A WEARABLE SENSOR SYSTEM FOR AUTOMATIC FOOD INTAKE DETECTION AND ENERGY INTAKE ESTIMATION IN HUMANS

by

ABUL BARKAT MOLLAH SAYEED UD DOULAH

EDWARD SAZONOV, COMMITTEE CHAIR
MEGAN A. MCCRORY
XIANGRONG SHEN
FEI HU
AIJUN SONG

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical & Computer Engineering
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2018

ABSTRACT

Accurate measurement and estimation of energy intake (EI) are important for the understanding of energy balance and body weight dynamics. Food is the primary source of energy in the body. Therefore, the objective monitoring of food intake patterns and eating behavior is necessary, as excessive EI leads to medical conditions such as obesity and overweight. Traditionally, self-report methods (e.g. food records, 24-hour recall) have been used for EI measurement. Most of these methods rely on a participant's own declaration in one form or another and suffer from misreporting of EI. To lessen the misreporting problems, various methods have been proposed ranging from image-assisted estimation to wearable on-body sensors, each with its own strengths and limitations. While some of the methods show great promise under certain circumstances, objective, accurate and cost-efficient methods for estimating of EI are yet to be developed. Towards automatic food intake detection, this dissertation first explores the desired time resolution of sensor-based food intake detection to characterize meal microstructure. This dissertation then investigates a wearable sensor system for automatic food intake detection, microstructure parameter estimation, passive image capture, and EI estimation. The automatic food intake detection with this system is developed by monitoring chewing activity associated with ingestion. Along with an accelerometer sensor, a novel chewing sensor is introduced to capture chewing information and detect food intake events. A wearable camera, capturing passive images in 15sec intervals, is used to acquire food and non-food images. A visual review of the food intake images is performed to identify food items, estimate portion size and validate that food was consumed. A study was performed with participants wearing the wearable sensor system (Automatic Ingestion Monitor, AIM-2) for 24h in pseudo-free-living and 24h in a free-living environment. Classification models were developed for automatic food intake detection. The estimation of chew counts was obtained using the new chewing sensor which was not adhesively attached to the body. The dissertation further presents an EI estimation method using both sensor-extracted features and image estimated portion size. Results suggest the potential of the wearable system to estimate EI in a free-living environment.

DEDICATION

This work is dedicated to my beloved parents, brother, wife, friends, and all my teachers.

ACKNOWLEDGMENTS

My time spent at the University of Alabama has been an opportunity to grow in many ways academically, personally, and spiritually. First of all, I want to thank God Almighty (Allah) for guiding my path, hearing my prayers, and providing the strength to pursue my goals. Likewise, to my parents for always encouraging me to do better in life, be my inspiration, and continue to provide your unconditional love and support.

I cannot express enough thanks to my advisor Dr. Edward Sazonov, whose direction and persistent support made this work possible. Dr. Sazonov was very cooperative, motivating, patient and greatly knowledgeable with his doors always open whenever I had doubts. I want to acknowledge the assistance and support of Dr. Muhammad Farooq, who was my friend/mentor in the early days of my graduate studies. I would also like to thank all my thesis committee members Dr. Megan McCrory, Dr. Xiangrong Shen, Dr. Aijun Song and Dr. Fei Hu for their insightful comments and constructive review in the writing of this dissertation. I have been blessed with several friends that I came to know during my studies at UA and made this journey even more stimulating. First, I would like to express my humble appreciation to those friends that helped me in times of difficulties and always found a way to lift me up: Muhammad Farooq, Habib Omar, Abu Shahab. My friends from my research group, Nagaraj Hedge, Masudul Imtiaz, Delwar Tapu, Dr. Sawal Ali and Dr. Raul Ramos for their support and productive discussion, which always directed me to better solutions to the problems at hand. I would also like to thank the community members in Tuscaloosa and Islamic Society of Tuscaloosa who were kind-hearted, welcoming and made my life at USA livable.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Motivation and Goals

An important factor in maintaining a balanced diet and healthy lifestyle is the balance between the Energy Intake (EI) and Energy Expenditure (EE). Irregularities in energy balance can lead to problems towards the development of chronic diseases such as obesity, overweight, diabetes, high blood pressure and heart diseases. Humans need energy in the body for basic metabolic processes, physical activity, physiological functions, heat production of new tissues, and maintenance of good health [1]. Food is the source of EI and the main source of nutrients necessary to maintain daily life. After ingestion, food energy is converted into different energy forms including thermic and mechanical energy. Thus, the body energy is derived from all foods and caloric beverages by oxidation. On the other hand, EE, a measure of energy burnt by the body, comprises of resting metabolic rate (the amount of energy expended while at rest), the thermic effect of food (required to digest and absorb foods and assimilate the energy derived from foods), and physical activity thermogenesis. When EI is higher than EE, the result is positive energy balance and storage of the excess energy as fats. A sustained fat accumulation can potentially lead to chronic conditions like obesity and overweight. Therefore, it is critical to accurately measure EI and EE patterns to gain insights to reveal important information about obesity and overweight-related disease.

