CHAPTER

7.4

Detection and Characterization of Food Intake by Wearable Sensors

Juan M. Fontana and Edward Sazonov

The University of Alabama, Tuscaloosa, Alabama, USA

1. INTRODUCTION

Food intake is the primary source of energy and nutrients necessary to maintain life. Monitoring of daily food intake and ingestive behavior is an important area that has direct implications on human health, as inadequate or excessive energy intake may lead to development of medical conditions such as malnutrition and underweight, or overweight and obesity, respectively. Understanding ingestive behavior is also a key in diagnosis and treatment of eating disorders such as anorexia, bulimia, and binge eating.

Food provides the chemical energy needed for functioning of the vital organs and performing physical activity, with the excess energy being stored in glycogen and adipose tissue for future use. The balance between energy intake from food and energy expended on basal metabolism and physical activity is an essential factor for maintaining a steady body weight in humans. A persistent imbalance between these two components is the cause of long-term changes in body weight, potentially leading to abnormal weight loss or gain. While hunger and malnourishment still remain an issue for a large part of the world's population, obesity has recently overtaken hunger as a global health threat.

Obesity, defined as the excessive accumulation of body fat, is the result of a chronic weight gain produced when the energy obtained from foods overcomes the energy expended. Excessive food intake (especially intake of calorie-rich foods now widely available on a global scale) may be a major contributor to the obesity epidemic. For example, in the United States, the prevalence of obesity reached a total of 35.5% among adults and 16.9% among adolescents in 2009–2010 [1]. Individuals suffering from obesity may potentially face a number of health issues ranging from cardiovascular problems to diabetes, and can expect a reduction in their life expectancy [2].

Eating disorders are serious mental disorders that cause disturbances in eating habits or weight-control behavior of individuals [3]. Anorexia nervosa, bulimia nervosa, and binge eating are the most common eating disorders with lifetime prevalence ranging from 0.6 to 4.5% in the United States [4]. Individuals with anorexia nervosa restrict their food intake by dieting, fasting, or excessive physical activity from the fear of gaining weight and distorted perceptions of their body shape and size. People with bulimia nervosa have periods of excessive food intake (binging) followed by a feeling of guilt that leads to extreme ways to compensate for binging (deliberate vomiting, crash dieting, and strenuous exercise). Binge eating is similar in terms of symptoms to bulimia nervosa, but does not include the extreme compensatory reactions typical to bulimia.

Both obesity and eating disorders are medical conditions highly resistant to treatment and can have severe physical and physiological health consequences [5]. The monitoring of food intake is considered to be the basis for behavioral treatment of obesity and eating disorders, which can be managed with dietary modification and control. Thus, monitoring of food intake is extremely important for identifying, understanding, and correcting food-intake patterns of individuals.

Traditionally, the ingestive behavior in humans is assessed by the means of self-monitoring. Various self-reporting methods were developed to estimate the timing and duration of intake as well as to characterize the intake in terms of amount of food consumed, energy, and nutrient intake [6]. Methods such as dietary records, 24-hr dietary recall, food frequency questionnaires, and diet history are widely used. All of these methods rely on a person's own declaration of what was eaten, when, where, and how much food was consumed. However, they suffer from underreporting, which is considered to be about 20% on average but may be as high as 50% [7,8]. The low accuracy of self-reporting food intake is mainly determined by two factors. First, there is a change in the eating behavior of individuals when they know they are being observed (the observation effect). Second, there is the tendency to either underestimate portion sizes or avoid reporting certain foods (reporting effect). For example, it has been shown that people tend to misreport or not report snacking at all, which may contribute significantly to the daily energy intake [9].

Thus, there is a critical need for developing methods for objective assessment of food intake, especially under free-living conditions. Such methods must provide accurate detection and characterization of food intake with minimal or no conscious effort from the subjects. The fundamental assumption is that the objectivity of monitoring will minimize or eliminate the reporting effect, while removing the conscious effort will minimize the observation effect.

Wearable sensors present a compelling possibility for monitoring of food intake. The dramatic technological advances in the past few decades allow building of miniature devices that can potentially detect the process of food ingestion and further characterize the ingested foods. A wearable ingestion sensor can potentially be very objective and capture all ingestion events, regardless of how short or insignificant they may seem. A wearable sensor could potentially capture timing, duration, and microstructure of food-intake episodes, characterize rate of ingestion, ingested mass, and nutritional and energy contents of food, without creating a reporting burden for the user. The individuals still need to comply with the requirement of wearing the sensor, but with an ergonomic design the wearer's burden can be minimal. The ingestion sensor can potentially recognize when it is being worn and when it is not, thus measuring the compliance of the individual.

Use of wearable sensors for monitoring of ingestion faces a number of challenges, with the main challenge being the great variety of foods that humans consume. Food is made from the most creative combinations of multiple ingredients, each with its own energetic and nutritional contents. Modifying just one ingredient may substantially change the nutritional properties of food (e.g., excluding butter or other forms of fat from a recipe may dramatically reduce the energy content). Foods have different physical properties and can be solid, liquid, semi-liquid, dry, moist, crispy, soft, and chewy, just to name a few possibilities. These physical properties are not necessarily correlated to the energy and nutritional content of food. The physical properties may also depend on whether the food is served raw or cooked in a certain way. Such variability of foods and food properties complicates and inevitably introduces errors even into the most sophisticated types of analyses.

The second challenge is the diversity of ingestive behaviors characteristic of humans. During an ingestion episode, we may eat a single food item or feast on a dozen or more of different food items. We may eat a little bit or a whole lot. We eat with hands, spoons, forks, chopsticks, or drink our food. We may have fairly stable meal times or customarily skip meals or eat at different times every day. Some individuals will eat fast, while others will eat slowly. Some will take breaks between different parts of a meal, while others will immediately start the next food item. We may eat during the waking hours or during the night. While most of us will consume a meal while seated, eating may also take place on the go or even lying down. If anything, the ingestive behavior is as variable and unpredictable as humans.

These two fundamental challenges create a variety of technological requirements for implementing wearable devices for food-intake monitoring. What sensor modality should be used? What are the capabilities and inherent limitations of such sensors? Can a given technology be made convenient, miniature, and lightweight to be worn for prolonged periods of time? What is the battery life of such a device? These and other questions need to be answered for any potential solution for food-intake monitoring.

This chapter presents an overview of the wearable sensors and accompanying methodologies proposed for food-intake monitoring. The overview is based on the recent research literature, with the focus on devices and methods developed in the Computer Laboratory of Ambient and Wearable Systems at The University of Alabama. Here, the task of monitoring ingestive behavior is considered as consisting of two subtasks: detection of food intake and characterization of food intake. Specifically, the task of detection of food intake includes:

- Detection and timing of each food-intake episode
- Measuring duration and microstructure of each episode

The task of characterization of food intake includes:

- Estimating the number and type of food items in a meal
- Estimating the mass and volume of ingestion
- Estimating caloric and nutritional content of a meal
- Measuring the rate of ingestion of each episode

This chapter is organized as follows. First, a description of the different sensor modalities developed for monitoring food intake is presented. Second, a description of the signal processing and pattern-recognition algorithms for automatic food-intake detection is presented. Third, methods of characterization of food intake are described. Fourth, an example of applying a wearable sensor for monitoring of food intake in free living is presented together with the challenges. Finally, future directions along with limitations and open issues concerning food intake are summarized at the end of the chapter.

2. WEARABLE SENSORS

Our society is increasingly dependent on wearable and mobile products that can offer effective ways to process information and interact with people. Many new approaches for monitoring of food intake are based on wearable sensors. These approaches can be divided in two main groups depending on how they are worn by individuals: handheld devices (such as mobile phones) and body-attached sensors.

2.1 Handheld Devices

Handheld devices are widely used to report and self-monitor daily intake by incorporating dietary software programs (electronic diaries). The use of mobile phones and personal digital assistants (PDAs) has significantly increased over the last decade, offering a new alternative for recording food intake. These computationally powerful devices can connect to the Internet anywhere and anytime, allowing Web-based interventions that can improve clinical management [10,11]. The advantage is that cell phones are carried by individuals most of the time, allowing them to make immediate food-intake annotations, which may help to improve cooperation and accuracy.

Dietary software programs have been developed and implemented into these electronic devices for a simpler and less burdensome food-intake monitoring (Figure 1). Such programs include a source database with thousands of food items with the corresponding nutritional information. For example, many diaries use the U.S. Department of Agriculture database, which has approximately 6,000 food items [12]. The integration of a food database into a user-friendly mobile application reduces the amount of time and labor required to report the intake when compared to paper-and-pencil diaries. Subjects simply log the foods consumed without the burden of searching and calculating the nutritional content of the meal. The software embedded in the electronic device automatically computes the total energy consumed in the meal from the amount and type of food reported. Electronic food diaries can also store the date and time of each entry, which can be used by clinicians or researchers to validate the reported data.

