Automatic analysis of food intake and meal microstructure based on continuous weight measurements

Vasileios Papapanagiotou, *Student Member, IEEE,* Christos Diou, *Member, IEEE,* Ioannis Ioakimidis, Per Södersten, and Anastasios Delopoulos, *Member, IEEE*

***Abstract*—The structure of the cumulative food intake (CFI) curve has been associated with obesity and eating disorders. Scales that record the weight loss of a plate from which a subject eats food are used for capturing this curve; however, their measurements are contaminated by additive noise and are distorted by certain types of artifacts. This paper presents an algorithm for automatically processing continuous in-meal weight measurements in order to extract the clean CFI curve and in-meal eating indicators, such as total food intake and food intake rate. The algorithm relies on the representation of the weight-time series by a string of symbols that correspond to events such as bites or food additions. A context-free grammar is next used to model a meal as a sequence of such events. The selection of the most likely parse tree is finally used to determine the predicted eating sequence. The algorithm is evaluated on a dataset of 113 meals collected using the Mandometer, a scale that continuously samples plate weight during eating. We evaluate the effectiveness for seven indicators, and for bite-instance detection. We compare our approach with three state-of-the-art algorithms, and achieve the lowest error rates for most indicators (24 g for total meal weight). The proposed algorithm extracts the parameters of the CFI curve automatically, eliminating the need for manual data processing, and thus facilitating large-scale studies of eating behavior.**

***Index Terms*—biomedical signal processing, food intake, Mandometer, context free grammar**

I. Introduction

THE devastating effects of obesity (OB) on individual and public health, including increasing prevalence worldwide and associated cardiovascular and metabolic morbidities, are too well known to need re-reviewing [1]. The lack of effective treatments, save gastric surgical intervention, is also well known. For this reason, treatments based on eating behavior are receiving increased focus [2]; the periods from 18 to 29 years old [3] and from 12 to 26 [4] are identified as "high risk" and thus should be heavily targeted.

A new approach emerged from the treatment of patients with anorexia nervosa [5] (the prototypical eating disorder from which the other eating disorders, such as bulimia nervosa, emerge). According to this study, dieting, an important cause of underweight as well as overweight [6], rapidly changes eating behavior. For this reason, eating behavior is the target of a treatment that has markedly improved outcome [7] while the same method has also improved the treatment of OB in children [8]. This method involves the Mandometer, a scale that continuously measures the weight of the plate, connected to a monitoring device, such as a computer or a smart phone.

Besides treatment however, prevention is equally, if not more, important, given the growing spread of OB [1]. Prevention can benefit significantly by robust monitoring systems. During the last decade, new intervention methods have been introduced, as technological advancements of smart-phones and wearable sensors and devices have enabled monitoring of human, and in particular eating, behavior. To this end, a wide spectrum of single and multi-sensor approaches can be found in the literature. In-ear microphone sensors have been proposed multiple times in the literature and are one of the most studied sensor type for chewing detection [9], [10]. Microphone sensors have been also used to detect swallowing sounds when placed near the throat [11]. Strain sensors [12] and photoplethysmography [13] have also been used to detect chewing activity. Combinations of such sensors have also been proposed, for example a strain sensor, a proximity sensor, and an accelerometer have been combined in [14], while an in-ear microphone, a photoplethysmography sensor, and an accelerometer have been combined in [15]. Finally, algorithms based on wrist-mounted sensors (accelerometers and gyroscopes) have also been proposed to detect food intake cycles [16]–[18].

These systems focus on discriminating eating (from non-eating) activity, and thus mainly detect eating sessions (meals, snacks, etc); in some cases, additional in-meal indicators can be extracted, such as duration of the eating session, or individual bites. As an alternative, recording the weight of the plate at regular time intervals (e.g. every 1 sec) during a meal can yield (after some processing) the cumulative food intake (CFI) curve, a weight versus time curve that corresponds to the total weight of food that has been consumed since the beginning of the meal. Based on this CFI curve, a number of in-meal indicators can be extracted, such as total meal duration, total weight of consumed food, and number of bites (see Section IV). One additional such in-meal indicator is the eating pattern, which can be linear or decelerated [19]. According to [19], the CFI curve can be modeled using a smoothed approximation: a quadratic curve of the following

V. Papapanagiotou, C. Diou and A. Delopoulos are with the Multimedia Understanding Group, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece. E-mail: vassilis@mug.ee.auth.gr, diou@mug.ee.auth.gr, adelo@eng.auth.gr

I. Ioakimidis and P. Södersten are with the Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden. E-mail: ioannis.ioakimidis@ki.se, per.sodersten@ki.se

Manuscript received ; revised .

Fig. 1. The Mandometer, version 5.

form

$$w_{\mathrm{q}}(t) = \alpha t^2 + \beta t \tag{1}$$

where α is the intake weight acceleration, β is the initial (at $t = 0$ sec) rate of intake weight, and $t \in [0, t_{\mathrm{dur}}]$ denotes time (with t_{dur} being the duration of the meal). There is no constant term in the quadratic model, as the origin is required to belong to the curve by the assumption that no food has been consumed at the beginning of the meal, i.e. $w_{\mathrm{q}}(0) = 0$.

Based on α, one can be classified as a linear or decelerated eater [20], [21]. In linear eating, a person consumes food (measured by weight) at a constant rate, while in decelerated eating the eating rate decreases over the course of the meal [20], [21]. People suffering from OB or any ED tend to eat in a linear fashion ($\alpha \approx 0$), while healthy individuals tend to eat in a decelerated manner ($\alpha < 0$) [21]. In [8], decreasing an obese person's eating rate and the amount of food to be consumed was found to contribute to weight loss, supporting its suggested role in treatment [22]. Remarkably, this intervention also normalized hormonal secretions [23]. Conversely, increasing the speed of eating and the amount of food to be consumed increases body weight and restores the health of severely emaciated anorexic patients [5].

Extracting the eating pattern during a meal requires continuous measurement of the weight of the consumed food. The Mandometer is a device than can be used for this task (Figure 1); it is a plate scale that is placed on the table, and the plate with food is placed on scale. The device samples the weight of the plate and the food $w_{\mathrm{raw}}[n], n = 0, 1, \ldots, N_{\mathrm{raw}} - 1$ at a constant rate (1 Hz in the latest version). Based on these measurements and a video recording of the meal, the CFI curve is calculated as $w_{\mathrm{CFI}}[n], n = 0, 1, \ldots, N_{\mathrm{CFI}} - 1$, which corresponds to the total weight of food that has been consumed by the subject at a given time. An example of a recorded meal and its CFI curve are shown in Figure 2.

Even though the Mandometer is a small and portable device, its application is limited by the need for expert manual annotation of the recording. The standard process in the clinic involves setting up one or more video cameras around the table that record the entire eating session. Afterwards, the recording is synchronized with the video stream, and a clinical expert manually watches and annotates bites, which are then used to manually process the recording in order to derive the CFI curve. For use of the device at home, an expert has to obtain the data and process them without any video recordings, which can decrease accuracy. Once the CFI curve is available, the quadratic model and any additional indicators can be extracted directly. However, the requirement for manual processing limits the scale at which the Mandometer can be applied. In addition, performing large-scale statistical analysis

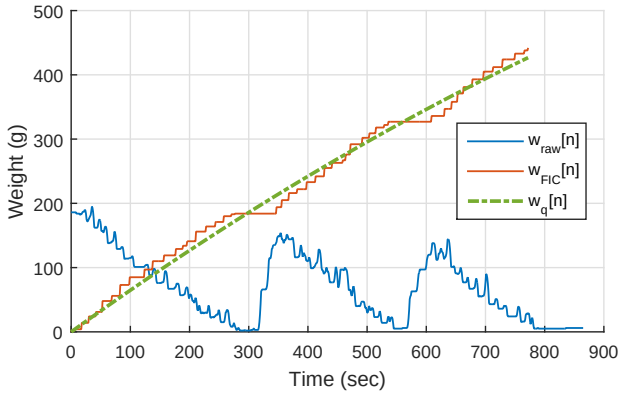


Fig. 2. An example of a raw Mandometer recording of a meal; the raw recording $w_{\mathrm{raw}}[n]$ is generally decreasing as food is consumed, and there are two food additions at 320 and 590 sec. The corresponding CFI curve $w_{\mathrm{FIC}}[n]$ quadratic model $w_{\mathrm{q}}[n]$ are also shown.

is also subject to the burden of manual processing hundreds or thousands of meal recordings.