The monitoring of EI includes detection of eating episodes, food items consumed, and eating patterns, and computing EI. To date, the most precise way of estimating EI is to use an indirect measurement through the use of doubly-labeled water (DLW). However, this method is expensive for regular use and cannot identify individual food intake events. Traditionally, EI is assessed by means of self-report methods such as food records, food frequency questionnaires and 24-hr dietary recall [2]–[4]. The food record methods rely on subject's declaration about the time, type and amount of food items consumed, as well as details about the food item such as serving size, amount of serving actually consumed, any condiments used, container type (for beverages) etc. The 24hr dietary recall method requires the subject to recall the food items they consumed in the last 24hrs, therefore highly dependent on the subject's memory. Due to the direct involvement of subject, several studies suggest that during self-report, subjects are prone to error in reporting and tend to misreport, particularly underreport their intake where the error may reach up to 50% [5]. With the advent of technology, the limitations of self-report methods led many investigators to take advantage of technological innovations to develop new methods to improve the accuracy of EI monitoring and assessment. The easy use of the internet and available smartphones has allowed for the collection of detailed food intake information with lower costs and burden, and facilitating timely approaches to EI analysis [6], [7]. Many new methods have been proposed ranging from image-assisted estimation to wearable on-body sensors. EI estimation through food imagery utilizes both manual review of images and automatic image processing techniques. In the manual review, the food images are analyzed by an expert nutritionist to identify different types of foods, their portion size, and energy content. Automatic image recognition and processing relies on computer algorithms to segment food images, recognize foods, estimate portion size/volume and compute energy value [8]. Both approaches have their own limitations. Manual review of a large number of images in a dataset may not be feasible. On the other hand, an automatic food recognition algorithm that requires a fiducial marker (known as dimensional and color references) in images may not be readily available. To offer another domain of EI estimation, various wearable sensors have been proposed to detect different stages of eating i.e. bites, chewing and swallowing of food and thus providing an understanding of the microstructure of eating [9]. The microstructure information can then be utilized to estimate EI. However, wearable sensors typically do not provide explicit information about the foods being eaten.

Our research group previously proposed a piezoelectric strain gauge based chewing sensor system "Automatic Ingestion Monitor (AIM)" for monitoring food intake and estimating EI [10]. In the study [11], EI was estimated using chewing and swallowing parameters and compared with weighed food records and self-report. In laboratory settings, participants were asked to consume their meals wearing the sensor systems. The participants' eating behavior was recorded through a video monitoring system. Features extracted from sensors and video annotation were used to build EI models. The EI was obtained by multiplying the mass intake with an energy density of the food item. The reporting error (%) of the best EI (kcal) model based on chews counts was 30.42 ± 23.08 (mean ± standard deviation).

1.2 The Big Picture

Several methods based on images and wearable sensors have demonstrated the potential for detecting food intake events, classifying food types and estimating EI. Some of the wearable sensor-based methods also extract microstructure parameters to further characterize food intake events. To improve dietary assessment, few proposed works utilize a wearable camera-based sensor system. Nevertheless, for these systems to be reliably usable in free-living conditions, there is a need for methods that offer automatic detection of food intake, recognition of food items in images, microstructure information, a way to alleviate privacy concerns, and EI estimation. Automatic detection of food intake can be achieved by monitoring chewing during food intake. To approach this problem, this dissertation proposes an algorithm for detection of chewing using a new chewing sensor. Classification models are proposed to classify segments of the sensor signals either as eating or non-eating. To accurately detect meal microstructure, the desired time resolution of sensors for food intake detection is investigated. For reducing the burden of reviewing a large number of images dataset, the number of images is significantly reduced by considering only the images that depict food intake. To estimate EI, a novel method is proposed that utilizes sensor features to estimate portion size and images to estimate both portion size and energy density of food items. Figure 1-1 shows an overview of the proposed system.

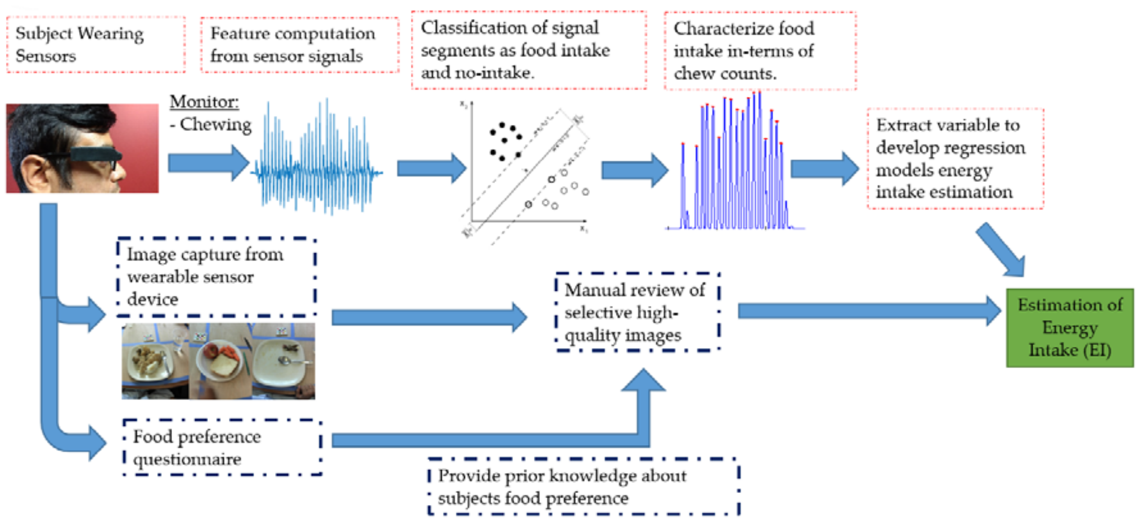

Figure 1-1. Overview of the proposed system. Wearable sensors are used to monitor the ingestive behavior of the subject. Features extracted from the sensor signals are used to train classifiers to identify between segments as food intake or non-intake. Chew count estimation is done using chewing sensor. Wearable camera capturing passive images in 15sec intervals are manually reviewed to recognize food items and portion size. Food preference questionnaire helps to identify unknown food items. Sensor-extracted features are used to

build a model that provides an estimation of ingested mass. Both images and sensor predicted portion size are used to estimate Energy Intake (EI).