Another advantage of electronic diaries is that they can provide real-time feedback to the users about their nutrient intake such that individuals can make adjustments to their food intake to meet daily intake goals (e.g., calories, fat, carbohydrates, etc.). Additionally, some programs can connect to the Internet and upload the nutritional information for

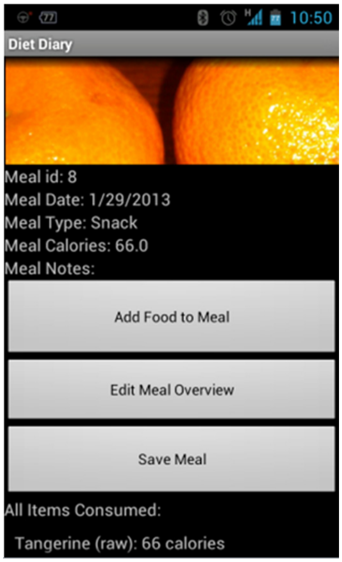

FIGURE 1 Example of an electronic diary implemented as a mobile phone application. Critical information about the consumed meals are recorded and stored in the phone.

further analysis by dietitians and clinicians. Finally, electronic diaries eliminate the issue of illegible handwriting present in paper-based methods [13].

Currently, most mobile phones are equipped with digital cameras that allow individuals to take high quality pictures of their meals. This technological advance has helped to improve the electronic diaries by providing additional information about a meal. Food-intake detection and characterization through food imagery is accomplished by taking photographs of the serving plate with the foods selected by an individual and of the plate's waste. Portion sizes of the food selection are estimated by a trained dietitian in the laboratory who compares the photographs taken by individuals against reference portion photographs of known quantities [14,15]. Estimation of the type and amount of foods consumed can also be done automatically by using methodologies based on image processing (see section 4.2). These estimates are entered into a food analysis program to derive the total mass and energy consumed along with macro- and micronutrients of food selection based on a source database. Accurate results are obtained when the pictures are taken always at the same angle (usually 45 degrees) with the serving plate occupying the entire field of view. This methodology has provided reliable results when used to measure energy intake of adults and children under different settings [16,17]. However, it may suffer from the limitations of self-reported intake as individuals need to remember to take pictures of the foods before and after consumption.

2.2 Body-Attached Sensors

Body-attached sensors monitor physiological processes in the body and focus on one or more stages of the food consumption process: hand-to-mouth gestures, bites, chewing, or swallowing.

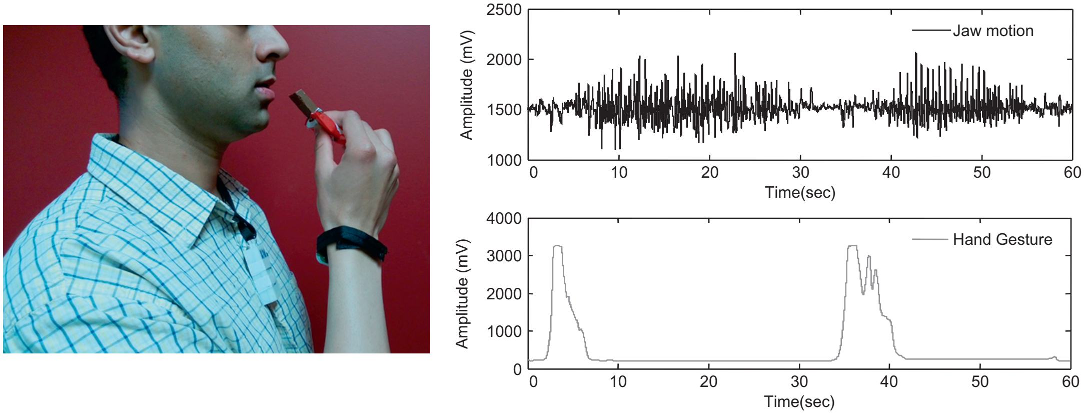

FIGURE 2 Left: Typical hand-to-mouth gesture occurring during food intake. Right: Hand gesture signal (bottom plot) collected during a short episode of food intake. The increase in signal amplitude occurs during a hand-to-mouth gesture related to a bite. The top plot illustrates a chewing sequence (captured by a jaw motion sensor) that follows the detected hand-to-mouth gesture.

2.2.1 Monitoring of Hand Gestures

Quite often, ingestion of food begins with a hand-to-mouth gesture followed by a bite. Consequently, ingestion events can potentially be detected by monitoring hand gestures occurring in daily living. Researchers have proposed a number of approaches integrating different sensor modalities into a wearable device [18−20].

In one approach, a proximity sensor operating in the radio frequency identification band was proposed to monitor hand-to-mouth gestures. This sensor consisted of an RF transmitter module worn on the inner side of the dominant arm and an RF receiver, which is a flat-loop antenna located in a package worn as a pendant around the neck (Figure 2, left) [21]. The location of the antennas was strategically selected such that the highest magnitude of the signal (when the antennas are aligned in parallel) is obtained for a typical hand-to-mouth gesture observed during food intake (Figure 2, left). The behavior of the proximity sensor is illustrated in Figure 2 (right), where a segment of the RF signal received by the antenna (bottom plot) shows two hand-to-mouth gestures. To help visualize these gestures within a food-intake episode, the top plot of Figure 2 (right) shows the chewing sequences occurring immediately after the hand-to-mouth gestures and captured by a jaw motion sensor. While the hand gesture sensor signal may not be sufficient by itself to discriminate hand-to-mouth gestures on various origins, it provides important features that can be used in a pattern-recognition algorithm [22].

In another approach, a watch-like device (Figure 3) was developed for measuring intake through an automatic tracking of wrist motions during hand-to-mouth gestures or "bites" [20]. This device incorporated a micro-electromechanical gyroscope to monitor the radial velocity of the wrist during food intake. The main advantages of this device are the low cost of the sensor and need for only a single wearable item (as compared with the RF sensor).

FIGURE 3 A wearable sensor for monitoring hand-to-mouth gestures associated with food intake incorporates a miniature gyroscope for capturing the motion of the wrist. *(Image courtesy of Dr. Adam Hoover [20].)*

Finally, accelerometers have been used for recognition of hand gestures [18,23]. Typically, the systems implementing this type of sensor use 3-axes accelerometers located on the arm to measure acceleration. The resulting signals are transmitted to a PC or smartphone via bluetooth or other wireless protocol for automatic detection of hand gestures.

A common limitation of hand gesture monitoring is that not only food intake, but many other activities can produce hand gestures that may be indistinguishable from gestures during food intake. This limitation stipulates either the use of additional sensors to capture independent indicators of food intake (such as in [24]) or the use of self-report in the form of turning the sensor on/off for each ingestion episode (such as in [20]). Another limitation is that people may use both hands during solid and liquid intake, which would require the use of two devices to reliably detect hand gestures.

2.2.2 *Monitoring of Chewing*

In a typical ingestion episode, a bite is followed by a sequence of chews and swallows, and this process is repeated throughout an entire meal. Thus, the ingestion of solid foods can be detected by monitoring the chewing process. Several sensing options have been evaluated with this purpose. Surface electromyography (EMG) can sense the activation of jaw muscles during mastication by placing electrodes over the skin surface [25]. Multi-point sheet-type sensors [26] and strain-gauge abutments [27] can also be used to measure bite and chewing forces. However, these sensors are likely to produce variations in an individual's normal mastication patterns because they are placed between the teeth.

Another sensing option is based on the monitoring of sounds produced in the chewing process. During food breakdown, portions of foods are crushed between the teeth, generating vibrations that are conducted through the teeth, the mandible, and the skull [28]. These vibrations are also propagated through the ear canal, which forms a natural cavity where they are audible. Therefore, an acoustic transducer can be placed in the proximity of the ear canal to capture the sounds of chewing. Several studies have been conducted based on this approach [29–32]. To reduce effects of environmental noise, a miniature in-ear microphone is used along with a reference microphone for noise cancellation (Figure 4). The acoustic signal is then processed in order to automatically detect and characterize food intake.

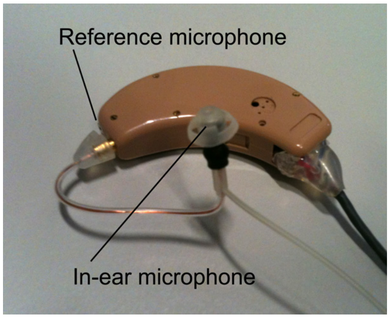

FIGURE 4 A wearable sensor system consisting of: 1) an in-ear microphone to monitor the chewing and swallowing sounds, and 2) a reference microphone to capture environmental sounds. *(Image courtesy of Dr. Sebastian Päßler [32].)*

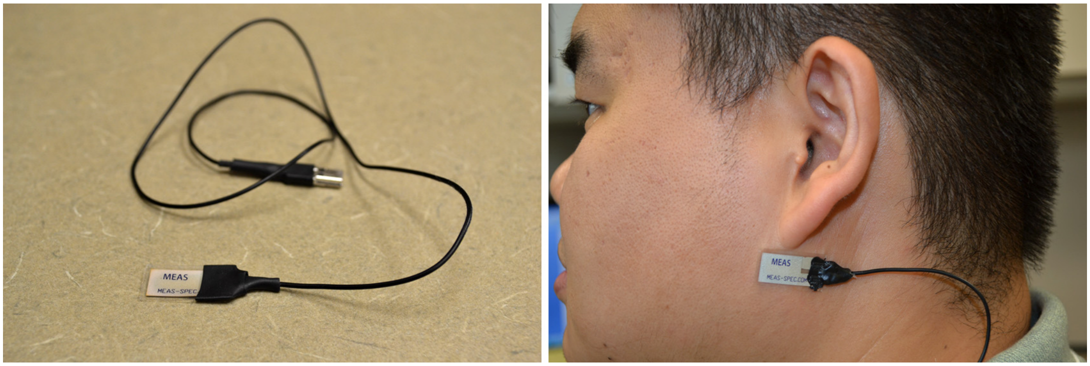

FIGURE 5 Left: Jaw motion sensor consisting of a piezoelectric film element that generates a voltage signal when flexed. Right: Sensor is placed below the earlobe and attached to the skin's surface using medical adhesive or medical tape.