Alternatively, the CFI curve can be obtained automatically, without the requirements for video recording and laborious manual annotation and processing, using an algorithm [24], [25]. Obtaining the CFI curve however is not a trivial task. The recorded Mandometer measurements $w_{\mathrm{raw}}[n]$ do not only contain the discrete weight decreases that correspond to bites being removed from the plate to be consumed, but are also contaminated by events that do not contribute to food consumption. Such events can correspond to the pressure applied by the utensils to the plate, temporarily resting utensils on the plate, etc. In non-controlled environments, additional events can contaminate the measurements, such as adding food to the plate (see Figure 2), taking a bite from a large piece of food and then putting the remaining food back on the plate, or taking several bites before putting the remaining food on the plate, etc.

In this work, we present an algorithm that automates the extraction of meal-related indicators by calculating the CFI curve of a meal based only on continuous weight measurements from the Mandometer. The algorithm is based on modeling a meal with a context-free grammar (CFG). After an initial pre-processing stage, the raw recording is partitioned into segments and a CFG terminal symbol is assigned to each segment, thus forming a string. All possible parse trees of the string are evaluated and the most likely tree is selected to interpret the meal events. Based on the obtained tree, we calculate the CFI curve and extract various in-meal indicators from the CFI. We evaluate our algorithm on a dataset of 113 non-controlled meals collected in the context of the EU SPLENDID project[1] [26], and compare with three state-of-the-art algorithms [24], [25], [27].

The rest of this paper is organized as follows: Section II presents related work for processing meal weight recordings. Section III presents the proposed algorithm that calculates

[1]The aim of SPLENDID is to provide personalized services guiding adolescents and young adults to healthy eating and activity behaviors, preventing the onset of obesity and eating disorders, by leveraging existing and novel prototype sensors. http://splendid-program.eu/

the CFI curve while Section IV presents the list of extracted eating-behavior indicators. Section V presents the collected evaluation dataset, the evaluation methodology, and the experimental results. Finally, Section VI concludes the paper.

II. Related work

Regarding computing the CFI curve from Mandometer measurements, we have already published two works [24], [25]. In [24], three preliminary algorithms are presented. The first one, named rule-based (RB) is a set of heuristic rules for detecting food additions; additional heuristics for detecting inactivity at the beginning and end of the meal are also part of the algorithm. The RB algorithm relies heavily on predefined thresholds. In order to reduce the need for calibrating thresholds, the second algorithm, named rule-based quadratic fitting (RBQF) uses thresholds that yield high recall for detecting candidate food additions. Then, all possible combinations of candidate food additions are considered, and the mean-squared error of each CFI curve and fitted quadratic model is used to rank the combinations; the one achieving the lowest error is selected as the final result. Finally, the third algorithm, named greedy quadratic fitting (GQF) operates iteratively on the measurements; once a candidate food addition is detected, the previous measurements are used to accept or reject it, based on two mean-squared errors of two different quadratic curves (one corresponding to accepting the candidate food addition and one to rejecting it).

In [25], a PCFG is introduced to interpret the events that occur during a meal, in particular bites, food additions, and artifacts (pressure on the plate while cutting or using the fork, or temporarily rest of a utensil on the plate). Each weight measurement is directly assigned one terminal symbol based on the derivative (sample difference). All parse trees are then computed, and a probability is assigned to each based on the probabilities of all events (bites, food additions, and artifacts) of the parse tree. The probability for each of the three event types is modeled using a parametric function; the parameters depend on various attributes of each event (e.g. for the bite event type, the function depends on the weight of the bite). The most probable parse tree is selected as the interpretation of the meal measurements and is used to calculate the CFI curve.

Another device similar to the Mandometer is the universal eating monitor (UEM), a table-embedded scale that also records food weight during eating. The UEM can support an entire tray that holds both plates and glasses. In [27], an algorithm is proposed that automatically processes an UEM meal session recording to detect bites. The algorithm initially detects "stable" time intervals of the recording during which the weight standard deviation is less than an accepted threshold. Then, the transitions from a stable time interval to another are examined, and a set of threshold rules are used to detect three different bite types: single food bites, food mass bites, and drink bites. Authors report a recall of 0.39 for single food and food mass bites, and also 0.39 for drink bites.

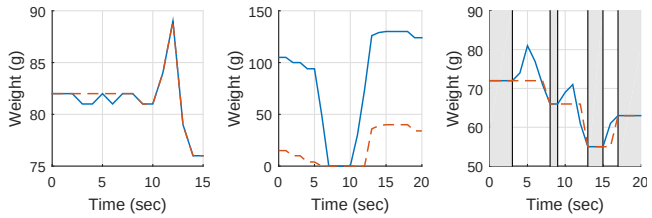
A preliminary version of our approach has been reported in [25] where we use a PCFG to model a meal. CFGs are extended to PCFGs by assigning a probability to each rule; these probabilities must satisfy certain requirements [28], [29]. However, using a PCFG as a method to select the parse tree that "best" interprets a meal is not very suitable. First, a PCFG assigns a probability to each rule so that the sum of probabilities for all rules that substitute the same non-terminal symbol must equal one, essentially defining a distribution among the possible substitutions of that non-terminal. Although one can argue that bites occur more frequently than a food addition, the distribution can depend on other parameters such as the total length of the meal, and thus might not be easy to estimate. Furthermore, PCFGs assign a constant probability to each rule, however, the rule $B \rightarrow b$ (this rule assigns a bite B to a weight decrease b, see Section III-B) can have different probability depending on the weight decrease b corresponds to (e.g. a bite of 6 g is more likely than a bite of 100 g). Finally, due to independence assumption of PCFGs, longer (sub)strings tend to be less probable. The main differences of our approach and the one presented in [25] are

1) All signal processing before applying the grammars is different. In particular, [25] uses morphological operators (opening) to manipulate the raw data into a suitable form for the PCFG, while the algorithm proposed in this paper uses a different, three-stage pre-processing approach (see Section III-A).
2) The terminal symbols are applied in a different manner. In particular, [25] assigns one terminal symbol to each sample of the raw recording, while the algorithm proposed in this paper partitions the raw recording into segments of varying length, and assigns a terminal symbol to each segment.
3) The grammar is expanded with a rule for bites where some food is picked up from the plate, a bite is taken, and the remaining food is placed back on the plate. Also, some pruning is performed to reduce the computational complexity of parsing.
4) The ranking method for the parse trees is different. In particular, [25] uses ad-hoc distributions to select the probability of each rule, while the algorithm proposed in this paper is data-driven; known distributions are fitted to training data and are then used to derive the likelihood for each event (bite, food addition, and artifact).
5) The post-processing is completely different. In particular, [25] removes the detected food additions and artifacts and then extracts the CFI curve, while the algorithm proposed in this paper directly computes the CFI curve based on detected bite events.