Chapter 2 presents a systematic review which examines emerging technology-driven methodologies that cover the estimation of EI in humans. Chapter 3 details the desired time-resolution for sensor-based food intake detection required to detect meal microstructure. Chapter 4 describes a methodology developed to separate food images from non-food images using a clustering algorithm. Chapter 5 describes a novel wearable system and method to accurately detect food intake, estimate chew counts and provide passive images of foods being consumed in a free-living environment. Chapter 6 describes a method to estimate EI in free-living utilizing wearable sensor and passive images. Chapter 7 concludes the dissertation and gives future directions.

CHAPTER 2
A SYSTEMATIC REVIEW OF TECHNOLOGY-DRIVEN METHODOLOGIES FOR ESTIMATION OF ENERGY INTAKE

This work is submitted to the IEEE Access and is under review.

Abstract

Accurate measurement and estimation of Energy Intake (EI) are important for the understanding of energy balance and body weight dynamics. Traditional measurements of EI rely on self-report, which may be inaccurate and underestimate EI. The imperfections in traditional methodologies such as food recall, food diaries, and food frequency questionnaires stipulate development of technology-driven methods that rely on wearable sensors, food images, biomarkers, mathematical modeling and other ways to provide an objective and accurate assessment of EI. The aim of this research was to systematically review and examine peer-reviewed papers that cover the estimation of EI in humans, with the focus on emerging technology-driven methodologies. Five major electronic databases were searched for articles published from January 2005 to August 2017: Pubmed, Science Direct, IEEE Xplore, ACM library, and Google Scholar. Twenty-nine eligible studies were retrieved that met the inclusion criteria. The review identified that while the current methods of estimating EI show promise, accurate estimation of EI in free-living individuals presents many challenges and opportunities. So far, the best result for EI (kcal) estimation had an average accuracy of 94%. However, results were obtained from limited food items, in a limited number of participants, sample sizes, and in controlled laboratory conditions. Therefore, new

methods that accurately estimate EI over long time periods in free-living conditions are urgently needed.

2.1 Introduction

Energy intake (EI) is an important component of energy balance affecting the regulation of body weight. Objective assessment of EI is needed to understand body weight dynamics associated with underweight, overweight, and obesity in humans [1]. Underweight generally results from an inadequate EI compared with, or energy expenditure, energy utilization in the body. Being underweight increases health risks such as malnutrition, bone fractures, heart irregularities, inability to fight infection, and even premature death [2]. Over 815 million people in developing countries are affected by the burden of underweight [3]. According to the National Health and Nutrition Examination Survey (NHANES) 2011-2012, approximately 1.7% of US adults are underweight [4].

On the other hand, overweight and obesity are primarily associated with excessive EI and low energy expenditure. These health conditions present a growing global epidemic with substantial individual and public health consequences. The World Health Organization in 2016, estimated that more than 1.9 billion adults (nearly 39% of the world's population) were overweight and over 650 million of those adults were obese [5]. Obesity and overweight are major contributors to chronic diseases such as type 2 diabetes, asthma, cardiovascular diseases, cancers, and musculoskeletal disorders, and caused 3.4 million deaths in 2016 [6]. Obesity imposes a considerable economic impact on individuals. Obese individuals spend 30% more per capita on medical costs than healthy individuals [7]. Apart from direct costs, obesity also results in indirect costs due to a work productivity losses, disability, and premature mortality [8]. In the United States, such productivity losses cost an estimated

30.15 billion USD at the national level. Due to obesity-related absence from work alone, the nationwide productivity costs range between $3.38 billion and $6.38 billion [9]. Several approaches have been tried to combat the obesity epidemic; however, a significant linear trend still exists in obesity prevalence [10]. Along with eating disorders, overweight and obesity are health conditions highly resistant to active medical treatment [11].

Based on the energy balance, the association between EI and energy expenditure are important for understanding both underweight and obesity [12]. With the advent of micro-electronic accelerometers, there has been rapid progress in the methods for assessing free-living energy expenditure [13]. However, the tools for monitoring of EI are far less developed. Accurate measurement of EI and characterization of related eating behaviors are one of the biggest challenges in nutrition research.

A range of EI assessment methods, including food frequency questionnaires (FFQs), 24 h dietary recalls, weighed food records (WFRs), and short assessment screens, have historically been used for EI measurement, each with its own strengths and limitations [14], [15]. Most of these methods rely on participants own recollection and declaration (self-report) in one form or another and suffer from underreporting, misreporting, and not-reporting of EI [16].

From the perspective of accurate evaluation of EI, energy expenditure measured by lithium and b-carotene have been investigated. These suffer from the limitation that one must consume enough foods containing these markers, either in pre-dosed foods or exogenously from the foods already eaten, relatively, and they require biological samples at specific times for accuracy which limits their use in truly free-living situations. Doubly Labeled Water (DLW) method has been identified as the gold standard for providing an estimate of true energy expenditure over 10 days. If a person is dosed with DLW and remains weight stable (i.e. is in energy balance), it is assumed that EI is equivalent to EE. This method is accurate assuming that the person is in energy balance but EE and EI fluctuate from day to day so this method gives a broad, undefined view of EI only and cannot be used to assess eating behavior of specific nutrient intake as it can only estimate overall energy. Eating behavior and nutrient density are important factors that affect chronic health and need to be considered when assessing EI. In addition, this method cannot be used in situations where energy balance does not hold true such as growing children or during active weight loss or weight gain. In addition, the dosing and timing of the samples required for this method is difficult to employ in free-living situations. Similarly, all biomarkers require biological samples at specific times, expensive measurement instruments, and procedural requirements. Thus, there is still a need for the development of lower-cost techniques for assessment of energy expenditure as a biomarker which can be employed to measure EI.