Monitoring of the motion of the mandible (jaw) occurring during chewing is another reliable approach for capturing food intake. Monitoring of jaw motion, rather than monitoring chewing sounds, may result in a simpler and more accurate sensor system, as the sensor is less sensitive to environmental noise. The jaw motion can be detected by monitoring changes in skin curvature due to changes in distance between the mandible and temporal/occipital bones of the skull during chewing. Foil-strain gauges have been tested providing reasonable results but with unacceptably high energy consumption. A good option was found to be an off-the-shelf piezoelectric film strain gauge sensor (Figure 5, left), which can monitor dynamic skin strain with low power consumption [33]. The changes in skin curvature are detected by placing the strain sensor immediately below the earlobe (Figure 5, right). The simple structure of the chewing strain sensor may result in a less intrusive and simpler way to detect food intake. Additionally, this sensor is simply attached to the skin by medical adhesive or medical tape, and may last for a 24-hour period with a low risk of sensor detachment. Finally, the strain sensor is easily available in the market at a low cost.

Chewing has also been used in conjunction with digital imagery through the integration of miniature, lightweight cameras into wearable sensors to facilitate the assessment of dietary intake [34]. These cameras can potentially be worn by individuals throughout the day and can be triggered to automatically take pictures without requiring user input, thus reducing the misreporting bias [35]. An example of this trend is a wearable sensor platform [36] designed as a headset worn on the ear. The sensor consists of a microphone for detection of the chewing sound and a miniature camera directed toward the table where foods are consumed. The chewing activity triggers the camera to capture a video sequence of the food container. This action creates a food-intake log having a series of snapshots and time-stamps that are saved into a file to keep record of the food consumption history.

2.2.3 Monitoring of Swallowing

The swallowing process is a sequential, semi-automatic contraction and relaxation of muscles of the tongue, pharynx, and esophagus. Spontaneous swallowing is a functional act involving automatic and involuntary transport of saliva from the mouth to the stomach. When a person is awake, spontaneous swallowing of saliva occurs at a rate that varies from 1 to 2 swallows per minute [37–39]. During sleep, the production of saliva virtually ceases and spontaneous swallows are occasional, occurring generally in association with movement arousals [40,41]. During food intake, the swallowing process involves the passage of the bolus of food (or liquid) from the mouth to the stomach. Studies have demonstrated that the swallowing rate significantly increases during eating, suggesting that monitoring of swallowing episodes may provide suitable information to detect ingestion of food.

Videofluoroscopy and electromyography are considered the gold standard methods for swallowing studies [42,43]. However, the development of a wearable monitor based on any of these methods may not be feasible due to their invasiveness (subcutaneous EMG) and their dependency on bulky, expensive, and possibly unsafe equipment (videofluoroscopy). To overcome this issue, different sensor modalities have been proposed to monitor a particular phenomenon occurring during the swallowing process. For example, neck mounted accelerometers were used to capture throat vibrations, microphones were used to record sounds produced during swallowing, an electroglottograph device was evaluated to monitor changes in the electrical impedance of the neck, and magnetic sensors were implemented to measure the movement of the thyroid cartilage. This section describes some of these sensors in more detail and highlights their advantages and disadvantages.

The use of accelerometers has been proposed as a non-invasive, low-cost alternative for studying swallowing. The acceleration signal is produced during laryngeal displacement and the magnitude of such signal is a measure of the elevation of the larynx during swallowing [44]. Typically, miniature accelerometers are placed on the skin, over the throat, and at the level of the thyroid cartilage of the subject. However, this location may not be suitable for obese individuals due to accumulation of body fat under the chin. Additionally, the sensitivity of accelerometers to body motion and to orientation in the gravity field may also jeopardize the implementation of this sensor modality for freely moving individuals.

Monitoring of swallowing by acoustical means has been relying on miniature microphones that are able to record the sounds generated when the bolus of food passes through the pharynx in the pharyngeal phase of the swallowing process. This characteristic sound can be captured from different locations including inside the ear canal, on the mastoid bone, or over the laryngopharynx. Research has shown that microphones placed over the

FIGURE 6 Left: Miniature microphone used to monitor swallowing events. Right: Acoustic sensor is placed over the laryngopharynx using a collar.

laryngopharynx (Figure 6) offer the advantage of stronger signals as the sensors are located closer to the origin of the swallowing sound [45].

A novel approach for swallowing detection is based on an electroglottograph (EGG) device. The EGG sensor measures the variation in the transverse electrical impedance across the neck at larynx level. The EGG signal is recorded by applying a high frequency signal (3 MHz) to guard-ring electrodes placed on the surface of the neck. EGG has been widely used for speech and swallowing analysis [46,47], but only recently has been applied to food-intake detection [48]. The passage of the bolus of food during the swallowing process causes significant variations in the electrical impedance across the larynx from the baseline, allowing detection of swallowing. The advantage of this sensor modality is that the EGG signal is virtually insensitive to external acoustic noise because of the physical principles used in EGG measurement. Disadvantages of this approach are related to motion artifacts that may arise during free living and may affect the EGG signal.

Recently, a magnetic sensor system was proposed to detect swallowing events by monitoring the movement of the thyroid cartilage [49]. This system also contained a piezoelectric microphone to detect swallowing sounds. The microphone and the magnetic coils were mounted in a holding unit that was positioned at the neck level. The distance between two coils (oscillation and detection coils) placed on each side of the thyroid cartilage was used for detection of swallowing. The level of voltage induced in the detection coil by a 20 kHz magnetic field produced in the oscillation coil was proportional to the length between coils.

3. SIGNAL PROCESSING AND PATTERN-RECOGNITION METHODS FOR AUTOMATIC DETECTION OF FOOD INTAKE

Wearable sensors convert a physiological event associated with ingestion into an electrical signal. Most often, the sensor signal carries not only information about ingestion but also information about other actions of the individual. For example, an acoustic swallowing sensor will register not only spontaneous and food-related swallowing but will also respond to speech, walking, and other activities. Thus, specialized signal processing and pattern-recognition algorithms are required for discriminating swallowing from other activities, and for detection of food intake from the time sequence of swallows. The following is a description of commonly used signal processing and pattern-recognition methods for analysis of the signals collected by different sensor modalities and automatic detection of food intake.

3.1 Food-Intake Detection from Imagery

Capturing an image of the foods being consumed and writing an entry in a multimedia food diary is, by itself, a method for food-intake detection that captures information about event timing, and in some cases, event duration. Using a digital diary requires no specialized processing to detect food intake. This simplicity carries a major limitation: individuals need to always carry the device and remember to take pictures of every food item they eat, subjecting the method to all of the relevant limitations of self-report. To alleviate this problem, a potential solution is to utilize automatic image recognition to detect food intake from continuous video monitoring. In [50] a wearable device was developed for food-intake detection from imagery. The principal assumption of this approach is that different features from the video footage such as motion patterns, presence of circular objects, surrounding sounds, and time of the day can provide valuable information to detect ingestion episodes. The food-intake detection faces many challenges, including privacy concerns for continuously captured video and the ability to automatically recognize the great variety of foods and make judgment whether the food is being consumed or just viewed by the user.

3.2 Detection of Hand Gestures

The process of ingestion begins when individuals take foods and put them into their mouth. This simple event requires a hand-to-mouth gesture. Automatic detection of hand-to-mouth gestures is thus a fundamental component of any system including a gesture sensor.

Automatic detection of gestures in the RF-based proximity system [21] can be reliably achieved by a simple algorithm detecting when the strength of the proximity signal is above a predefined threshold. Such simplicity of the gesture detection is an advantage of the RF sensor.

Use of inertial sensors in [20] demonstrated a signal pattern associated with a bite that can be detected by more sophisticated signal analysis. The comparison of orientation of the wrist when picking up the food with the orientation of the wrist when putting the food into the mouth found that a roll of the wrist occurred in most cases independently of the utensils or fingers used. Based on this motion pattern, a pattern-recognition algorithm was developed and implemented in a microcontroller to automatically detect bites. The algorithm looked at the velocity of the rolling motion measured by the sensor and at the minimum time between the two rolls present in a single bite. The methodology was tested under laboratory settings to determine the accuracy of bite detection.

Results showed a sensitivity of 94% for "bite" counting during controlled meals and of 86% during uncontrolled meals.

Neither of the described hand gesture detection algorithms can automatically differentiate between food-related and non-food-related hand gestures (e.g., using a napkin, head scratching, etc.) and are thus incapable of directly detecting food intake. Although food-intake episodes may potentially have a relatively high density of hand-to-mouth gestures, the experimental data show a high incidence of hand gestures during the waking day that does not allow simple differentiation based on the rate of gesture occurrence. Nevertheless, hand gestures carry important information that can be used in multi-sensor food-intake detection.

3.3 Food-Intake Detection from Chewing

With a few exceptions, chewing is a reliable indicator of solid food intake. Therefore, many methods attempt to detect food intake through computerized recognition of chewing.