III. Food intake curve algorithm

A. Pre-processing and segmentation

Initially, we perform three pre-processing steps to prepare the raw data for segmentation, as shown in Figure 3. First, the recording $w_{\text{raw}}[n]$ is pre-processed to remove "jitter"; jitter can occur during inactive moments when the recorded weight might increase or decrease by 1 g, and a few samples later might return to the original value, without any other weight change in-between. Next, the plate weight which is sampled



(a) Jitter (b) Plate removal (c) Stable samples

Fig. 3. Examples of the pre-processing steps. Input and output of each step is drawn with a continuous blue and a dashed red line respectively. Grey areas in Figure 3c mark stable time intervals.

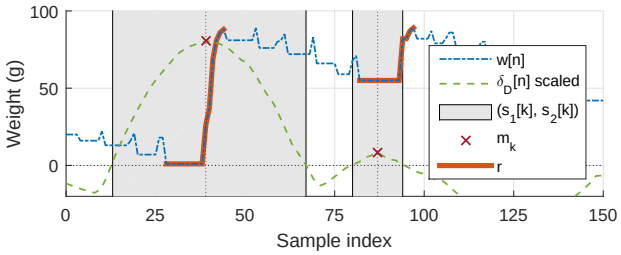


Fig. 4. An example of detecting r-intervals (bold red line). Based on $\delta_D[n]$ (dashed green line), two time intervals (gray areas) are detected: $(s_1[0], s_2[0]) = (13, 67)$ and $(s_1[1], s_2[1]) = (80, 94)$. The maximum $\delta_D[n]$ per time interval are $m_0 = 30$ and $m_1 = 87$ (burgundy "x"), around which the r time intervals are detected.

before the meal starts is subtracted from all weight samples, and negative values are set to zero. Finally, samples are characterized as stable or unstable; a sample is stable if its weight is equal to the weight of either the previous or the next sample (i.e. time intervals where the recorded weight is constant, see Figure 3c). The weight of each unstable sample is then explicitly set equal to the weight of the closest past stable sample. The meal recording after the pre-processing steps is denoted $w[n]$.

We then segment $w[n]$ into three types of intervals, one type for each terminal symbol of the CFG; r-intervals ("important" weight increase/rise), d-intervals (big weight decrease), and b-intervals (small weight decrease). The need for these three terminal symbols and the way that they are used in the CFG are described in the next Section (Section III-B). Segmenting $w[n]$ is based on the forward derivative $\Delta[n]$ and the delta coefficients $\delta_D[n]$, which are computed as

$$\Delta[n] = w[n+1] - w[n], n = 0, 1, \ldots, N-2 \quad (2)$$

$$\delta_D[n] = h_D[n] * w[n], n = 0, 1, \ldots, N-1 \quad (3)$$

where $*$ denotes convolution and $h_D[n]$ is the impulse response of an FIR filter with $2D+1$ taps given by

$$h_D[n] = \frac{n}{\sum_{i=-D}^{D} i^2}, n = -D, \ldots, D \quad (4)$$

The delta coefficients have been used extensively in speech processing [30], [31], and capture the "trend" of $w[n]$ (increasing or decreasing) using an interval of $2D+1$ samples, thus providing an estimation of the derivative that is less sensitive to short-term changes (depending on the value of D). This property is very useful in identifying events such as food additions; food additions are recorded as sudden increases in weight, however, similar increases can be recorded when pressured is applied by fork and knife before a bite. In both cases the derivative $\Delta[n]$ exhibits high values during the weight increase, however the delta coefficients $\delta_D[n]$ exhibit high values only in the case of food addition (see also Figure 4). In our work, we have selected a window of 1 minute for the estimation of $\delta_D[n]$, and we have thus set $D = 30$ given the 1 Hz sampling rate; thus, the denominator of Equation 4 is constant.

Important weight increases are detected using the following method (Figure 4 illustrates an example). Non-overlapping time intervals $[s_1[k], s_2[k]]$, are detected such that all delta coefficients are positive, i.e. $\delta_D[n] > 0$ for all $s_1[k] \leq n \leq s_2[k]$. From each such time interval k we mark the index m_k of the highest coefficient in the interval, i.e. $\delta_D[m_k] \geq \delta_D[n]$ for $s_1[k] \leq n \leq s_2[k]$. If $\Delta[m_k] > 0$ then we find the longest interval around m_k in which $w[n]$ is non-decreasing, and assign the terminal symbol r to it, which corresponds to a critical weight increase (rise). If $\Delta[m_k] \leq 0$ then we assign no symbol.

Finally, b-intervals and d-intervals are intervals where weight strictly decreases. For such a strictly decreasing interval, if the total weight decrease is small enough, i.e. less than a threshold Δ_{thr}, the interval is regarded as a b-interval, otherwise as a d-interval. The threshold Δ_{thr} is selected by thresholding the cumulative normalized histogram of ground truth bites at 95% (Figure 8b); typical values are close to 20 g when computed in leave-one-subject-out (LOSO) fashion (see Section V-B).

An artifical example of a meal is illustrated in Figure 5. The identified intervals of the segmentation are noted with different symbols and colors. The curve does not exhibit a clear decreasing trend, as it is polluted with a food addition, an artifcact, and a burger bite. Figure 6 shows the same data after remove the effect of all non-eating behavior; five bites are clearly visible.

B. Modeling meals as strings

We identify the following three event types: bite (B), food addition (F), and artifact (A), where the letters in parenthesis correspond to the respective non-terminal symbols of the CFG. Thus, a meal can be modeled as a sequence of events of B, F, and A (Equation 5).

A bite (B) event type denotes the action of picking up food from the plate in order to consume it. In it's simplest form, a bite causes the measured weight to be decreased by the weight of the bite (Equation 6, $B \rightarrow b$). If multiple bites are taken very quickly, it is possible that a single, longer, and greater weight decrease is measured (Equation 6, $B \rightarrow d$). However, such greater weight decreases can be the result of other event types such as removal of a utensil (e.g. Equation 8). Sometimes, pressure applied by utensils before a bite can cause weight to increase momentarily before decreasing. Usually, the weight increase is filtered out by the third pre-propcessing step (Figure 3c). However, for the cases where the weight increase survives the filtering, we introduce the third bite rule which describes

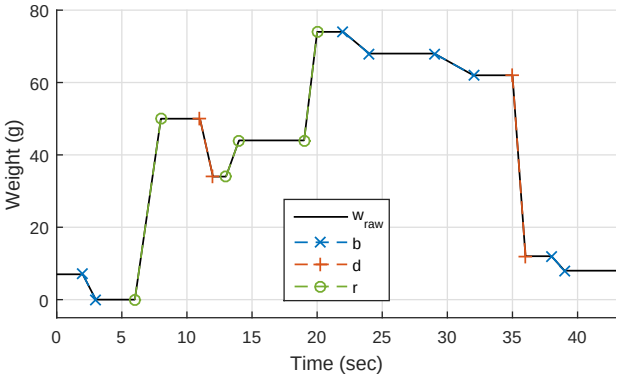


Fig. 5. An artificial example of a meal (black line represents the raw recording). The segments for each terminal symbol are denoted with different symbols and colors of dashed lines. Contrary to intuition, the curve is not decreasing; various events have polluted the curve. In particular, after an initial bite (blue, 2-3 sec) a utensil is placed on the plate (green, 6-8 sec) and is removed later on (red, 35-36 sec). A bite of the form dr (Equation 6) occurs at 11-14 sec, and then a food addition (green, 19-20 sec). Three more bites occur at 22-24, 29-32, and 38-39 sec (blue). The extracted string is $brdrrbbdb$. Figure 6 shows the same data after removing the effect of food additions, artifacts, and simplifying the dr bite.