The limitations of existing methods led to attempts to utilize technological innovations to improve the accuracy of EI assessment. In particular, the use of the internet and smartphones has resulted in the improvement of traditional methods, allowing the collection of detailed nutritional information with lower costs and burden, and facilitating timely approaches to EI estimation [17], [18]. To lessen the problems associated with self-reported EI, various methods have been proposed ranging from image-assisted reporting to wearable on-body sensors. The development of technology-driven, quantitative methods of EI estimation has received significant attention as an active area of discovery. In this article, we provide a systematic review of methodologies used to estimate EI in humans, with the focus on recent technology-based developments.

2.2 A Brief Overview of Concepts Related to Energy Intake

2.2.1 Energy Balance

Energy balance is the difference between EI and energy expenditure. The first law of thermodynamics is the law of conservation of mass and energy, which states that energy can be transferred from one form to another, but it cannot be created or destroyed. In the context of human physiology, this is expressed as:

$$ES = EI - EE \tag{2.1}$$

where ES = rate of energy stored in kilocalories (or kJ) per day, EI = rate of energy intake in kilocalories per day, and EE = rate of energy expended in kilocalories per day [12]. It is to be noted that the variables are time-dependent and therefore energy balance depends on the time domain during which they are measured [19]. At any given time, the human body is considered an open system in which energy can be added by intake of foods or subtracted by expending energy through basal metabolism, exercise or physical activity related thermogenesis, non-exercise or physical activity thermogenesis, and dietary-induced thermogenesis [20]. The energy balance equation describes what occurs to ES when EI and EE are in balance or imbalance over a given period of time (Figure 1) [12]. In balance, the EI equals energy expenditure. An imbalance of energy implies a change in the energy content of the body [1] which consequently changes energy expenditure due to a change in body weight.

2.2.2 Energy Intake, Calorie and Energy Density

Humans need energy in the body for basic metabolic processes, physical activity, physiological functions, generation of new tissues, and maintenance of good health. EI represents the energy content (expressed as kilocalories/kcals or kJ) derived from all foods and beverages by oxidation. The energy is released after ingestion of foods and converted into different forms including thermal and mechanical energy.

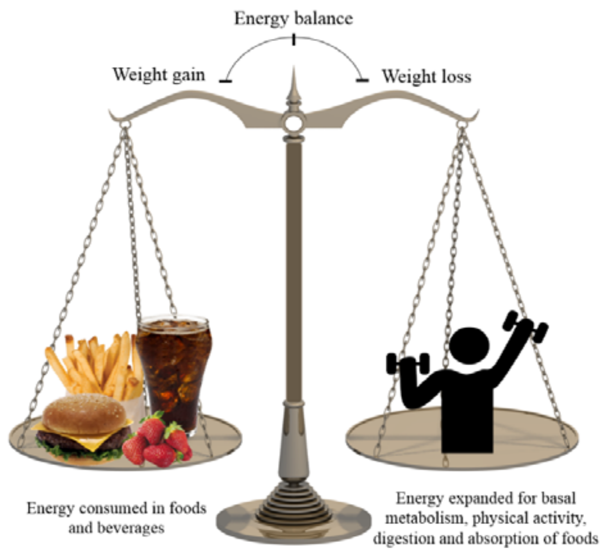

Figure 2-1. Representation of the energy balance paradigm.

A kilocalorie (kcal, or Calorie) is defined as the amount of heat needed to raise the temperature of 1kg of water 1º Celsius and also represents a metric unit of energy measurement [21].

Energy density is defined as the amount of energy per unit weight of food and is usually expressed in kcal/g or kJ/g [22]. The main sources of energy are provided by the three macronutrients: carbohydrates, protein, and fat. The average values of macronutrients that are 4 kcals per gram for carbohydrate and protein, and 9 kcals per gram for fat [23]. So, the macronutrient composition of foods is primary determinants of energy density [24]. Another primary determinant of energy density is the water content of food, as it adds weight but no energy [24], [25]. The fiber content of foods also acts as a determinant of energy density. Non-digestible or low-digestible compounds such as dietary fiber reduce the energy density of foods because they contain fewer than 4 kcals/g [26]. For example, the energy derived from the fiber is estimated as 2 kcal/g [27].

2.3 Methods

This was a systemic review study and the study protocol was developed in accordance with the Preferred Reporting Items for Systematic Reviews (PRISMA) statement [28] and agreed to by all authors.

2.3.1 Search Strategy and Data Sources

Given the research topic of finding available technology-driven methods for EI estimation, the most frequently used keywords related to EI were first identified. These keywords were then searched in five databases (Pubmed, Science Direct, IEEE Xplore, ACM library, and Google Scholar) from January 2005 to August 2017. The final list of keywords that were used to conduct the full search to get the primary selected papers is in Table I. A total of 482 articles was identified for complete screening. Based on the title and abstract, the articles were checked and screened. The primary selected articles were further screened based on the inclusion and exclusion criteria discussed in the following subsections. Cross-references were screened for additional eligible articles. The final selection of articles was decided after reading the full articles.