Automatic recognition of chewing was achieved by processing the signal from a piezoelectric strain sensor by a pattern-recognition algorithm [33]. Analysis of the jaw motion signals in the frequency domain showed a strong signal in the frequency range between 1.25 Hz to 2.5 Hz, which results from the rhythmic up-and-down and side-to-side movements of the lower jaw during chewing. Such characteristic jaw movements are absent or less marked during inactivity, talking, or walking. Frequency and time domain features extracted from the strain sensor signal were used to train a subject-independent classifier based on support vector machines (SVM). Results indicated that the models were able to discriminate between chewing and other activities with an averaged accuracy of around 81% and a time resolution of 30 s. Later analysis [51] further improved the recognition results by incorporating features extracted from the voice frequency band in the range of 100 to 250 Hz.

Similar attempts have been performed in automatic recognition of the chewing sounds. Signal processing and pattern-recognition algorithms have been developed to discriminate between chewing sounds and sounds not related to eating or drinking (i.e., speech, ambient noise, etc.). One of the earlier studies implemented a naïve Bayes classifier trained with features from the frequency domain [31]. The pattern-recognition system achieved an overall food-intake recognition rate of 86.6% averaged across two subjects that completed a total of 375 chewing sequences, most of them in laboratory settings. A later analysis of the chewing sound revealed that most of energy is found in the frequency range below 4 kHz [32] and that no characteristic frequency with high energy content can be used to distinguish the chewing process from other activities, thus requiring noise cancelation techniques to reliably detect chewing sounds. Using the signals from the sensor shown in [52], chewing detection was performed by a signal processing algorithm developed to compute the energy of the acoustic signals from each microphone. The ratio between energy values was compared to an adaptive threshold value to discriminate food-intake sounds with an overall accuracy of 85%.

Overall, chewing is a fairly reliable indicator of food intake that can be accurately detected by non-intrusive wearable sensors. This factor makes recognition of chewing one of the more popular approaches to detection of food intake. At the same time, chewing is not a perfect predictor of food intake. Most of liquids are not being chewed during ingestion and thus cannot be detected through chewing. Many semi-liquids (e.g., yogurt) and even some of the solid foods (e.g., fufu) are also not being chewed during consumption. At the same time, there are several non-food items that may be chewed continuously but not swallowed (e.g., chewing gum, chewing tobacco, betel, etc.) thus potentially triggering a false detection. Therefore, food-intake detection from chewing may need additional indicators of food intake (such as swallowing or hand gestures) for reliable detection.

3.4 Food Intake Detection from Swallowing

Swallowing is potentially the most reliable indicator of food intake as all of the ingested food is eventually swallowed. Since spontaneous swallowing happens naturally during the waking hours, detection of food intake from swallowing can be a two-step process: first, individual swallows are recognized and, second, food intake is detected from the time series of swallows. This section presents methods developed for both swallowing recognition and food-intake detection from swallowing.

Several studies performed automatic recognition of swallowing events using accelerometers [53,54]. However, they were mostly focused on establishing objective criteria for detecting swallowing disorders rather than on detection of food intake. Pattern-recognition algorithms were implemented to discriminate normal swallows, dysphagic swallows, and artifacts showing promising results that can potentially be extended to detect food-intake swallows [55−57].

A two-step methodology for automatic detection of food intake based on the acoustical swallowing sensor data was recently presented. The first step used Mel-scale Fourier spectrum features and SVM signal classification to recognize individual swallowing events [58] with average accuracy of 84.7% on a dataset containing more than 64 hours of acoustic data and about 10,000 swallows from 20 subjects. Results indicated that the highest accuracy of identifying individual swallowing events was observed during quiet periods of no food intake (88%) and the lowest accuracy was observed during food-intake periods, where talking and/or background noise was present (82.9%). The swallowing recognition model was further improved by incorporating principal component analysis (PCA) and a smoothing algorithm to improve recognition rates for intra- and inter-subject models [59].

The second step performed the detection of food intake based on the frequency of swallowing. The instantaneous swallowing frequency (*ISF*) averaged over a sliding window of 30 s duration (*EISF*) was used as the key predictor for discriminating between the "intake" and "no intake" classes. The best subject-independent model used a decision threshold $T_{FL}{}^{\mathrm{INGEST}}$ obtained by multiplying the floating average of the swallowing frequency over several epochs with a scaling factor α. The resulting model had the advantage of self-adjusting to individual variations in the swallowing rates. Training and validation of this model demonstrated an average detection accuracy of 87% [39].

Another study used the time sequence of swallows as the main predictor of food-intake activity [60]. The absolute difference in time between a swallow occurring at time t_i and d neighboring swallows was used as features to create food-intake detection models.

A supervised machine learning technique (SVM) was used to create a subject-independent model whereas an unsupervised clustering technique (K-means) was used to create individual models. The swallowing data used to train the models was collected from 18 subjects and contained 4,045 spontaneous saliva swallows and 5,811 food-intake swallows. Results showed that unsupervised individual models presented a better performance than the subject-independent model (93.9% vs. 89.9%), most likely due to ability of unsupervised methods to adapt to individual traits in the data.

Food-intake detection from swallowing has also been studied in conjunction with an electroglottograph sensor and a one-step approach that does rely on recognition of individual swallowing events [48]. A pattern-recognition algorithm was developed to detect food-intake based on the changes in electrical impedance across the larynx. Signals captured by the EGG device were divided into non-overlapping windows of 30 seconds and features were extracted by performing wavelet decomposition. Subject-independent food-intake detection models were trained with these features and artificial neural networks. The models were created using data from a study involving 30 subjects performing unrestricted consumption of four meals each under laboratory settings. The study also included the monitoring of swallowing sounds through a miniature microphone. The performance of the EGG method was compared to an acoustic method using the same one-step algorithm. Results showed a statistically significant difference between the average food-intake detection accuracy for the EGG-based method (90.1%) and the acoustic-based method (83.1%).

In summary, food-intake recognition from swallowing achieves results similar to those from chewing. Also, similarly to chewing, the fundamental challenge of the signal processing and pattern-recognition is differentiation of swallowing and/or food intake from other activities of daily living. Swallowing is present in consumption of all foods: liquids, semi-liquids, and solids, and thus deserves to be one of the prime predictors of food intake.

4. METHODS FOR CHARACTERIZATION OF FOOD INTAKE

Successful detection of food intake is the first step for understanding the ingestive behavior. Once the fact of food intake is established, the process of ingestion and the consumed meal need to be characterized. Particularly, the number and type of food items in a meal, the mass and volume of ingestion, caloric and nutritional content of a meal, and the rate of ingestion can be estimated by analysis of the sensor data. This section presents a review of several approaches aimed at food-intake characterization.

4.1 Recognition of Number of Foods in a Meal

A methodology for recognizing the number of foods consumed is an important first step in characterization of food intake. As with other aspects of technology for food-intake monitoring, the available methods either rely on analysis of imagery or on analysis of physiological signals collected by wearable sensors.

Image processing methods attempt to automatically segment the image into regions that may be attributed to different foods, and potentially identify the food based on image features [61,62]. The recognition procedure starts by determining the region of interest, where the food may be located on an image. This is typically done by detecting round objects such as plates, bowls, or glasses and analyzing the regions inside of the enclosing circle. The next step of the procedure is the multi-scale image segmentation (such as normalized cut [63]) that uses both coarse and fine details to identify image segments that may contain food items. After the segmentation, color and texture features extracted from the image segment are used to recognize a particular food item that may be represented by this segment. The classification may be performed by an SVM or some other robust classifier. Since an image segment may not contain the entire food item, the next step of the procedure merges the segments that are in spatial proximity to each other and have the same class label (belong to the same food item). The outcome processing is the number of food items found in the image and the food type of each food item. While image segmentation is capable of estimating the number of foods in a meal, the authors of [61,62] quantified their approach in terms of accuracy in recognizing specific food types, which varied between 56% for 19 food items [62] to 44% for 32 food items [61].

Use of physiological signals to quantify the number of foods in a meal can be illustrated by a method based on monitoring of chewing and swallowing [64]. The fundamental assumption for the proposed method is that food of different physical properties result in varying chewing and swallowing patterns, allowing differentiating food items without prior knowledge of quantity of the food items consumed. To demonstrate feasibility of this approach, chewing and swallowing data was collected from 17 subjects participating in experiments involving the consumption of five different foods: cheese pizza, yogurt, apple, peanut butter sandwich, and water. Three different types of features were used to segment a meal into different food items. The first feature was the location in time of each swallow, which was important for grouping swallows associated to a certain food type. The second feature was the time to preceding swallow (TPS), which indicated the time difference between a swallow occurring at time t_i and the previous swallow occurring at time t_{i-1}. The last feature was the number of chews preceding a swallow (CPS), which was an indication of the number of chews observed between two consecutive swallows.

Since chewing and swallowing patterns for a given food may vary from individual to individual, the collected data was analyzed with two different unsupervised clustering techniques to identify groups of food within a meal: affinity propagation (AP) and agglomerative hierarchical clustering (AHC). The use of unsupervised learning enabled adaptation to individual traits while preserving the generality of the approach. Results showed that an overall accuracy of 95% was obtained when estimating the number of foods using a model created with AHC technique. On the other hand, an accuracy of 90% was obtained for a model created with AP technique. A limitation of the proposed clustering approach was that the food items were consumed in a predefined sequence and unmixed to eliminate the uncertainty caused by inter-food variation. Further studies are needed to test the clustering algorithm's performance on unrestricted meals with larger variability of food items. However, this methodology for food clustering should be directly applicable to unrestricted meals as food intake is cumulative over time and the consumption normally happens sequentially in a bite-by-bite manner.

