $E[n]$ is computed within each frame. Chewing events are detected when the signal energy reaches a maximum value and at the same time exceeds a predefined threshold E_t. A chewing event is registered only if the maximum is not exceeded by λ subsequent frames. In particular, a chew is detected for the n-th frame if the following expression holds

$$E[n] > E_t \text{ and } E[n] > E[n+i] \text{ for } i = 1, \ldots, \lambda \qquad (1)$$

An important drawback of the algorithm when applied to audio signals is the confusion of chewing peaks with peaks found in speech signals. This drawback however is completely eliminated as the PPG sensor does not capture talking.

The second algorithm, called Low-Pass Filtering Algorithm (LPFA), is presented in [14] and is also applied to audio signal. A band-pass filter with a very narrow frequency range (close to the chewing frequency of 1.5 to 2.5 Hz) is applied to a rectified version of the audio signal. For the case of PPG, no rectification is required since PPG signals exhibit much smoother behaviour compared to audio. By detecting peaks in the filtered version of the signal, the algorithm is able to identify individual chews. The ideal chewing frequency can be determined in a subject dependent way [14] or set to a constant value [7].

The third algorithm, called Chewing-Band Power Algorithm (CBPA), computes the time-varying spectrum of the signal, based on Welch's method, i.e. by taking the ensemble average of the FFT over several overlapping windows for each frame. In particular, let $X_n[k]$ denote the DFT coefficients for $k = 1, 2, \dots, N$ for the n-the window. We select a number of $2q+1$ windows to perform spectrum estimation using

$$S_n[k] = \frac{1}{2q+1} \sum_{i=-q}^{q} \|X_{n+i}[k]\|^2 \quad (2)$$

The chewing band energy for the n-th window is calculated as the sum of $S_n[k]$ for those values of k that correspond to frequencies in the 1.1 - 2.5 Hz band; it is then used as a criterion for decision, based on a threshold. The threshold can easily be estimated based on the overall signal power over an extended time window. Given the chewing regions, one can detect individual chews by selecting the maxima of the chewing signal in the region.

Regardless of the choice of the main processing algorithm, a series of detected chew activations are available to the pipeline, in the form of a binary signal $b(n)$. Median filtering is then applied to $b(n)$ to remove some scarce, false positive activations, yielding the filtered signal $b'(n)$ which is a signal of pulses (Fig. 5). These pulses are in fact the detected chews, and are directly compared and evaluated with ground truth chews, as detailed in Section IV.

Additionally, the third stage integrates chews to chewing bouts, and subsequently chewing bouts into snacking events (snacks). Chews are integrated into bouts by first applying a filter to the time-series of chews that removes isolated chews. To this end, we define a chew density by a maximum duration of l seconds (in our case 10) and a minimum number of chews m (in our case 8) for that duration. For every chew that we can find a time window of at most l seconds that includes at least m chews (including the current one), we retain it; for each chew that we can't find such a window, we discard it. This essentially removes some false-positive detections that can occur due to abrupt environmental light changes. Then, these dense sequences of chews are directly aggregated into chewing bouts. Finally, nearby bouts are merged into a snack if they are no more than l' minutes (in our case 2) apart.

III. Experimental Setup

The prototype chewing sensor along with the processing pipeline were evaluated on a dataset recorded at the Wageningen University, Netherlands, in the framework of EU funded program SPLENDID [15]. It contains recordings of 21 individuals wearing the prototype PPG sensor connected via wire to a prototype data-logger. Various activities were performed by each subject in randomised sequences, including pauses, talking, listening to another person speaking, coughing, and consumption of a variety of different foods and liquids, such as apple, lettuce, potato chips, toffee, water, milk, etc. The recording for each subject lasts approximately 30 minutes. The dataset is an expanded version of the one presented in [8].

Ground truth is available from diaries recorded by experts supervising the experiment, as well as from Surface EMG (sEMG) signals from sensors placed on the points of the face of each subject. The sEMG signals are processed so as to generate detailed ground truth chews. We use a modified version of the standard technique presented in [16], extended with linear prediction from auxiliary sEMG channels. The chewing regions are then aggregated based on a combination of the recorded diaries, as well as the integration approach presented in Section II-B (third stage), to yield the ground truth for bouts and snacks.