2.3.2 Inclusion and Exclusion Criteria

All primary selected articles were checked for eligibility and acceptability by adhering to the following inclusion criteria: 1) study participants were 18 years or older; 2) study participants had a Body Mass Index (BMI) over 18 kg/m^2; 3) studies aimed at estimating/predicting EI in humans (i.e. in kcal or kJ per day/meal/food item) ; 4) studies aimed at estimating EI-related measures such as amount consumed, energy density, food recognition; and 5) studies which proposing approaches that utilize technology for EI estimation. The aspects of technology for EI estimation encompass methods that rely on image analysis and/or analysis of physiological signals obtained from wearable sensors. Full texts of the articles that met the inclusion criteria were obtained for review.

In the field of dietary assessment, most research considers two major aspects: 1) Food intake detection that deals with identification of all eating episodes, measurement of duration and microstructure parameters of each eating episode etc., and 2) Characterization of food intake which deals with estimation of food items, mass, volume, energy, and nutrients consumed during each eating episodes. In this systematic review, the studies which purely dealt with the food intake detection were excluded. A good review of food intake detection and existing methods can be found in [29]. Since the focus of the present study was to review methodologies related to EI estimation, the studies that were solely concerned with the validation of an existing method of EI estimation were also excluded. The excluded validation studies were primarily focused on comparing the performance of self-reported EI against 1) estimated EI required during weight maintenance; 2) estimated EI measured from EE and, 3) presumed energy requirements estimated from basal metabolic rate (BMR). All duplicate papers were also excluded from the analysis.

2.3.3 Data Extraction

The full text of each selected article was thoroughly reviewed to extract relevant data using a standardized data extraction procedure. The data extraction included information on the following domains: study design, study characteristics, method/algorithm used, data analyses, and significant findings.

2.4 Results

2.4.1 Study Selection

Table 1 summarizes the article selection for the review. After the primary selection, Figure 2 illustrates the PRISMA diagram and finally selected articles for the review. We identified 6029 articles include 1402 from Pubmed, 521 from Google Scholar, 1556 from IEEE Xplore, 1198 from Science Direct, 1352 from ACM library. Of the 6029 papers, 1,943 were duplicates and removed from the searches. After title and abstract review of primary selected articles, 448 articles were excluded leaving 29 eligible studies. The last search for the study was conducted on 12 August 2017. The study analyzed and tabulated the important data from eligible studies: lead author's name, publication year, gender and sample size, mean age of subjects, study design, EI estimation in each method.

Table 2-1. Keywords Used to Obtain Candidate Articles

Keywords	Databases					Total publication identified	Full-text article, Peer-reviewed Journal articles, Conference proceedings	Primary selection
	Pubmed	Google Scholar	IEEE Xplore	Science Direct	ACM Library			
Energy intake estimation	656	119	1044	755	863	3437	410	227
Nutrient intake estimation	164	99	184	265	78	790	85	62
Food amount consumed estimation	144	88	95	45	98	470	65	48
Energy intake model	233	201	185	65	217	901	105	95
Mathematical modelling Energy intake	205	14	48	68	96	431	288	50
Total	**1402**	**521**	**1556**	**1198**	**1352**	**6029**	**953**	**482**

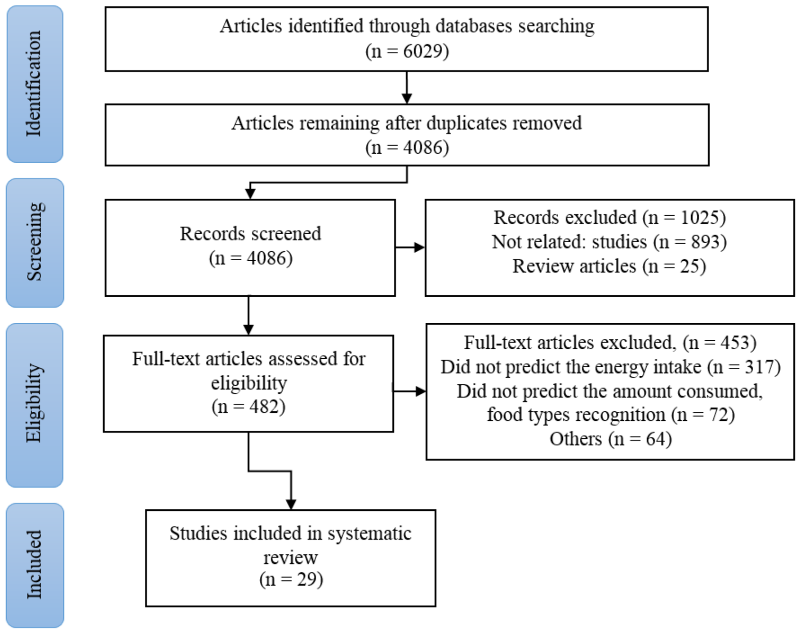

Figure 2-2. Flowchart of the systematic review outlining the study selection process.

2.4.2 Methodologies to Estimate EI and Associated Parameters

Based on the review of the selected articles, the methodologies for estimating EI and related metrics (i.e. amount consumed, food type recognition, portion size estimation) were divided into five main groups based on methods: traditional self-report, DLW, image-assisted, mathematical modeling of the intake/balance equations, and wearable sensor.