Non-invasive monitoring of chewing sounds is another alternative for recognizing the number of foods in a meal. Recognition of the different food consumed in a meal may also be possible as the chewing of foods with different properties (i.e., moisture, crispiness, crunchiness, etc.) may produce characteristic sounds. By means of pattern-recognition techniques, foods with similar chewing sound patterns can potentially be grouped into categories containing foods with similar physical properties. This approach has been studied in [32], where 51 participants consumed seven different types of solid foods and a drink. The chewing sounds were monitored by a miniature microphone inserted into the ear canal, with hidden Markov models recognizing chewing and swallowing events. The chewing sound recognition achieved 79% classification accuracy on the food items tested. These results indicate that analysis of chewing sound is likely to be able to detect the number of foods in the meal, as long as these foods possess distinct physical properties and are consumed with chewing.

4.2 Estimation of Ingested Mass and Energy Intake from Imagery

The human brain can rapidly discern the energy content of foods by a simple visual inspection of different items in a process that involves object categorization, reward assessment, and decision making [65]. In an attempt to mimic this process, image-processing techniques have been used to estimate the food's portion size from imagery, which is then used to calculate the amount of energy intake by using a food database.

The estimation of the portion size from a picture is a difficult task due to the lack of distance information in a typical image. The same object may look bigger or smaller depending on the distance to camera, camera angle, and focal length of the camera. A typical solution to this problem is use of dimensional referents (fiducial marks) in each picture that allow the estimation of the food volume by computer algorithms. Several approaches have been investigated that included different types of referents, such as a checkered tablecloth, calibration cards, checkerboards, and circular plates of known dimensions [15,66,67]. Another option is to use lights from LEDs or laser diodes to produce a spotlight in the field of view when the picture is taken [68]. In most cases, a single picture was used to estimate food's volume and volumes computed from before and after the meal pictures were used to estimate amount of food consumed.

Different approaches have been proposed to estimate food's volume from a single picture [15,68,69]. One approach evaluated different algorithms to determine the location and orientation of an object plane based on the intrinsic parameters of the camera [68]. The results obtained were used to measure dimensional variables, such as length and thickness, by selecting various feature points in the food image. The values obtained for these variables provided an estimate of the food volume. Another approach evaluated different image-processing algorithms to segment (separate) the food items from the background of the picture [15]. The segmented image was used as the input of a process involving camera calibration and 3-D volume reconstruction that automatically estimated the volume of foods. One of the recent studies [70] used image processing techniques to automatically estimate portion weights from the mobile phone images. For the set of 19 different foods, the ratio of predicted weight to known weight varied from 0.89 to 4.61, while for a smaller subset of nine foods the ratio was in the range of 0.8 to 1.2. The largest errors were observed for foods such as lettuce, French fries, and garlic bread, while the smallest errors were obtained for strawberry jam, milk, orange juice, and cheeseburger sandwich. Use of 3-D models that fit the shape of the 2-D food image to a pre-constructed shape such as a cylinder, etc., has been reported to produce more accurate results (Figure 7). In [71], the error of estimating volume for 17 different foods was reported as 3.7%, which may be considered as highly accurate. Comparable results (7.2% average error for five food items) were reported in [72] for another 3-D-model-based method.

Although these methods show promising results towards the automatic estimation of portion sizes, food portion estimation from imagery needs further development. Computer algorithms can use images to estimate the volume of food consumed; however, the research combining automatic weight estimation with food item recognition for automatic energy content estimation is yet to emerge. Indeed, there are many challenges to overcome both in improving the accuracy of volume and mass estimation and in accurate identification and energy density estimation for the great variety of foods that we may find on the table. As a potential solution, food-intake characterization from imagery can be implemented in a semi-automatic process, where a human nutritionist visually identifies each food consumed and the computer algorithm automatically estimates the volume of food and calculates calories and nutrient values from the food database. A promising new direction is use of 3-D cameras or camera motion to assess the volume of food.

4.3 Estimation of Ingested Mass and Energy Intake from Counts of Chews, Swallows, and Hand Gestures

Characterization of food intake has also been studied using metrics derived from chewing, swallowing, and hand-to-mouth gestures. Mathematical models were proposed to estimate mass and energy consumed in a meal using counts of chews and swallows captured by wearable sensors.

In the approach presented in [39], individualized (subject-dependent) linear models were developed based on the hypothesis that chews and swallows can estimate the total mass of food ingested with acceptable accuracy. A model for estimation of the total

FIGURE 7 Example of food's volume estimation using 3-D models. *(Image courtesy of Dr. Mingui Sun [71].)*

ingested mass of solid food used the total number of chews and swallows within a period of ingestion as predictors was

$$M_S = \frac{1}{2}(\overline{M}_{sw}^S \times N_{sw}^S + \overline{M}_{chew} \times N_{chew})$$

where $\overline{M}_{sw}^S$ was subject's average mass per swallow of solid food, N_{sw}^S was the total number of swallows of solid foods, $\overline{M}_{chew}$ was the subject's average mass per chew, and N_{chew} was the total number of chews. The model to estimate the mass of liquids ingested used only the number of swallows as liquid consumption does not involve chewing. Thus, the mass of liquids was calculated as

$$M_L = \overline{M}_{sw}^L \times N_{sw}^L$$

where $\overline{M}_{sw}^L$ was the subject's average mass per swallow of liquid and N_{sw}^L was the total number of swallows of liquid. The values of the parameters $\overline{M}_{sw}^S$, $\overline{M}_{sw}^L$, and $\overline{M}_{chew}$ were statistical estimates of the average mass consumed by an individual in a swallow or a chew. Using individualized instead of population-based models is required due to the high inter-subject variability of the mass-per-chew and mass-per-swallow parameters [39]. These models were implemented and evaluated on data from 16 subjects consuming five different food items in a sequential manner. Results showed that the mass estimation model achieved an average accuracy of 91.8% for solid-food intake and an average accuracy of 83.8% for liquid intake.

In further development of this approach by the same investigator team, a similar method was developed to estimate the total amount of energy consumed in an unrestricted meal [24]. This new method involved two steps. First, the amount of mass ingested was estimated for each food item using the individual models based on chews and swallows. Second, the energy intake was estimated by multiplying the estimated mass by the caloric density (CD) of each consumed food, which was extracted from a nutritional analysis of food imagery by a qualified nutritionist. The total energy consumed was calculated by summation of energy estimates for each food ingested. Prediction models were developed based on training and validation meals. The training meals consisted of three meals of identical size and content that were consumed in three separate visits. Counts of chews and swallows for the training meals were used to estimate the model parameters $\overline{M}_{sw}^S$, $\overline{M}_{sw}^L$, and $\overline{M}_{chew}$. The validation meal was a new meal selection that included different solid foods and drinks from the training meal. The prediction models in this experiment were evaluated in a more realistic scenario of food intake where the dataset included information from a wide variety of foods (a total of 45 different foods) and assumed no restriction in the way the food was consumed. The performance of the models was compared to diet diaries and photographic food records to determine how well the estimation of energy intake matched the estimations obtained with weighed food records (gold-standard method).

Results showed that the models estimated the energy consumed in the training meals with a reporting error of 15.83%, which was significantly lower than the reporting errors obtained for the other methods (27.86% for diet diaries and 19.95% for photographic food records). Additionally, the proposed models presented the lowest reporting bias

(−8.6 kCal compared to −60 kCal for diet diary and 83.6 for photographic records). On the other hand, estimation of energy intake for the validation meals was not significantly different from either diet diaries or photographic methods. This was caused by an increase in the reporting error for the validation meal probably due to the different physical properties of the food items consumed during the training and the validation meals. Additionally, in the validation stage, it was observed that the accuracy of the self-report methods was very high given that this was just a single meal in a controlled setting. As the accuracy of self-report tends to decrease with the duration of the recordings [7] (due to either underestimation or subjects forgetting to take pictures of food), it is reasonable to think that, over several days of observation, energy estimation models based on chews and swallows may produce better results than self-report as they do not rely on subject participation.

The use of chewing and swallowing for mass estimation has been also explored in the sensors relying on acoustical detection of chewing sound. In [73], bite weight prediction was attempted based on four features extracted from the chewing sounds. On a test with three foods (potato chips, lettuce, and apple), the reported error varied between 19% to 31% for individual, subject-calibrated models, and between 41% to 62% for subject-independent models. Even though these results were obtained on a small set of foods and smaller population (eight subjects), they largely confirm the individuality of mass ingestion as expressed in the counts of chews.