IV. Evaluation & Results

Evaluation is performed at three different levels, and evaluation results are presented in Table I. At the first level, detected chews are evaluated directly, using the sEMG ground truth chews. A simple matching technique is employed that

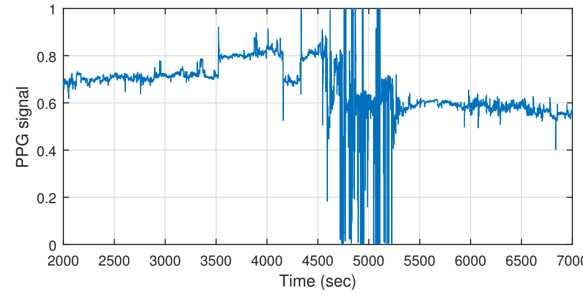


Fig. 3: Raw signal from the PPG chewing sensor. Abrupt changes are caused by the adaptive amplifier.

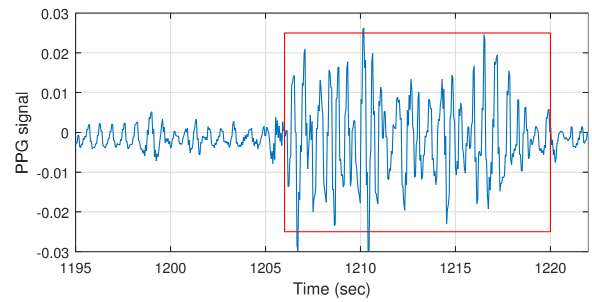


Fig. 4: High-pass filtered signal from the PPG chewing sensor. The red box indicates chewing activity.

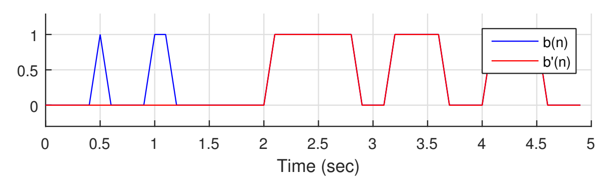


Fig. 5: The boolean signal $b(n)$ that is the output of the second stage of the pipeline, and the median filtered $b'(n)$.

TABLE I: Precision and Recall for detecting chews, bouts, and snacks. Algorithms: Maximum signal energy (A), Low-pass filtering (B), Chewing band power (C).

	Chews		**Bouts**		**Snacks**	
	Prec.	Recall	Prec.	Recall	Prec.	Recall
A	**0.706**	0.395	**0.862**	0.653	**0.928**	0.916
B	0.534	**0.763**	0.731	**0.993**	0.640	**0.987**
C	0.634	0.565	0.855	0.730	0.824	0.968

allows small drifts for the start and end time-stamps of each chew. From the detected chews, the ones that are matched with ground truth chews are labelled as true positives (TP), and the remaining as false positives (FP). Finally, the non-matched ground truth chews are labelled as false negatives (FN).

At the second level, bouts are produced based on each algorithm's detected chews, and ground truth bouts based on the sEMG chews. In order to account for the variations of the bout duration we compute precision and recall based on duration. Thus, TP is the total duration where both detection and ground truth yield chewing, FP is the total duration where detection yields chewing but ground truth does not, and FN is the total duration where detection yields no chewing while ground truth does. Finally, on the third level, snacks are produced based on bouts, both for the algorithm and the ground truth. Evaluation is again performed on the basis of duration.

Regarding individual chews, our system achieves average results. MSEA achieves relatively high precision, where as LPFA achieves higher recall. For chewing bouts detection, the proposed pipeline exhibits significant improvements both for precision and recall, regardless of the algorithm used for chewing detection. In particular, when LPFA is employed, the system achieves remarkable recall of 99.3% while maintaining precision at 73.1%. Finally, for the last level of snack event detection, which is the primary mission of our system, precision and recall are even higher compared to the already high results for chewing bout detection. MSEA yields precision and recall both higher than 91%, and LPFA almost 99% recall. These results are particularly encouraging, since snacking occurrences are an important behavioural indicator for dietary monitoring and early OB risk detection. Furthermore, the PPG chewing sensor outperforms the acoustic sensor on the same dataset, as presented in [8].

V. Conclusions

In this work we have presented a novel chewing sensor based on PPG and mounted on an ear hook. The main design of the sensor has been based on our motivation about ear blood flow relation to chewing activity. We have applied the sensor in an experiment of 21 individuals, and validated its effectiveness on three different levels: chews, bouts, and snacks. The sensor has yielded satisfying results, especially for the main task of snack detection, where both precision and recall of more than 91% are achieved for on of the alternative signal processing pipelines we have suggested. We thus support the argument that it can be used for robust, objective dietary monitoring in real-life conditions. Furthermore, it's mounting design allows it to be combined with the already explored audio-based chewing sensor. Future work includes designing the pipeline that fuses information from both sensors and further increases robustness and effectiveness.