2.4.2.1 Food Diary, Recall, and Web-Based Methods

Traditionally, the estimation of EI in humans, especially in free-living conditions, is done by means of self-reporting. Various self-reporting methods such as food diaries, 24-hour dietary recalls, and food frequency questionnaires have been developed and widely used to estimate EI [14], [15], [30]. With food diaries, participants are asked to report all food items consumed during each eating occasion with the amount consumed (portion size) and other details such as if the food was regular or low-fat, how it was cooked, and additions such as condiments. To estimate EI, the recorded information is reviewed by an expert nutritionist and entered into a software program containing a food dataset that contains the energy density of thousands of individual foods and performs the nutritional analysis. However, the estimation of EI from food diaries may be inaccurate due to several factors one being lack of sufficient detail for the nutritionist to perform an accurate estimate. For example, in one study, [31] about 47% of food diaries didn't contain sufficient information for accurate EI estimation. These errors can result in underreporting of EI when compared with expected energy requirement. Unlike food diaries, 24-hour dietary recalls require participants to recall all foods and beverages consumed in the past 24 hours and has the advantage that this is often obtained via phone directly with the help of a nutritionist so more complete data can be gleaned. 24-hour recall also results in underreporting error. Despite lowering the burden of recording information after each occasion, individuals may have difficulty remembering, or not want to reveal, all items consumed in the amounts consumed. With food frequency questionnaires (FFQ), participants are provided a finite list of foods or groups of foods and asked to report how frequent they consumed each food over a defined period of time such as the past 1 year or 1 month. Semi-quantitative FFQs also include an estimate of portion consumed, either as a standard portion or participants select from given portion sizes. EI can be estimated from semi-quantitative FFQs by incorporation of frequency portion consumed, and known energy content of a given portion. However, this assessment involves difficult, abstract concepts and math. If the time frame of the questionnaire, for example, 6 months, crosses seasons, say summer to fall, the participant must compute how often over both seasons they ate seasonal items such fruit. Similar to food diaries and 24-hour recall, FFQ also includes participants own declaration and therefore produce a large error in the estimation of EI. FFQs typically result in greater underestimates of EI than 24h recalls or food diaries due to the finite list of foods provided.

During behavioral interventions aimed at weight loss/gain, the self-report based methods can provide important information about EI. With the advent of technology, web-based 24-hour recall software programs have been developed. For example, the Automated Self-Administered 24- hour Dietary Assessment Tool (ASA24) developed by the National Cancer Institute has freely available and increased access to 24h recall software over other software with 24h recall capability but that has a cost associated with it. The ASA24 allows participants to enter dietary intake during the past 24 hours and portion size estimation is assisted by images of typical portion sizes [32]. Alternately, the researcher may use the ASA24 to conduct an interview-assisted 24h recall. The ASA24 has been shown to provide comparable EIs with its predecessor, the USDA multiple pass method whose software was previously only accessible to the USDA [33]. Several smartphone applications have been developed in recent years. These are not considered research-quality and are meant for consumers to track their own dietary intake [34], [35].

Despite these advantages, due to a variety of factors, self-report methods grossly underestimate EI [36]–[41]. These factors include lack of knowledge about portion size, and unintentional (due to forgetting) and intentional misreporting. Weighed food intake records may also result in undereating (not eating as much as usual) and appear as underreporting. The authors in [42] demonstrated that the degree of misreporting, including undereating and under recording, may range from 20% to 50% for individuals in the normal to overweight range.

The authors of [36] comprehensively reviewed the effect of underreporting and the biases in estimating EI from self-reported methods. The review also explored the interpretation of data produced by the underreporting of participants.

2.4.2.2 Doubly Labeled Water (DLW)

The DLW technique is considered the gold standard for measuring EE under free-living conditions over a specified time period [36]. If a person is in energy balance, the EI can be regarded equal to measured EE by DLW. With this technique, a dose of water enriched with the stable isotopes deuterium (^{2}H) and oxygen 18 (^{18}O) is consumed. Over time, the ^{2}H and ^{18}O mix with water in the body (including the more abundant ^{1}H and ^{16}O in the body, but also basal amounts of ^{2}H and ^{18}O which are present in our everyday water supply), and several biological samples of fluid are collected over the next several days. The fluid sampled is, typically urine, as it is the least invasive, but the fluid sampled could also be blood, saliva, or breast milk. Fluids are collected over a defined period of time. Typically, this technique involves collecting 7 urine samples over two weeks although other dosing, sampling, and timeframes have been validated ranging from 7 to 21 days [43]. Urine samples are typically analyzed by mass spectrometry to determine the isotopic elimination from the body. Deuterium is lost in water only, whereas oxygen 18 is lost in both water and carbon dioxide. The rates of disappearance measure the body's water and water-plus-carbon dioxide turnover rates, from which carbon dioxide production can be calculated by difference. Total Energy Expenditure (EE) is calculated from carbon dioxide production by applying classical indirect calorimetric equations [44].

The estimation of EI from the measured EE involves the usage of the fundamental physiological equation of energy balance. For individuals in energy balance, when the body weight is constant the EI equals the EE. DLW has been shown to estimate EI (kcal) within 2% to 8% in inpatient clinical studies and within 8% to 15% in field studies [45]. Because of its high degree of accuracy, the DLW method is often used as an EI benchmark to measure total EI in free-living [36], [45], [46]. However, because of the high cost associated with obtaining the isotopes, the DLW method is only available to the most well-funded and equipped laboratories and clinics. Many researchers fall back on the use of various methods of self-monitoring to obtain estimates of free-living EI such as food diaries. One of the major benefits of the DLW method is that despite its cost and effort, it can be used to validate the performance of novel methods that estimate EI[36].