5. APPLICATIONS

5.1 Laboratory Vs. Free-Living Monitoring

Under laboratory conditions, monitoring of food intake through paper-based methods, electronic diaries, or wearable sensors may provide reasonable agreement between the food records and the actual amounts eaten. However, laboratory research presents significant deficiencies when compared to free-living research [74]. Most of these deficiencies are caused by real-world variability that directly affects the eating behavior of individuals and is usually missing in laboratory settings. Food selection, timing of meals (i.e., determined by external schedules, work, or social commitments), food-intake environments, and food-intake behavior may vary dramatically day-to-day and between individuals in free living, while lab experiments typically are severely constrained. Additionally, important changes in the eating behavior of an individual during the course of a day and over different days (weekdays vs. weekends) occurring in free living may be controlled or eliminated in laboratory experiments [75]. Finally, the importance of variables can be overestimated and important effects can be missed as laboratory experiments are usually short in duration. Consequently, monitoring of eating behavior of individuals under free-living conditions is far more challenging than laboratory experiments.

Monitoring of food intake in the real world and over extended periods of time generates more complex datasets than monitoring in the laboratory. Methodologies developed for automating the process of food-intake detection should be able to handle the inherent intra- and inter-subject variability. The intra-subject variability is increased by the different activities performed by an individual over the course of a day (i.e., walking, talking, eating, sleeping, working, etc.). The inter-subject variability is reinforced by the diversity of the population, which, in some cases, may have different eating patterns and lifestyles. These issues make it difficult to generalize food-intake models created with laboratory data to free-living data while maintaining an acceptable performance. Ideally, methods of food-intake detection and characterization should be created using innovative methodologies based on free-living data.

5.2 Wearable Devices for Free-Living Monitoring

A novel wearable device, the Automatic Ingestion Monitor (AIM), has been developed and evaluated for objective monitoring of food intake under free-living conditions (Figure 8) [24]. AIM presented three major benefits over self-reported intake. First, AIM is a wearable device that has the ability to monitor 24 hours of ingestive behavior without relying on self-report or any other actions from subjects. AIM wirelessly integrated three different sensor modalities for an accurate monitoring: a jaw motion sensor to monitor chewing, a proximity sensor to monitor hand-to-mouth gestures, and an accelerometer to

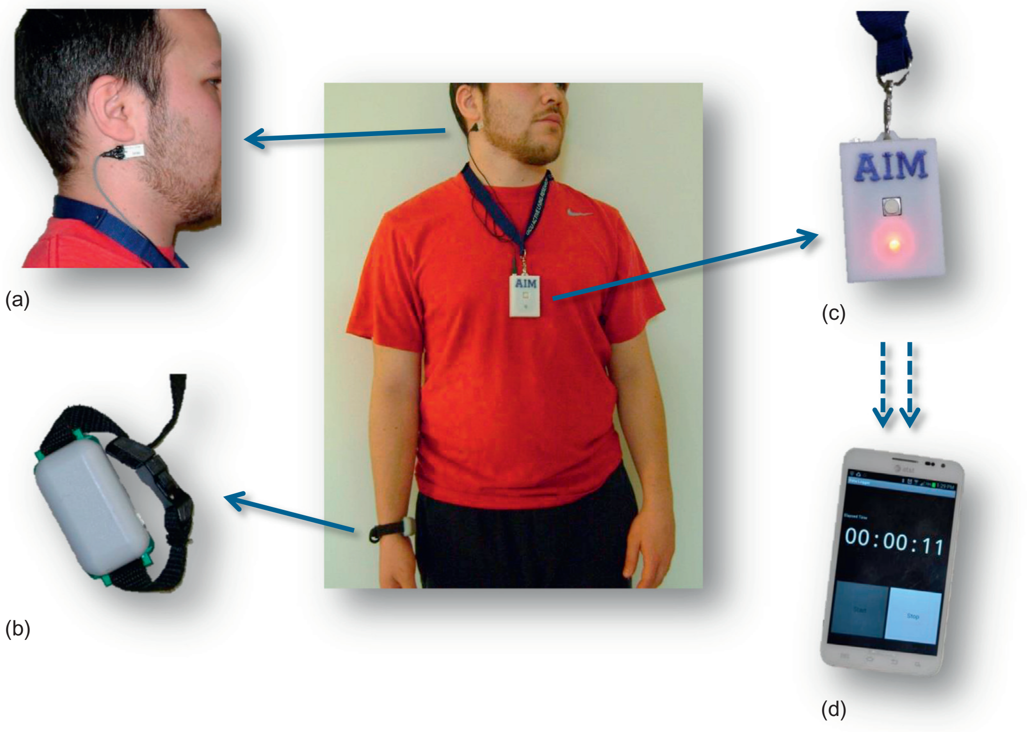

FIGURE 8 The Automatic Ingestion Monitor (AIM) consisting of four main parts: (a) the jaw motion sensor, (b) the wireless module, (c) the proximity sensor, and (d) the smartphone.

monitor body motion. Second, AIM is able to reliably detect food-intake episodes in the presence of real-life artifacts using a robust pattern-recognition methodology for detection of food intake. The detection methodology contains several steps, such as sensor information fusion, feature extraction, and classification. The sensor fusion step removes portions of the signal that cannot be food intake based on statistically derived rules. For example, it is highly uncommon to eat solid foods during moderate to vigorous exercise, or during sleep. Both of these activities (exercise and sleeping) can be reliably detected from the accelerometer signal and corresponding signal intervals not included into further consideration for food-intake detection. The feature extraction step computes a number of time, frequency, and time-frequency domain features from the sensor signals. Food-intake detection is based on an artificial neural network implementing a subject-independent classification model that requires no individual calibration. Third, the AIM device and food-intake detection methodology were validated in an objective study where an average food-intake recognition rate of 89.8% was achieved. Individuals with origins from five different countries and having different lifestyles and ingestive behaviors participated in the validation study. They wore AIM in free living during 24 hours without any restrictions on their eating behavior and activities.

The results of the validation study revealed that AIM can potentially provide an accurate prediction of the food-intake episodes occurring over the course of a day in a free-living population. However, several questions remain to be answered. One question is related to the capability of AIM for detecting liquid intake. In the validation study, the results showed the recognition rate for solid food intake only. Previous studies suggest that certain intake of liquids (such as gulping large quantities of a drink) may be detected through the monitoring of jaw motion [33] while others, such as sipping, may be undetected. Another question is related to the acceptance of the device by subjects. AIM was designed as a pendant device worn on a lanyard around the neck, which intended to satisfy the need for a socially acceptable device; further miniaturization of the device is needed to make it less obtrusive. Finally, although the food-intake detection was performed offline, the ultimate goal of AIM is to perform real-time recognition and characterization of food intake and to deliver feedback about an individual's intake behavior.

6. SUMMARY AND CONCLUSIONS

Detection and categorization of food intake is a difficult task given the variability and complexity of human food-intake behavior. New methodologies are constantly being developed attempting to achieve objective and accurate measurements of ingestion. The current trend is to replace inaccurate methods relying on self-report by new, less burdensome methods that require minimal participation of subjects.

Historically, portable electronic food diaries first facilitated the reporting task by allowing individuals to simply search and log the consumed foods on mobile phones or PDAs. The electronic diaries evolved to include image capture to further improve assessment of food intake. Complex image processing algorithms are being developed to facilitate image analysis. However, even modern food diaries and food imagery require active participation of the subjects in reporting their intake, thus misreporting and underreporting may still affect the measurements.

Development of body-worn sensors for food-intake detection and characterization has gained substantial interest in recent years. From the functional perspective, the approaches presented in this chapter show that wearable sensors such as microphones, accelerometers, and piezoelectric and magnetic sensors are capable of capturing information about events occurring during the eating process (hand-to-mouth gestures, bites, chewing, and swallowing). Most of the sensor modalities presented in this chapter also respond to excitations from the sources other than the physiological processes of interest, thus requiring sophisticated signal processing and pattern-recognition algorithms to separate artifacts from food intake. The future in development of wearable sensors for monitoring of ingestion lies in further miniaturization and advancements in reliability and functionality, especially with the focus on characterization of food intake.

The social acceptability of wearable devices is another open issue that needs to be carefully addressed in future research. Most of the current research focuses mainly on assessing the functionality of the device and does not present comprehensive analysis on the sensor's burden. Wearable sensors for monitoring ingestive behavior should be non-invasive, unobtrusive, comfortable, and should minimally impact the way people eat. If some of these requirements are missing, people would tend to be non-compliant and abandon the device. Moreover, if the ingestive behavior is significantly altered by the presence of the sensors, then the benefits of using such sensors may be greatly reduced.

Computer algorithms process signals captured by the sensors to detect and discriminate ingestion episodes from other activities. In the majority of the approaches food intake is detected from recognition of chewing or swallowing, potentially in combination with each other or with hand gestures. To date, no published research has reported on the use of hand gestures alone for detection of food intake. Automatic recognition of jaw motion during chewing or chewing sounds translates into 80% to 90% food-intake detection accuracy in most of the reported studies, typically performed under laboratory conditions. Similar accuracies have been reported for swallowing-based approaches. The variability of free-living behaviors may negatively impact the accuracy of these methods, and thus most of them need a thorough testing in the real world. Another major consideration is use of subject-dependent (individual) models vs. subject-independent (group) models. For example, acoustical recognition of swallowing demands an individual model as the swallowing sound varies from individual to individual. On the other hand, jaw motion during chewing exhibits lower variability and thus is easier to use in a group model. Also, use of unsupervised machine learning methods may help with adapting to individual traits in the sensor signals. Finally, the computational demands of the algorithms should also be taken into consideration. Many of the reported algorithms rely on computationally intensive time-frequency decompositions (e.g., wavelets) and classifiers (e.g., support vector machines) that may be unsuitable for limited computational resources of a wearable sensor. Future research should strive to develop robust, subject-independent, and computationally efficient methods of food-intake detection and characterization.