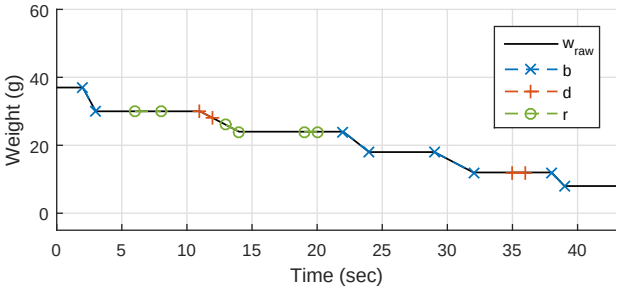


Fig. 6. The data of the meal of Figure 5 after removing the effect of food additions, artifacts, and simplifying the dr bite. The colors for detected segments are retained for direct comparison with Figure 5. Five bites are now clearly visible at 2-3, 11-14, 22-24, 29-32, and 38-39. The disimilarities between this curve and the curve of Figure 5 (even though it is an artificial example) are an indication of the challenges of the algorithm.

bites as a weight increase followed by a big weight decrease (Equation 6, $B \rightarrow rd$). Finally, we define a fourth type of bite, that describes bites taken typically from food types such as burgers or chicken wings: a large portion of the food is picked up from the plate, registering a significant decrease in measured weight, a bite is taken, and the remaining food is placed back on the plate, registering a weight increase that is however less than the previous weight decrease. Thus, this type of bites is described by Equation 6, $B \rightarrow dr$.

An important note is that a CFG bite event can actually correspond to multiple "actual" bites, e.g. in the case of many bites taken too quickly and resulting in a single d, or in the case of picking up food such as a burger, taking multiple bites and then putting down the remaining food, resulting in dr. For such cases, any approach relying exclusively on measured weight cannot accurately detect all actual bites. This is a known limitation, however its impact on effectiveness is minimal: if the weight difference is estimate correctly, the algorithm will accurately estimated most indicators (such as total food weight), albeit for a few missed bites.

Food addition (F) refers to the act of adding more food on the plate (from some other container), thus increasing the measured weight (Equation 7). The weight usually increases by greater quantity (compared to the weight decrease of a bite), while the duration varies.

Finally, the artifact (A) event type corresponds to more than one action; however, all these actions share the common characteristic that they result in a weight increase, followed (instantly or later) by an almost equal weight decrease. Examples of such actions are temporarily resting a utensil or other item on the plate, resting one's hand, etc. For such types of actions, any combination of meal events can occur between the placing and removal of the extra weight. All this behavior can be modeled by Equation 8; r and d correspond to the weight increase and decrease, while S allows for any (or none) events to take place in-between.

Thus, the CFG we use is formally defined as $\mathbf{G} = \{\mathbf{V}, \mathbf{\Sigma}, \mathbf{R}, S\}$, where $\mathbf{V} = \{S, B, F, A, e, b, d, r\}$ is the alphabet, $\mathbf{\Sigma} = \{b, d, r, e\}$ is the set of terminal symbols, $\mathbf{R} \subseteq (\mathbf{V} - \mathbf{\Sigma}) \times \mathbf{V}^\star$ is the following set of rules

$$S \rightarrow BS|FS|AS|e \quad (5)$$

$$B \rightarrow b|d|rd|dr \quad (6)$$

$$F \rightarrow r \quad (7)$$

$$A \rightarrow rSd \quad (8)$$

S is the start symbol, e is the empty string symbol, and * is the Kleene star operator [32]. Note that $\mathbf{L}(\mathbf{G}) = \mathbf{\Sigma}^\star$, i.e. the language generated by $\mathbf{G}$ is the set of all possible strings that can be obtained by concatenating any combination of the CFG terminal symbols.

An efficient algorithm for parsing CFGs that does not require the CFG to be in Chomsky normal form [32] has been proposed by Jay Earley [33]. We use our Java implementation available on GitHub[2] to compute all parse trees $\mathcal{T}_j$ of the string s. An example of a parse for the artificial meal of Figure 5 is shown in Figure 7. Each tree corresponds to a different interpretation of the meal. Thus, an interpretation is equivalent to a sequence of events $E[i]$, $i = 0, 1, \ldots, N_E - 1$, where each event is a B, an F, or an A. In Figure 7 the sub-tree is denoted with blue color and corresponds to the sequence B_1, A, B_2 (subscripts of B_1 and B_2 are only used for presentation). However, since A can spawn an entire meal using the recursive rule of Equation 8, all possible sub-trees must be taken into account. In Figure 7 the interpretation of the sub-tree (shown in green color) is the sequence $B_3 F B_4 B_5$ (again, the subscripts are only used for presentation).

C. Maximum likelihood interpretation

Once the string representation of a meal is obtained, and all parse trees are available, a method is required to evaluate each parse tree and select the most likely one. We thus use a modified version of CFGs, in which we estimate the likelihood of a parse tree based on parameters derived from $w[n]$. In particular, the likelihood $\mathcal{L}(\mathcal{T}_j)$ of the parse tree $\mathcal{T}_j$ is obtained

[2]https://github.com/mug-auth/pcfg-toolbox

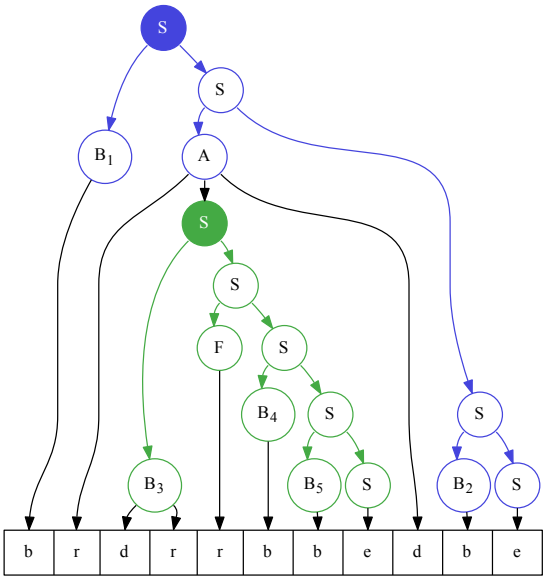


Fig. 7. A parse tree for the meal of Figure 5. The blue S is the tree root. The meal is interpreted as the sequence of B_1AB_2, while the artifact A spawns a sub-tree with the green S as root. The sub-tree is interpreted as $B_3FB_4B_5$. The empty string symbol e has been inserted in the string where it was required.

based on the event interpretation of the meal, i.e. the sequence of detected events $E[i]$, as

$$\mathcal{L}(\mathcal{T}_j) = \prod_{i=0}^{N_E-1} \mathcal{L}(E[i]) \tag{9}$$

using the naive assumption that each event is independent of all other events (note that for selecting the most likely parse tree, our implementation uses in fact the sum of log-likelihood, which is equivallent and numerically stable). We thus need to define the likelihood functions for each event type.

We model the likelihood of a bite event B by the following exponential distribution function

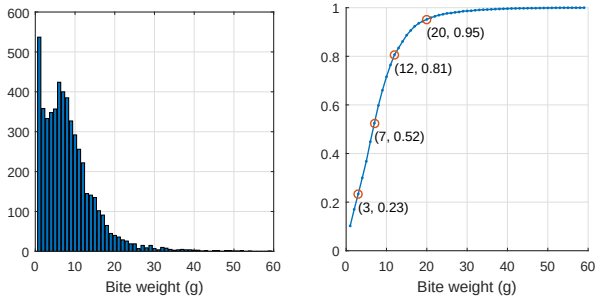
$$\mathcal{L}(B;\theta_B) = \lambda e^{-\lambda\theta_B} \tag{10}$$

where θ_B is the bite weight. Note that bite weight corresponds directly to the weight decrease for bites described by b and d, while it corresponds to the total weight decrease for rd and dr bites (i.e. weight decrease of d minus weight increase of r). This formula reflects our intuition that bite weights are in general small, and greater bite weights are less likely to occur (see Figure 8a). The value of parameter λ is selected by fitting an exponential distribution on the normalized histogram of ground truth bites; typical values for λ when estimated in LOSO fashion are close to 8.4.