2.4.2.3 Mathematical Modeling and Intake-Balance Method

The intake-balance method of EI estimation utilizes the energy conservation principle: EI over a given period equals to the sum of the change in EE and the corresponding changes in body energy stores (ES) [19]. Therefore, EI can be estimated if both EE and ES are available. The measurement of the change in EE can be obtained using the DLW method over the period of the experiment. The change in body ES can be estimated from the changes in body composition. Several methods can be used to measure body composition such as multiple dual-energy X-ray absorptiometry (DXA) scans, deuterium dilution, underwater weighing, or air displacement plethysmography [47]. Change in body weight may also be used, with an assumption about the composition of weight change. Theoretically, the intake-balance method could provide an accurate estimation of EI over extended time periods if the measurements of EE and ES are accurate. However, this indirect method typically involves DLW and DXA methods which are highly expensive (DLW) and requires specialized trained individuals (DLW and DXA).

In the recent past, mathematical models of human EI, metabolism, and body dynamics have been proposed as a relatively inexpensive possible alternative to the intake-balance method. Mathematical models for cumulative intake curve (CIC) were investigated. In this systematic review, two studies related to EI modeling and one related to the CIC were identified. Table 2-2 illustrates the overview of included studies using mathematical modeling approaches. Table 2-3 illustrated the summary of methods for EI and CIC estimation.

Table 2-2. Studies Using Mathematical Modelling Approaches

Study	**Sample size**	**Mean age**	**BMI / BW range**	**Notes about subjects**	**Study design**	**Dietary EI estimation**	**Primary outcome measures**
Hall (2011) [19]	3	NA	NA	Obese	Controlled cohort study	ΔEI (kcal)	Estimating ΔEI in the free-living environment during the weight-management program
Thomas (2011) [48]	36	NA	NA	Obese	Controlled cohort study	EI (kcal)	Predicts individual weight changes
Thomas (2017) [49]	22	NA	NA	Over-weight to obese	Randomized controlled trial	CIC (g)	Estimate the cumulative intake curve

Table 2-3. Mathematical Modelling Methods for EI and CIC Estimation

Study	**Equations**	**Dependent variables**
Hall (2011) **[19]**	$\Delta EI = \varepsilon(BW_i - BW_0) + \rho \frac{dBW}{dt}$	BW: Body weight (kg) ρ : Effective energy density (kcal/kg) ε : parameter defines how EE depends on BW (kcal/kg/d)
Thomas (2011) **[48]**	$EI = c_l \frac{dFFM}{dt} + c_f \frac{dFM}{dt} + EE$ $EE = RMR + PA + SPA + DIT$	FFM: Fat-free mass (kg) FM: Fat mass (kg) c_l: energy density of FFM (kcals/kg) c_f: energy density of FM (kcals/kg) RMR: Resting metabolic rate (kcals/day) PA: Physical activity (kcals/day) SPA: Spontaneous physical activity (kcals/day) DIT: Dietary induced thermogenesis (kcals/day)

***Thomas (2017)* [49]**	$I(t) = \dfrac{E_{max}\theta(e^{\frac{t(E_{max}r+\theta)}{E_{max}}} - 1)}{\theta(e^{\frac{t(E_{max}r+\theta)}{E_{max}}} + E_{max}r)}$	$I(t)$: Cumulative food intake (g) θ: The initial rate of eating (g/min) E_{max}: Maximal food intake (g) r: Eating duration (min)

In [48], the authors presented a computational model using a differential equation of energy balance for determining individual EI during weight loss. The differential equation model for EI solved the value of EI through the knowledge of ES and EE. The EI model also required the use of BW, age, height, and sex. The ES was modeled by summing the changes in ES from fat-free mass (FFM) and fat mass (FM). The EE was modeled as the sum of resting metabolic rate (RMR), volitional physical activity (PA), diet induced thermogenesis (DIT) and spontaneous physical activity (SPA). Finally, the model was tested using data from a 24-week calorie restriction study. The validity of the EI model was then tested against 1) EI estimated from provided foods and 2) EI quantified from DLW/DXA.

The authors of the study [19] introduced a method to estimate changes in EI (ΔEI) using longitudinal body weight measurements and its 95% confidence interval in individual participants. To develop the ΔEI model, the study solved the energy balance equation for free-living ΔEI as a function of body weight (BW) and the rate of its change. The study proposed a low-cost and readily available body weight scale to assess the changes in body composition (necessitating an assumption about the composition of weight change), avoiding expensive DXA scans. The performance of ΔEI method was evaluated both in a controlled inpatient feeding study and simulated free-living virtual study. In the controlled study, EI (around 800 kcal/day) was held constant and in the free-living virtual study, significant variations in day-to-day EI were considered. The authors proposed that daily BW measurement over at least 28 days is sufficient to estimate ΔEI in free-living adults. Also by summing the variance of ΔEI and variance of initial energy-balanced EI, the 95% CI of the total EI can be estimated.

The cumulative intake curve (CIC) modeling provides a quantitative description of food intake. In the early 80s, the authors of [50] introduced a quadratic equation $y = ax^2 + bx + c$ for modeling CIC, where y was food intake, a was the rate of deceleration, b was the initial speed of eating and c = initial amount of food. Recently, the study [49] proposed a CIC model based on a differential equation. To define an individual's CIC, the proposed model relied on three key parameters: initial rate of eating (θ), eating duration (r) and maximal food intake or highest amount of food that an individual can eat (E_{max}). The model exhibited better CIC than the quadratic model in 94% of the subjects and 43% improved quality of fit.