Food-intake characterization is still one of the biggest challenges facing eating behavior research. The diversity of foods makes it problematic to automatically and accurately capture energy intake through wearable sensors. Although many complex algorithms have been investigated, obtaining accurate and reliable algorithms for estimating energy intake is an active topic of research. Multi-modal fusion and merging of information acquired from different sources, such as physiological sensor and food imagery, is a promising approach in improving the accuracy of the mass, caloric, and nutritional intake estimates. New sensor modalities that directly or indirectly measure the nutritional content of food, possibly through intra-oral sensing, may need to be developed to improve reliability of such estimates. Similarly to food-intake detection, use of group, rather than subject-dependent models, is an important consideration in the algorithm development.

Given the early age of wearable sensors, most of the studies to date have relied on laboratory experiments to test the feasibility of suggested approaches. Food-intake detection and characterization becomes substantially more complex in free living due to the presence of artifacts originated by real-world activities that are controlled or missing in a laboratory. Food intake in free living is also more variable and complex than any laboratory or cafeteria study. Therefore, it is critical to test any proposed methodologies in realistic conditions of free living, and the expectation is that future research will reveal more and more of such experiments.

In summary, ongoing research on use of wearable sensor technology for study of ingestive behavior is addressing the need for innovative methodologies for dietary assessment. The development of wearable devices that integrate different non-invasive sensor modalities seems to be the main route selected by researchers towards an objective and accurate monitoring of food intake. However, a substantial number of challenges have to be resolved before a practical, accurate, and non-invasive monitoring of food intake becomes reality. A successful implementation of such devices and methodologies would have tremendous impact on the population as it may help to study and correct ingestive behaviors associated with obesity and eating disorders.

References

[1] K.M. Flegal, M.D. Carroll, B.K. Kit, C.L. Ogden, Prevalence of Obesity and Trends in the Distribution of Body Mass Index Among US Adults, 1999-2010, JAMA J. Am. Med. Assoc. 307 (5) (2012) 491−497.

[2] S.J. Olshansky, D.J. Passaro, R.C. Hershow, J. Layden, B.A. Carnes, J. Brody, et al., A potential decline in life expectancy in the United States in the 21st century, N. Engl. J. Med. 352 (11) (2005) 1138−1145.

[3] C.G. Fairburn, P.J. Harrison, Eating disorders, Lancet 361 (9355) (2003) 407−416.

[4] J.I. Hudson, E. Hiripi, H.G. Pope Jr, R.C. Kessler, The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication, Biol. Psychiatry 61 (3) (2007) 348−358.

[5] D. Sánchez-Carracedo, D. Neumark-Sztainer, G. López-Guimerà, Integrated prevention of obesity and eating disorders: barriers, developments and opportunities, Public Health Nutr. 15 (12) (2012) 2295−2309.

[6] F.E. Thompson, A.F. Subar, Dietary assessment methodology, Nutrition in the Prevention and Treatment of Disease, second ed., Academic Press, San Diego, CA, 2008.

[7] A.H.C. Goris, E.P. Meijer, K.R. Westerterp, Repeated measurement of habitual food intake increases under-reporting and induces selective under-reporting, Br. J. Nutr. 85 (05) (2001) 629−634.

[8] A.H. Goris, M.S. Westerterp-Plantenga, K.R. Westerterp, Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake, Am. J. Clin. Nutr. 71 (1) (2000) 130−134.

[9] S.D. Poppitt, D. Swann, A.E. Black, A.M. Prentice, Assessment of selective under-reporting of food intake by both obese and non-obese women in a metabolic facility, Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes 22 (4) (1998) 303−311.

[10] M.-J. Park, H.-S. Kim, K.-S. Kim, Cellular phone and Internet-based individual intervention on blood pressure and obesity in obese patients with hypertension, Int. J. Med. Inf. 78 (10) (2009) 704−710.

[11] S.-I. Kim, H.-S. Kim, Effectiveness of mobile and internet intervention in patients with obese type 2 diabetes, Int. J. Med. Inf. 77 (6) (2008) 399–404.
[12] NDL/FNIC Food Composition Database Home Page. [Online]. Available: <http://ndb.nal.usda.gov/>. (Accessed: 02.07.14)
[13] C. Johannes, J. Woods, S. Crawford, H. Cochran, D. Tran, B. Schuth, Electronic versus paper instruments for daily data collection, Ann. Epidemiol. 10 (7) (2000) 457.
[14] C.K. Martin, H. Han, S.M. Coulon, H.R. Allen, C.M. Champagne, S.D. Anton, A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method, Br. J. Nutr. 101 (03) (2009) 446–456.
[15] F. Zhu, M. Bosch, I. Woo, S.Y. Kim, C.J. Boushey, D.S. Ebert, et al., The Use of Mobile Devices in Aiding Dietary Assessment and Evaluation, IEEE J. Sel. Top. Signal Process. 4 (4) (2010) 756–766.
[16] D.A. Williamson, H.R. Allen, P.D. Martin, A.J. Alfonso, B. Gerald, A. Hunt, Comparison of digital photography to weighed and visual estimation of portion sizes, J. Am. Diet. Assoc. 103 (9) (2003) 1139–1145.
[17] J.A. Higgins, A.L. LaSalle, P. Zhaoxing, M.Y. Kasten, K.N. Bing, S.E. Ridzon, et al., Validation of photographic food records in children: are pictures really worth a thousand words? Eur. J. Clin. Nutr. 63 (8) (2009) 1025–1033.
[18] M. Popa, Hand gesture recognition based on accelerometer sensors, 2011 7th Int. Conf. Netw. Comput. Adv. Info. Manage. (NCM) (2011) 115–120.
[19] H. Junker, O. Amft, P. Lukowicz, G. Tröster, Gesture spotting with body-worn inertial sensors to detect user activities, Pattern Recognit. 41 (6) (2008) 2010–2024.
[20] Y. Dong, A. Hoover, J. Scisco, E. Muth, A new method for measuring meal intake in humans via automated wrist motion tracking, Appl. Psychophysiol. Biofeedback 37 (3) (2012) 205–215.
[21] P. Lopez-Meyer, Y. Patil, T. Tiffany, E. Sazonov, Detection of Hand-to-Mouth Gestures Using a RF Operated Proximity Sensor for Monitoring Cigarette Smoking, Open Biomed. Eng. J. 9 (2013) 41–49.
[22] J.M. Fontana, M. Farooq, E. Sazonov, Estimation of Feature Importance for Food Intake Detection Based on Random Forests, 2013 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.(EMBC) (2013).
[23] R. Xu, S. Zhou, W.J. Li, MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition, IEEE Sens. J. 12 (5) (2012) 1166–1173.
[24] J.M. Fontana, M. Farooq, E. Sazonov, Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior, IEEE Trans. Biomed. Eng. 61 (no. 6) (2014) 1772–1779 <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload = true&arnumber = 6742586>. (Accessed 02.07.14)
[25] K. Fueki, T. Sugiura, E. Yoshida, Y. Igarashi, Association between food mixing ability and electromyographic activity of jaw-closing muscles during chewing of a wax cube, J. Oral Rehabil. 35 (5) (2008) 345–352.
[26] K. Kohyama, E. Hatakeyama, T. Sasaki, T. Azuma, K. Karita, Effect of sample thickness on bite force studied with a multiple-point sheet sensor, J. Oral Rehabil. 31 (4) (2004) 327–334.
[27] V.A. Bousdras, J.L. Cunningham, M. Ferguson-Pell, M.A. Bamber, S. Sindet-Pedersen, G. Blunn, et al., A novel approach to bite force measurements in a porcine model in vivo, Int. J. Oral Maxillofac. Surg. 35 (7) (2006) 663–667.
[28] P.J. Lillford, The Materials Science of Eating and Food Breakdown, MRS Bull. 25 (12) (2000) 38–43.
[29] J. Nishimura and T. Kuroda, "Eating habits monitoring using wireless wearable in-ear microphone," in 3rd International Symposium on Wireless Pervasive Computing, 2008. ISWPC 2008, 2008, pp. 130–132.
[30] M. Shuzo, S. Komori, T. Takashima, G. Lopez, S. Tatsuta, S. Yanagimoto, et al., Wearable Eating Habit Sensing System Using Internal Body Sound, J. Adv. Mech. Des. Syst. Manuf 4 (1) (2010) 158–166.
[31] O. Amft, A wearable earpad sensor for chewing monitoring, 2010 IEEE Sens. (2010) 222–227.
[32] S. Päßler, M. Wolff, W.-J. Fischer, Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food, Physiol. Meas. 33 (6) (2012) 1073–1093.
[33] E. Sazonov, J.M. Fontana, A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing, IEEE Sens. J. 12 (5) (2012) 1340–1348.
[34] A.R. Doherty, S.E. Hodges, A.C. King, A.F. Smeaton, E. Berry, C.J.A. Moulin, et al., Wearable Cameras in Health: The State of the Art and Future Possibilities, Am. J. Prev. Med. 44 (3) (2013) 320–323.
[35] G. O'Loughlin, S.J. Cullen, A. McGoldrick, S. O'Connor, R. Blain, S. O'Malley, et al., Using a Wearable Camera to Increase the Accuracy of Dietary Analysis, Am. J. Prev. Med. 44 (3) (2013) 297–301.