Food addition weights lie almost uniformly on the entire range of 0 - 250 g (Figure 9a). Thus, we approximate the likelihood of a food addition by the sigmoid function

$$\mathcal{L}(F;\theta_F) = \frac{1}{1+e^{-a(\theta_F-c)}} \tag{11}$$

where θ_F is the weight of the food addition. Parameters a and c are obtained by fitting the sigmoid function on the normalized cumulative histogram (see Figure 9b). Typical values are close



(a) (b)

Fig. 8. Histogram (8a) and normalized cumulative histogram (8b) of ground truth bite weight.

to 0.02 for a and close to 115 for c, when estimated in LOSO fashion.

Finally, the likelihood of an artifact A is the product of two independent terms and defined as

$$\mathcal{L}(A;\theta_A) = e^{-\frac{(\theta_A)^2}{2\sigma^2}} \cdot \mathcal{L}(\mathcal{T}_A) \tag{12}$$

The exponential term is proportional to a Gaussian distribution with zero mean and standard deviation σ, and models the fact that placing and removing an item (e.g. a utensil) on the scale increases and then decreases the measured weight by (almost) the same amount; this corresponds to a substring of the form rSd (Equation 8) where the weight difference between r and d is close to 0 g. The Gaussian is in fact scaled so that $\mathcal{L}(A;0) = 1 \cdot \mathcal{L}(\mathcal{T}_A)$. The standard deviation σ cannot be directly estimated since there is no available ground truth for artifacts. However, it can be estimated indirectly based on the observation that there are substrings of the form rSd which are not artifacts and r corresponds to a food addition and d to a bite. Thus, we can estimate the distribution f_{F-B} of the "food-addition-weight minus bite-weight" combinations as[3]

$$f_{F-B}(x) = f_F(x) * f_{-B}(x) \tag{13}$$

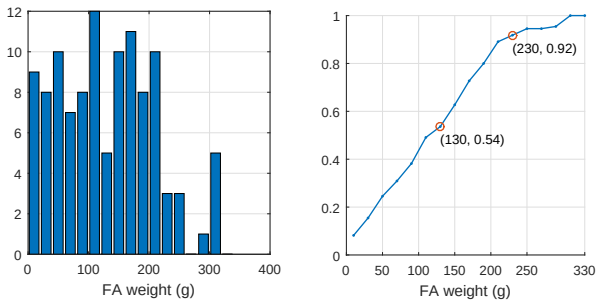
where f_F is the empirical distribution (histogram) of ground truth food addition weights, and f_{-B} is the empirical distribution of ground truth bite weights (note that ground truth bite weights are used with a negative sign). By thresholding the cumulative distribution of f_{F-B} at 5% we can obtain a value for parameter σ; typical values for σ when estimated in LOSO fashion are 4 and 5 g.

The term $\mathcal{L}(\mathcal{T}_A)$ is the likelihood of the sub-tree spawning from the S of Equation 8, and is computed in the same way as for any parse tree (Equation 9). If there is no sub-tree (S is eliminated using the last alternative of Equation 5), we set $\mathcal{L}(\mathcal{T}_A) = 1$.

D. Calculating the CFI curve

Based on the above, the most likely $\mathcal{T}_j$ is selected, and is used to construct the CFI curve based on the detected bites (since only bites correspond to intake activity). In particular,

[3]based on the fact that the probability density function (PDF) of the sum $X+Y$ of random variables X and Y is the convolution of their PDFs f_X and f_Y



(a) (b)

Fig. 9. Histogram (9a) and normalized cumulative histogram (9b) of ground truth food addition weight.

let the pairs $(t_B[i], w_B[i])$, $i = 0, 1, \ldots, N_B - 1$ denote the time-stamp and weight of each bite. The CFI curve $w_{\text{CFI}}[n]$ is obtained by sample-hold interpolation of $(t_B[i], w_B[i])$ at 1 Hz. Also, the time before the first bite (0 to $t_B[0] - 1$ sec) is removed, and thus $w_{\text{CFI}}[n]$ runs for $n = 0, 1, \ldots, N_{\text{CFI}} - 1$, where $N_{\text{CFI}} = t_B[N_B - 1] - t_B[0] + 1$ samples.

IV. In-meal indicators

Given a CFI curve $w_{\text{CFI}}[n]$, $n = 0, 1, \ldots, N_{\text{CFI}} - 1$ and the corresponding bites $(t_B[i], w_B[i])$, $i = 0, 1, \ldots, N_B - 1$, we extract the following seven in-meal eating-behavior indicators. Meal duration (in seconds) is obtained as

$$t_B[N_N - 1] - t_B[0] + 1 \tag{14}$$

Total meal intake (in g) is obtained as the sum of all bite weights $w_B[i]$. In addition, eating rate, which can be used to modify eating patterns towards healthier behavior [8], can be directly obtained as the ratio of total meal intake over meal duration.

We also compute the mean and standard deviation of bite weights $w_B[i]$ (both in g too), and the mean bite frequency (in Hz) as

$$\left(\frac{1}{N_B - 1} \sum_{i=0}^{N_B - 2} t_B[i+1] - t_B[i] \right)^{-1} \tag{15}$$

Finally, by solving the following minimization problem

$$\min_{\alpha,\beta} \sum_{i=0}^{N_{\text{CFI}}-1} \left(\alpha i^2 + \beta i - w_{\text{CFI}}[i] \right)^2 \tag{16}$$

we obtain the CFI acceleration α (in g/sec^2) that is used to characterize linear and decelerated eating patterns, and the initial CFI rate β (in g/sec).

V. Evaluation

A. Dataset

To evaluate our algorithm we use a dataset of 113 meals collected during three trials in the context of the EU funded project SPLENDID [26]. The trials took place in the Wageningen University, Netherlands, and in the Karolinska Institutet, Sweden. The 113 meals used in this work belong to 77 participants. Their demographic data are shown in Table I.

TABLE I
Demographics of the three recording trials: date of recording trials, number of male and female participants (M/F), mean and standard deviation of age (in years) and BMI (in kg/m^2), total number of meals and bites.

Date	M/F	Age	BMI	Meals	Bites
Apr-Sep 2014	14/13	26.2 (5)	23.9 (2.1)	52	2,615
Mar-Apr 2015	18/21	16 (0)	21.4 (2.5)	39	1,487
Apr-May 2015	0/11	22.8 (1.6)	21.7 (2.1)	22	1,153
Total	32/45	20.6 (5.6)	22.3 (2.6)	113	5,255

Most participants had normal body-mass-index (BMI) during the trials, and no problems were reported. In total, the 113 meals contain 5,255 bites. A subset of these meals (76) for which "risk" ground truth is available have been used in [34]. Another subset of these meals have been used in [35].

Eating during the trials was completely unrestricted. Participants were free to consume as much food as they wanted, and refill from a buffet. In particular, the second trial with the 39 young participants of 16 years old took place in the Internationella Engelska Gymnasiet Södermalm school, during the regular eating break at the school cafeteria. The types of consumed food include vegetables with chicken, tomato and mean soup, minced meat, fish and potatoes, beefburger/chickenburger with potatoes and salad. There was no limitations on utensils.

For each meal, the ground truth CFI curves for all the meals were calculated independently by the eating behaviour analysis scientists at KI. A validated methodological approach [36] was used, manually combining and correcting raw Mandometer data based on meal-event occurrences annotated on video recordings of the meals. In the past similar manual data analysis approaches have been used in a wide range of experimental settings and target populations [37], [38].

B. Evaluation methodology

We evaluate our algorithm by comparing the values of various in-meal indicators extracted from ground truth CFI curves and the algorithm's CFI curves. We use this evaluation approach since directly comparing the CFI curves (e.g. using mean squared error) can be misleading. For example, if the algorithm "misses" one bite, then all the remaining CFI is offset by the weight of the missed bite. The earlier the missed bite occurs, the greater it contributes to the mean squared error, which is counter-intuitive. However, we indirectly compare the curves on the basis of bite-weight sequences using the Earth's mover's distance (EMD) [39].