References

[1] J. B. Dixon, "Obesity in 2015: Advances in managing obesity," *Nature Reviews Endocrinology*, vol. 12, no. 2, pp. 65–66, 2016.

[2] T. H. Inge, A. P. Courcoulas, T. M. Jenkins, M. P. Michalsky, M. A. Helmrath, M. L. Brandt, C. M. Harmon, M. H. Zeller, M. K. Chen, S. A. Xanthakos, *et al.*, "Weight loss and health status 3 years after bariatric surgery in adolescents," *New England Journal of Medicine*, vol. 374, no. 2, pp. 113–123, 2016.

[3] W.-J. Lee, K.-H. Ser, Y.-C. Lee, J.-J. Tsou, S.-C. Chen, and J.-C. Chen, "Laparoscopic roux-en-y vs. mini-gastric bypass for the treatment of morbid obesity: a 10-year experience," *Obesity surgery*, vol. 22, no. 12, pp. 1827–1834, 2012.

[4] I. Ioakimidis, M. Zandian, C. Bergh, and P. Södersten, "A method for the control of eating rate: a potential intervention in eating disorders," *Behavior research methods*, vol. 41, no. 3, pp. 755–760, 2009.

[5] A. L. Ford, C. Bergh, P. Södersten, M. A. Sabin, S. Hollinghurst, L. P. mHunt, and J. P. Shield, "Treatment of childhood obesity by retraining eating behaviour: randomised controlled trial," *Bmj*, vol. 340, p. b5388, 2010.

[6] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster, "Analysis of chewing sounds for dietary monitoring," in *UbiComp 2005: Ubiquitous Computing*. Springer, 2005, pp. 56–72.

[7] S. Päßler and W.-J. Fischer, "Evaluation of algorithms for chew event detection," in *Proceedings of the 7th International Conference on Body Area Networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2012, pp. 20–26.

[8] V. Papapanagiotou, C. Diou, Z. Lingchuan, J. van den Boer, M. Mars, and A. Delopoulos, "Fractal nature of chewing sounds," in *New Trends in Image Analysis and Processing – ICIAP 2015 Workshops*, ser. Lecture Notes in Computer Science. Springer International Publishing, 2015, vol. 9281, pp. 401–408. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-23222-5_49

[9] E. S. Sazonov and J. M. Fontana, "A sensor system for automatic detection of food intake through non-invasive monitoring of chewing," *Sensors Journal, IEEE*, vol. 12, no. 5, pp. 1340–1348, 2012.

[10] M. Farooq and E. Sazonov, "Comparative testing of piezoelectric and printed strain sensors in characterization of chewing," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE, 2015, pp. 7538–7541.

[11] O. Rasmussen, F. Bonde-Petersen, L. Christensen, and E. Møller, "Blood flow in human mandibular elevators at rest and during controlled biting," *Archives of oral biology*, vol. 22, no. 8-9, pp. 539–543, 1977.

[12] S. Arberet, M. Lemay, P. Renevey, J. Sola, O. Grossenbacher, D. Andries, C. Sartori, and M. Bertschi, "Photoplethysmography-based ambulatory heartbeat monitoring embedded into a dedicated bracelet," in *Computing in Cardiology Conference (CinC), 2013*. IEEE, 2013, pp. 935–938.

[13] M. Lemay, M. Bertchi, J. Sola, P. Renevey, J. Parak, and I. Korhonen, "Applicarion of optical heart rate monitoring," in *Wearable Sensors: Fundamentals, Implementation and Applications*.

[14] M. Shuzo, S. Komori, T. Takashima, G. Lopez, S. Tatsuta, S. Yanagimoto, S. Warisawa, J.-J. Delaunay, and I. Yamada, "Wearable eating habit sensing system using internal body sound," *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, vol. 4, no. 1, pp. 158–166, 2010.

[15] C. Maramis, C. Diou, I. Ioakeimidis, I. Lekka, G. Dudnik, M. Mars, N. Maglaveras, C. Bergh, and A. Delopoulos, "Preventing obesity and eating disorders through behavioural modifications: the splendid vision," in *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on*. IEEE, 2014, pp. 7–10.

[16] J. Chen, "Food oral processinga review," *Food Hydrocolloids*, vol. 23, no. 1, pp. 1–25, 2009.