[36] J. Liu, E. Johns, L. Atallah, C. Pettitt, B. Lo, G. Frost, et al., An Intelligent Food-Intake Monitoring System Using Wearable Sensors, 2012 Ninth Int. Conf. Wearable Implantable Body Sens. Netw.(BSN) (2012) 154−160.

[37] C.S.C. Lear, J.B. Flanagan Jr., and C.F.A. Moorrees, The frequency of deglutition in man Arch. Oral Biol., 10, 1, pp. 83−99, IN13−IN15.

[38] M. Pehlivan, N. Yüceyar, C. Ertekin, G. Çelebi, M. Ertaş, T. Kalayci, et al., An electronic device measuring the frequency of spontaneous swallowing: Digital Phagometer, Dysphagia 11 (4) (1996) 259−264.

[39] E. Sazonov, S.A.C. Schuckers, P. Lopez-Meyer, O. Makeyev, E.L. Melanson, M.R. Neuman, et al., Toward Objective Monitoring of Ingestive Behavior in Free-living Population, Obesity 17 (10) (2009) 1971−1975.

[40] I. Lichter, R.C. Muir, The pattern of swallowing during sleep, Electroencephalogr. Clin. Neurophysiol. 38 (4) (1975) 427−432.

[41] W.J. Dodds, The physiology of swallowing, Dysphagia 3 (4) (1989) 171−178.

[42] T.A. Hughes, P. Liu, H. Griffiths, B.W. Lawrie, C.M. Wiles, Simultaneous electrical impedance tomography and videofluoroscopy in the assessment of swallowing, Physiol. Meas. 17 (2) (1996) 109−119.

[43] D.S. Cooper, A.L. Perlman, Electromyography in the functional and diagnostic testing of deglutition, Deglutition Its Disord. Anat. Physiol. Clin. Diagn. Manag. (1996) 255−285.

[44] N.P. Reddy, A. Katakam, V. Gupta, R. Unnikrishnan, J. Narayanan, E.P. Canilang, Measurements of acceleration during videofluorographic evaluation of dysphagic patients, Med. Eng. Phys. 22 (6) (2000) 405−412.

[45] E. Sazonov, S. Schuckers, P. Lopez-Meyer, O. Makeyev, N. Sazonova, E.L. Melanson, et al., Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior, Physiol. Meas. 29 (5) (2008) 525−541.

[46] J.L. Schultz, A.L. Perlman, D.J. VanDaele, Laryngeal movement, oropharyngeal pressure, and submental muscle contraction during swallowing, Arch. Phys. Med. Rehabil. 75 (2) (1994) 183−188.

[47] S. Nozaki, J. Kang, I. Miyai, T. Matsumura, Electroglottographic evaluation of swallowing in Parkinson's disease, Rinshō Shinkeigaku Clin. Neurol 34 (9) (1994) 922−924.

[48] M. Farooq, J.M. Fontana, E. Sazonov, A novel approach for food intake detection using electroglottography, Physiol. Meas. 35 (no. 5) (2014) 739 <http://iopscience.iop.org/0967-3334/35/5/739>. (Accessed 02.07.14)

[49] A. Kandori, T. Yamamoto, Y. Sano, M. Oonuma, T. Miyashita, M. Murata, et al., Simple Magnetic Swallowing Detection System, IEEE Sens. J. 12 (4) (2012) 805−811.

[50] M. Sun, J.D. Fernstrom, W. Jia, S.A. Hackworth, N. Yao, Y. Li, et al., A Wearable Electronic System for Objective Dietary Assessment, J. Am. Diet. Assoc. 110 (2010) 45−47.

[51] J.M. Fontana, E.S. Sazonov, A robust classification scheme for detection of food intake through non-invasive monitoring of chewing, 2012 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.(EMBC) (2012) 4891−4894.

[52] S. Passler, W.-J. Fischer, Food Intake Activity Detection Using a Wearable Microphone System, 2011 7th Int. Conf. Intell. Environ. (IE) (2011) 298−301.

[53] K. Takahashi, M.E. Groher, K. Michi, Methodology for detecting swallowing sounds, Dysphagia 9 (1) (1994) 54−62.

[54] J. Lee, C.M. Steele, T. Chau, Time and time−frequency characterization of dual-axis swallowing accelerometry signals, Physiol. Meas. 29 (9) (2008) 1105−1120.

[55] A. Das, N.P. Reddy, J. Narayanan, Hybrid fuzzy logic committee neural networks for recognition of swallow acceleration signals, Comput. Methods Programs Biomed. 64 (2) (2001) 87−99.

[56] J. Lee, S. Blain, M. Casas, D. Kenny, G. Berall, T. Chau, A radial basis classifier for the automatic detection of aspiration in children with dysphagia, J. NeuroEng. Rehabil. 3 (1) (2006) 14.

[57] S. Damouras, E. Sejdic, C.M. Steele, T. Chau, An Online Swallow Detection Algorithm Based on the Quadratic Variation of Dual-Axis Accelerometry, IEEE Trans. Signal Process. 58 (6) (2010) 3352−3359.

[58] E. Sazonov, O. Makeyev, P. Lopez-Meyer, S. Schuckers, E. Melanson, M. Neuman, Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior, IEEE Trans. Biomed. Eng. 57 (3) (2010) 626−633.

[59] O. Makeyev, P. Lopez-Meyer, S. Schuckers, W. Besio, E. Sazonov, Automatic food intake detection based on swallowing sounds, Biomed. Signal Process. Control 7 (6) (2012) 649−656.

[60] P. Lopez-Meyer, O. Makeyev, S. Schuckers, E. Melanson, M. Neuman, E. Sazonov, Detection of Food Intake from Swallowing Sequences by Supervised and Unsupervised Methods, Ann. Biomed. Eng. 38 (8) (2010) 2766–2774.

[61] F. Zhu, M. Bosch, N. Khanna, C.J. Boushey, E.J. Delp, Multilevel segmentation for food classification in dietary assessment, 2011 7th Int. Symp. Image Signal Process. Anal. (ISPA) (2011) 337–342.

[62] F. Zhu, M. Bosch, T. Schap, N. Khanna, D.S. Ebert, C.J. Boushey, et al., Segmentation Assisted Food Classification for Dietary Assessment, Proc. SPIE. 7873 (2011) 78730B.

[63] J. Shi, J. Malik, Normalized Cuts and Image Segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 22 (1997) 888–905.

[64] P. Lopez-Meyer, S. Schuckers, O. Makeyev, J.M. Fontana, E. Sazonov, Automatic identification of the number of food items in a meal using clustering techniques based on the monitoring of swallowing and chewing, Biomed. Signal Process. Control 7 (5) (2012) 474–480.

[65] U. Toepel, J.-F. Knebel, J. Hudry, J. le Coutre, M.M. Murray, The brain tracks the energetic value in food images, NeuroImage 44 (3) (2009) 967–974.

[66] W. Jia, Y. Yue, J.D. Fernstrom, Z. Zhang, Y. Yang, M. Sun, 3D localization of circular feature in 2D image and application to food volume estimation, 2012 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC) (2012) 4545–4548.

[67] M. Sun, Q. Liu, K. Schmidt, L. Yang, N. Yao, J. D. Fernstrom, M. H. Fernstrom, J. P. DeLany, and R. J. Sclabassi, Determination of food portion size by image processing in 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008. EMBS 2008, 2008, pp. 871–874. <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4649292&tag=1>. (Accessed 02-July-2014)

[68] W. Jia, Y. Yue, J.D. Fernstrom, N. Yao, R.J. Sclabassi, M.H. Fernstrom, et al., Imaged based estimation of food volume using circular referents in dietary assessment, J. Food Eng. 109 (1) (2012) 76–86.

[69] C. K. Martin, S. Kaya, and B. K. Gunturk, Quantification of food intake using food image analysis in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009. EMBC 2009, 2009, pp. 6869–6872. <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5333123>. (Accessed 02-July-2014)

[70] C.D. Lee, J. Chae, T.E. Schap, D.A. Kerr, E.J. Delp, D.S. Ebert, et al., Comparison of known food weights with image-based portion-size automated estimation and adolescents' self-reported portion size, J. Diabetes Sci. Technol. 6 (2) (2012) 428–434.

[71] H.-C. Chen, W. Jia, Y. Yue, Z. Li, Y.-N. Sun, J.D. Fernstrom, et al., Model-based measurement of food portion size for image-based dietary assessment using 3D/2D registration, Meas. Sci. Technol. 24 (10) (2013) 105701.

[72] C. Xu, Y. He, N. Khanna, C. Boushey, E. Delp, Model-based food volume estimation using 3D pose, Proc. 2013 IEEE Int. Conf. Image Process. (2013) 2534–2538September 15-18, Melbourne, Australia

[73] O. Amft, M. Kusserow, G. Troster, Bite Weight Prediction From Acoustic Recognition of Chewing, IEEE Trans. Biomed. Eng. 56 (6) (2009) 1663–1672.

[74] J.M. de Castro, Eating behavior: lessons from the real world of humans, Nutrition 16 (10) (2000) 800–813.

[75] J.M. de Castro, Seasonal rhythms of human nutrient intake and meal pattern, Physiol. Behav. 50 (1) (1991) 243–248.