To assess the effectiveness of our algorithm (or any of the algorithms we compare against, see Section V-C) for an indicator, we first compute the indicator value for each meal i twice: $I[i]$ based on the ground truth CFI curve (see Section V-A), and $\hat{I}[i]$ based on the algorithm's CFI curve. We then compute the absolute error of the algorithm as the absolute difference

$$e[i] = \left| I[i] - \hat{I}[i] \right| \tag{17}$$

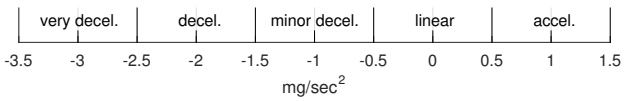


Fig. 10. Visualization of the classes used in evaluation of α. Each class is 0.5 mg/sec^2 wide, corresponding to half the width of five eating patterns: linear, decelerated (three levels), and accelerated.

We present the mean and standard deviation of absolute error across all meals of the dataset in Table II, for each indicator and algorithm pair.

We also evaluate indicators α and β with a second classification-based approach. Similar to [24], we define a set of classes for each indicator. In particular, α values (based on ground truth) are roughly within the range of -3.5 to 1 mg/sec^2. We thus partition this range into classes of 0.5 mg/sec^2 width. Based on ground truth and algorithm values for α we assign ground truth and predicted class labels, and compute the accuracy based on the confusion matrix (see Table III). The same method is used for β; β values are roughly within 0 to 2.5 g/sec, and we thus partition using classes of 0.25 g/sec width. Figure 10 shows the 10 classes we use and their correspondence in eating patterns.

We also evaluate accurate detection of bite instances. For a given meal, let $(t_B[i], w_B[i])$, $i = 0, 1, \ldots, N_B - 1$ be the ground truth bite pairs (time-stamp and weight), and similarly let $(t_{\hat{B}}[i], w_{\hat{B}}[i])$, $i = 0, 1, \ldots, N_{\hat{B}} - 1$ be the algorithm's detected pairs. We perform an one-to-one matching between the ground truth and detected bites so that each ground truth bite can be matched with at most one detected bite, and vice versa. The i-th ground truth bite is matched with the j-th detected bite if

$$\left|t_B[i] - t_{\hat{B}}[j]\right| \leq t_{\text{thr}} \tag{18}$$

$$\left|w_B[i] - w_{\hat{B}}[j]\right| \leq w_{\text{thr}} \tag{19}$$

i.e. if both the predicted time-stamp and weight absolute errors are less than two thresholds. We select $w_{\text{thr}} = 3$ g to allow some small tolerance for weight detection. We also select $t_{\text{thr}} = 10$ sec; this value might seem a bit high at first, however, there is an inherent delay between the moment the food is picked from the plate (and a correct detection is expected to occur) and the moment the bite is actually placed in mouth (and the clinical experts annotate as occurrence of a ground truth bite). Also note that the time mismatch for bite detection is also taken into account in EMD (Table III). Each matched pair contributes as one true positive (TP) detection. Each non-matched ground truth bite contributes as one false negative (FN) detection, and each non-matched detected bite contributes as one false positive (FP) detection. We then compute the total TP, FP, and FN across the dataset (by summation), and present precision and recall for bite detection in Table III.

Finally, we compare the sequence the time-weight sequences of bites using the EMD. Based on the notation of [39], signatures P and Q correspond to the sets of ground truth and detected bite sequences respectively

$$P = \{(t_B[i], w_B[i]), i = 0, 1, \ldots, N_B - 1\} \tag{20}$$

$$Q = \{(t_{\hat{B}}[i], w_{\hat{B}}[i]), i = 0, 1, \ldots, N_{\hat{B}} - 1\} \tag{21}$$

where the cluster representatives are the time-stamps of bites, and the bite weights are the cluster's weights. We then solve the minimization problem defined in [39] and compute the actual EMD. It is important to note that the total mass of each signature is not equal; the difference corresponds to the total meal weight error. Since the EMD does not take into account extra mass (i.e. incurs no penalty if the algorithm has detected more or less total mass), results should always be considered together with total meal weight errors.

In order to apply the algorithm, values for the following parameters are required: Δ_{thr} (Section III-A), λ (Equation 10), a and c (Equation 11), and σ (Equation 12). Thus, we apply the algorithm in LOSO fashion, where for each participant, the meals of all other participants are used to estimate the values of these parameters (as described in Section III-C). The algorithm is then applied on the meals of the left-out participant with the estimated values. This process is repeated for every participant.

C. Results & discussion

We evaluate our algorithm and compare its effectiveness with the following three algorithms: greedy quadratic fitting (GQF) from [24] (since it demonstrates the highest effectiveness among the three algorithms proposed in [24]), the original PCFG algorithm from [25], and our implementation of the UEM algorithm of Mattfeld *et al.* [27]. Since GQF and PCFG algorithms do not directly detect bites (even though PCFG adopts a bite event), we detect bites for these two algorithms by simply detecting weight increases in the CFI curve; despite this being a naive approach, it works very well for GQF (see Table III). For the algorithm in [27] we have implemented the "single bite" and "food mass bite" stages, and have omitted the "drink bite" stage, since only the plate is placed on the Mandometer. For the threshold values of parameters W_1, W_2, and W_3 (Equations (1-4) of [27]) we use the suggested values. We also set $3\sigma_{\text{noise}} = 1$ g to account for the recorded jitter (see Section III). Finally, for "single bite" we use only Equation (1) and not Equation (2) of [27], since the requirement of Equation (2) decreases the detected bites to almost zero.

Our proposed algorithm achieves minimum error rates for five out of the seven indicators. In particular, the mean error for total weight is less than 25 g, while the second-lowest result (achieved by PCFG) is more than 100 g. Meal duration error is 61 sec, more than 25 sec less than the second-lowest error of 88.5 sec (for PCFG). Mean bite weight error is 1.4 g for our approach, while the second-lowest error is achieved by Mattfeld *et al.* [27] and is 2.1 g. Note that t-test indicates a significant difference between our algorithm and each of the GQF, PCFG, and Mattfeld *et al.* algorithms ($p < 0.05$, for most cases p is even lower, e.g. $p < 0.01$), except for the following three cases: (a) STD bite weight, Mattfeld *et al.* vs. our approach, $p = 0.27$, (b) mean bite frequency, GQF vs. our approach, $p = 0.23$, and (c) mean bite frequency, PCFG vs. our approach, $p = 0.9$. This is highly encouraging since these three cases are the only ones that our algorithm does not achieve minimum error. Finally, standard deviation of error is

TABLE II
MEAN AND STANDARD DEVIATION OF ABSOLUTE DEVIATIONS FROM GROUND TRUTH INDICATOR VALUES FOR EACH INDICATOR-ALGORITHM PAIR. DIAMOND ($^\diamond$) INDICATES NO SIGNIFICANT DIFFERENCE ($p > 0.05$) OF THE CORRESPONDING ALGORITHM COMPARED TO OUR APPROACH. BOLD INDICATES TOP EFFECTIVENESS (MINIMUM MEAN ERROR).

	GQF [24]	**PCFG [25]**	**Mattfeld *et al.* [27]**	**Our approach**
Duration (sec)	108.2 (164.6)	88.5 (138.8)	102.7 (164.7)	**60.7** (108.6)
Weight (g)	225.5 (320.9)	103.2 (142)	118.6 (83)	**24.3** (47.1)
Mean bite weight (g)	7.67 (10.78)	3.04 (3.82)	2.1 (1.93)	**1.4** (1.91)
STD bite weight (g)	31.59 (45.84)	8.46 (9.58)	**1.67**$^\diamond$ (1.52)	2.08 (3.8)
Mean bite frequency (mHz)	**13.3**$^\diamond$ (17.8)	14.9$^\diamond$ (14.7)	23.3 (19.1)	15.1 (15.7)
α (mg/sec^2)	0.375 (0.587)	0.464 (0.577)	0.435 (0.607)	**0.244** (0.416)
β (g/sec)	0.21 (0.364)	0.255 (0.323)	0.422 (0.307)	**0.114** (0.154)

TABLE III
CLASSIFICATION ACCURACY FOR α AND β COEFFICIENTS, PRECISION AND RECALL FOR BITE DETECTION, AND EARTH MOVER'S DISTANCE. BOLD INDICATES TOP EFFECTIVENESS (MAXIMUM ACCURACY, PRECISION, AND RECALL, AND MINIMUM EMD).

	GQF [24]	**PCFG [25]**	**Mattfeld *et al.* [27]**	**Our approach**
α accuracy	0.602	0.434	0.442	**0.69**
β accuracy	0.619	0.451	0.204	**0.743**
Bites precision	**0.84**	0.635	0.428	0.793
Bites recall	**0.815**	0.611	0.343	0.744
EMD (sec)	14.92	22.39	32.23	**6.45**

the lowest for our algorithm for most indicators, indicating the increased robustness of our approach.

Regarding indicators α and β, our algorithm achieves the minimum error among the other three (Table II) as well as the highest classification accuracy (Table III); accuracy for our approach is 0.69 and 0.743 for α and β respectively; these values are 0.088 and 0.124 higher than the second most effective results of GQF.

Finally, for bite detection, our proposed algorithm achieves the second highest precision and recall, outperformed only by GQF. However, the achieved precision of 0.79 and recall of 0.74 are encouraging, while the effectiveness gain of GQF is not comparable to the significant gain of our algorithm for most indicators, such as meal duration, total meal intake, α, and β. This effectiveness (in bite detection) of our approach (as well as of GQF) is high enough to enable practical application in real-life. Indeed, Mattfeld *et al.* [27] report that their algorithm is able to measure just 39% percent of bites, however this is sufficient to accurately estimate average bite weight. In addition, we have observed that our algorithm sometimes detects one large bite that corresponds to two smaller ground truth bites; this leads to one FP and two FN bites in the confusion matrix (based on our strict evaluation methodology for bite detection, Section V-B). However, even though this decreases bite detection effectiveness, it does not affect the accurate estimation of other indicators (such as weight of meal) since the measured weight of consumed food is the same.

VI. CONCLUSIONS

In this work we propose an algorithm that automatically extracts the CFI curve and in-meal eating-behavior indicators from the Mandometer, eliminating the need for laborious manual work from clinical experts, and thus enabling its application on the large scale. The algorithm is based on a CFG; non-terminal symbols model the different event types that can occur during a meal, while terminal symbols are assigned to different parts of the recording based on delta coefficients. Meals are mapped to strings of terminal symbols which are then interpreted by means of (possibly multiple) parse trees. We select the most likely interpretation based on the likelihood of each event. The CFI curve is then constructed and we extract indicators such as meal duration and total meal intake, as well as the α and β coefficients of the quadratic intake model.

We evaluate our algorithm on a dataset of 113 meals and compare its effectiveness against three other algorithms, and find that it achieves significantly better results for most indicators. We also evaluate on bite-instance detection, where our algorithm achieves the second-best precision and recall rates.

ACKNOWLEDGMENTS

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 (http://splendid-program.eu), and from the European Community's Health, demographic change and well-being Programme under Grant Agreement No. 727688, 01/12/2016 - 30/11/2020 (http://bigoprogram.eu).

Mandometer was supplied by Mando Group AB.

REFERENCES

[1] G. . O. Collaborators, "Health effects of overweight and obesity in 195 countries over 25 years," *The New England journal of medicine*, vol. 377, no. 1, p. 1327, July 2017. [Online]. Available: http://europepmc.org/articles/PMC5477817

[2] N. Gletsu-Miller and M. A. McCrory, "Modifying eating behavior: Novel approaches for reducing body weight, preventing weight regain, and reducing chronic disease risk," *Advances in Nutrition: An International Review Journal*, vol. 5, no. 6, pp. 789–791, 2014. [Online]. Available: http://advances.nutrition.org/content/5/6/789.abstract

[3] M. C. Nelson, M. Story, N. I. Larson, D. Neumark-Sztainer, and L. A. Lytle, "Emerging adulthood and college-aged youth: An overlooked age for weight-related behavior change," *Obesity*, vol. 16, no. 10, pp. 2205–2211, 2008. [Online]. Available: http://dx.doi.org/10.1038/oby.2008.365

[4] D. Neumark-Sztainer, M. Wall, M. Story, and A. R. Standish, "Dieting and unhealthy weight control behaviors during adolescence: Associations with 10-year changes in body mass index," *Journal of Adolescent Health*, vol. 50, no. 1, pp. 80 – 86, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1054139X11001765

[5] C. Bergh, M. Callmar, S. Danemar, M. Hölcke, S. Isberg, M. Leon, J. Lindgren, Å. Lundqvist, M. Niinimaa, B. Olofsson *et al.*, "Effective treatment of eating disorders: Results at multiple sites." *Behavioral neuroscience*, vol. 127, no. 6, p. 878, 2013.

[6] I. Ioakimidis, M. Zandian, F. Ulbl, C. Bergh, M. Leon, and P. Södersten, "How eating affects mood," *Physiology & Behavior*, vol. 103, no. 3, pp. 290 – 294, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031938411000400

[7] P. Södersten, C. Bergh, M. Leon, U. Brodin, and M. Zandian, "Cognitive behavior therapy for eating disorders versus normalization of eating behavior," *Physiology & Behavior*, vol. 174, pp. 178 – 190, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031938416308824

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

[9] O. Amft*, M. Kusserow, and G. Trster, "Bite weight prediction from acoustic recognition of chewing," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 6, pp. 1663–1672, June 2009.

[10] O. Amft, "A wearable earpad sensor for chewing monitoring," in *2010 IEEE Sensors*, Nov 2010, pp. 222–227.

[11] E. S. Sazonov*, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, "Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, pp. 626–633, March 2010.

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

[13] V. Papapanagiotou, C. Diou, L. Zhou, J. van den Boer, M. Mars, and A. Delopoulos, "A novel approach for chewing detection based on a wearable ppg sensor," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 6485–6488.

[14] J. M. Fontana, M. Farooq, and E. Sazonov, "Automatic ingestion monitor: A novel wearable device for monitoring of ingestive behavior," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1772–1779, June 2014.

[15] V. Papapanagiotou, C. Diou, L. Zhou, J. van den Boer, M. Mars, and A. Delopoulos, "A novel chewing detection system based on ppg, audio, and accelerometry," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 607–618, May 2017.

[16] Y. Dong, A. Hoover, J. Scisco, and E. Muth, "A new method for measuring meal intake in humans via automated wrist motion tracking," *Applied Psychophysiology and Biofeedback*, vol. 37, no. 3, pp. 205–215, 2012. [Online]. Available: http://dx.doi.org/10.1007/s10484-012-9194-1

[17] K. Kyritsis, C. L. Tatli, C. Diou, and A. Delopoulos, "Automated analysis of in meal eating behavior using a commercial wristband imu sensor," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 2843–2846.

[18] K. Kyritsis, C. Diou, and A. Delopoulos, "Food intake detection from inertial sensors using lstm networks," in *New Trends in Image Analysis and Processing – ICIAP 2017 Workshops: ICIAP 2017 International Workshops, MADiMa, Catania, Italy, September 13-15, 2017, Proceedings*. Cham: Springer International Publishing, 2017 [accepted for publication].

[19] H. R. Kissileff, J. Thornton, and E. Becker, "A quadratic equation adequately describes the cumulative food intake curve in man," *Appetite*, vol. 3, no. 3, pp. 255–272, 1982.

[20] M. Westerterp-Plantenga, K. Westerterp, N. Nicolson, A. Mordant, P. Schoffelen, and F. ten Hoor, "The shape of the cumulative food intake curve in humans, during basic and manipulated meals," *Physiology & Behavior*, vol. 47, no. 3, pp. 569 – 576, 1990. [Online]. Available: http://www.sciencedirect.com/science/article/pii/003193849090128Q

[21] M. Zandian, I. Ioakimidis, C. Bergh, and P. Södersten, "Linear eaters turned decelerated: Reduction of a risk for disordered eating?" *Physiology & behavior*, vol. 96, no. 4, pp. 518–521, 2009.

[22] 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.

[23] J. Galhardo, L. P. Hunt, S. L. Lightman, M. A. Sabin, C. Bergh, P. Södersten, and J. P. H. Shield, "Normalizing eating behavior reduces body weight and improves gastrointestinal hormonal secretion in obese adolescents," *"The Journal of Clinical Endocrinology & Metabolism"*, vol. 97, no. 2, pp. E193–E201, 2012. [Online]. Available: +http://dx.doi.org/10.1210/jc.2011-1999

[24] V. Papapanagiotou, C. Diou, B. Langlet, I. Ioakimidis, and A. Delopoulos, *Automated Extraction of Food Intake Indicators from Continuous Meal Weight Measurements*. Cham: Springer International Publishing, 2015, pp. 35–46. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-16480-9_4

[25] ——, "A parametric probabilistic context-free grammar for food intake analysis based on continuous meal weight measurements," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 7853–7856.

[26] 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.

[27] R. S. Mattfeld, E. R. Muth, and A. Hoover, "Measuring the consumption of individual solid and liquid bites using a table embedded scale during unrestricted eating," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, no. 99, pp. 1–1, 2017.

[28] A. Stolcke, "An efficient probabilistic context-free parsing algorithm that computes prefix probabilities," *Comput. Linguist.*, vol. 21, no. 2, pp. 165–201, Jun. 1995. [Online]. Available: http://dl.acm.org/citation.cfm?id=211190.211197

[29] T. L. Booth and R. A. Thompson, "Applying probability measures to abstract languages," *IEEE Transactions on Computers*, vol. C-22, no. 5, pp. 442–450, May 1973.

[30] T. Masuko, K. Tokuda, T. Kobayashi, and S. Imai, "Speech synthesis using hmms with dynamic features," in *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, vol. 1, May 1996, pp. 389–392 vol. 1.

[31] K. Tokuda, T. Yoshimura, T. Masuko, T. Kobayashi, and T. Kitamura, "Speech parameter generation algorithms for hmm-based speech synthesis," in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*, vol. 3, 2000, pp. 1315–1318 vol.3.

[32] H. R. Lewis and C. H. Papadimitriou, *Elements of the Theory of Computation*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1997.

[33] J. Earley, "An efficient context-free parsing algorithm," *Commun. ACM*, vol. 13, no. 2, pp. 94–102, Feb. 1970. [Online]. Available: http://doi.acm.org/10.1145/362007.362035

[34] C. Diou, I. Sarafis, I. Ioakimidis, and A. Delopoulos, "Data-driven assessments for sensor measurements of eating behavior," in *2017 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, Feb 2017, pp. 129–132.

[35] B. Langlet, A. Anvret, C. Maramis, I. Moulos, V. Papapanagiotou, C. Diou, E. Lekka, R. Heimeier, A. Delopoulos, and I. Ioakimidis, "Objective measures of eating behaviour in a swedish high school," *Behaviour & Information Technology*, vol. 36, no. 10, pp. 1005–1013, 2017. [Online]. Available: https://doi.org/10.1080/0144929X.2017.1322146

[36] I. Ioakimidis, M. Zandian, L. Eriksson-Marklund, C. Bergh, A. Grigoriadis, and P. Södersten, "Description of chewing and food intake over the course of a meal," *Physiology & Behavior*, vol. 104, no. 5, pp. 761 – 769, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031938411003738

[37] I. Ioakimidis, M. Zandian, F. Ulbl, C. Âlund, C. Bergh, and P. Södersten, "Food intake and chewing in women," *Neurocomputing*, vol. 84, no. Supplement C, pp. 31 – 38, 2012, from Neuron to Behaviour: Evidence from Behavioral Measurements. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231211007351

[38] B. Langlet, P. Fagerberg, A. Glossner, and I. Ioakimidis, "Objective quantification of the food proximity effect on grapes, chocolate and cracker consumption in a swedish high school. a temporal analysis," *PLOS ONE*, vol. 12, no. 8, pp. 1–16, 08 2017. [Online]. Available: https://doi.org/10.1371/journal.pone.0182172

[39] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *International Journal of Computer*

Vision, vol. 40, no. 2, pp. 99–121, Nov 2000. [Online]. Available: https://doi.org/10.1023/A:1026543900054

Vasileios Papapanagiotou is a PhD student in the Multimedia Understanding Group (MUG) of the Information Processing Laboratory (IPL), Department of Electrical and Computers Engineering, Aristotle University of Thessaloniki (AUTh). He received his diploma from the same Department in 2013. His research interests are digital signal processing, wearable sensors, behavioral monitoring and analysis, supervised and semi-supervised machine learning, and concept-based image retrieval.

Christos Diou received his diploma and PhD from the Electrical and Computer Engineering Department of the Aristotle University of Thessaloniki, in 2004 and 2010 respectively. Since 2011 he has been working as a Research Associate at the Information Processing Laboratory of the Aristotle University of Thessaloniki in the areas of machine learning and signal processing, for applications such as multimedia analysis and retrieval, analysis of human eating and physical activity behavior, analysis of electrical energy consumption behavior of small-scale consumers as well as machine learning for computer security applications.

Ioannis Ioakimidis received his B.S. in biology from the University of Crete in Greece. Afterwards he received his PhD from Karolinska Institute in 2011 and he has since worked in the Department of Neurobiology, Care Sciences and Society, Division of Applied Neuroendocrinology first as a Postdoctoral fellow and then as an assistant Professor. His research interests focus on the objective quantification of human eating behavior with the use of new technologies, mainly in regards to obesity and eating disorders.

Per Södersten is professor of behavioral neuroendocrinology at Karolinska Institutet. He explored the neuroendocrinology of reproduction and, the other side of the same neurobiological coin, eating, over close to 50 years. With Dr Cecilia Bergh, the research was translated into the management of under- and overweight patients over 25 years. The work has resulted in some 200 hundreds publications, some of which are important as they provide the scientific basis for the clinical work that has restored the health of more than 1,000 patients.

Dr. Anastasios Delopoulos was born in Athens, Greece, in 1964. He graduated from the Department of Electrical Engineering of the National Technical University of Athens (NTUA) in 1987, received the M.Sc. from the University of Virginia in 1990 and the Ph.D. degree from NTUA in 1993. From 1995 till 2001 he was a senior researcher in the Institute of Communication and Computer Systems of NTUA. Since 2001 he is with the Electrical and Computer Engineering Department of the Aristotle University of Thessaloniki where he serves as an associate professor. His research interests lie in the areas of machine learning, signal and multimedia processing and computer vision. On the applied domain he works in the areas of multimedia retrieval, biomedical engineering and behavioral informatics. He is the (co)author of more than 80 journal and conference scientific papers. He has participated in 21 European and National R&D projects related to application of his research to entertainment, culture, education and health sectors. Dr. Delopoulos is a member of the Technical Chamber of Greece and the IEEE.