The summary of selected mathematical model equations of EI estimation and CIC is provided in Table 2-3. While the mathematically derived estimation models are yet to be tested as accurate tools for use in real-time settings and clinical conditions, they hold promise also alternate methods to estimate current EI. Again, these methods cannot measure or estimate diet quality or eating behavior so their use in clinical interventions or research settings is very limited. To the best of our knowledge, no studies have used these methods to objectively estimate EI to date.

2.4.2.4 Image-Based Methods

In the past decades, due to advancement in technology, the availability of image capturing devices has increased significantly. These personal devices with imaging capability are available in the form of mobile phones, smartphones, tablets, and wearable devices. By taking advantage of easy capture, images of food and beverage items consumed can be used to estimate EI. Typically, image-based methods acquire images of food items before, during, and after a consumption using either a mobile device or wearable camera. When an individual has to perform actions to activate image capture via mobile phone, or digital camera, the image capturing method is called ‘active’. Alternately, when the image is automatically acquired using wearable cameras that can capture images continuously or at pre-determined intervals, the image capturing method is called ‘passive’.

To estimate EI, the analysis of acquired images can be performed in two ways: manual annotation of the images and automatic image recognition and processing. During manual annotation, food images are analyzed by an expert nutritionist to identify individual foods, their portion size, and nutritional content. One potential advantage of this approach compared with traditional self-report methods is a reduction in participant burden. The studies of [51]–[53] validated manual annotation for estimating EI. In these studies, participants captured photos before and after meals and provided short descriptions of ingredients used if mixed item dishes such as stews were consumed. Then the images were retrieved and analyzed by the nutritionist using nutritional analysis software interfaced with a nutrient database. A review of manual review/visual inspection methods is presented in [54].

Automated image recognition and processing relies on computer algorithms to segment food images, recognize foods, estimate portion size/volume and compute a nutritional value. Some of the reported methods requested the user to mark foods in the image before processing, which can be categorized as a semi-automatic approach. After recognition, based on the type of foods and portion size/volume from imagery, the amount of EI is estimated using nutritional analysis software interfaced with a nutrient database. For all of these approaches, to assist the nutritionist or computer, fiducial markers (known dimensional and color references) are typically placed in the food image.

In this review, 17 papers were identified that adopted image-assisted approaches (i.e. manual review, semi-automatic and automatic image analysis). Table IV illustrates the overview of included studies. Table 2-5 summarizes the methodologies, key details and reported the accuracy of the reviewed methods.

The study of [55] presented an approach based on manual review named "Remote Food Photography Method". Before the study, participants were provided with necessary training about how to use a smartphone to capture images of their foods before consumption and any leftovers not consumed. The images were collected in free-living conditions and sent over a cellular network for review. Expert nutritionists analyzed food images to obtain EI. Around the same time, the study [56] presented a method for recognizing foods from video recordings of eating. The videos were captured and recorded by a web camera. A trained model was developed on restaurant food images and the food recognition on video frames was then carried out by image matching algorithm. After the food recognition, the method then estimated EI.

The authors of [17] described a novel mobile food record application that provided an account of daily food and nutrient intake. The method consisted of the segmentation of food items in the image, food recognition, automatic portion estimation, and extraction of nutrient information that was consumed. Two images were obtained during the start and end of each eating occasion, with a fiducial marker placed next to the food. After segmenting and recognizing food items, portion size was estimated through 3-D reconstruction. The estimation of EI was then obtained from the authors' customized food database. Instead of direct food recognition, the authors of [57] proposed a system that recognizes the dish contents from the image and computes energy and nutrients using a food-log system. The authors developed a dish segmentation algorithm using a Hough transformation and extracted rotation invariant features from images.

In [58] the authors proposed a camera phone based automatic food intake monitoring system (called DietCam). They adopted a Bayesian probabilistic approach to classify food items. Apart from regular features from images, the authors also utilized statistical features considering the noise produced by camera sensors. The authors in [59] proposed a system where the estimation of the volume of food items was done by stereoscopic imaging. A pair of images was captured from the left and right side of the food item placed on the plate. Next, feature matching between stereo images was performed. The authors estimated the volume without manual fitting of 3D models to the food items. Another study [60] proposed a method to estimate EI in a food's image by computing the volume of food portions using image segmentation followed by SVM classification. To measure the food volume, the authors took two pictures of the food items: one from the top and one from the side. A thumb beside the dish served as a reference object. To compute the volume of the food portion, the authors proposed a grid superimposed on the top view of the image. The number of pixels in the grid was then calculated to compute the food portion area. The side of view of the image was used to calculate the depth of the food. The volume was computed by multiplying the area and depth. Once the food types and food volume were obtained, nutrient composition tables were used to estimate the energy content of food items.

Table 2-4. Overview of Included Studies Using Image Assisted Approaches

Study	System design	Data Acquisition	Study setting	# of persons	# of days	Food types included in study	Total Number of image included	Food recognition	Volume/ Portion size estimation	Mass estimation
Martin (2009) [55]	Smartphone / PAD	Pre and post image from food plate	Both free-living & Lab	50	3	Full meal	NA	NA	NA	NA
Wu (2009) [56]	Web cam	Video	Both free-living & Lab	NA*	NA	Solid, liquid	101	Yes	No	No
Zhu (2010) [17]	Smartphone -Server	Pre and post image of meal	Lab	NA	NA	Solid, liquid	3000	Yes	Yes	Yes
Wazumi (2011) [57]	Smartphone	Single pre-meal image	NA	NA	NA	Full meal	350	Yes	No	No
Kong (2012) [58]	Smartphone / video	3 images before the meal, 3 images after the meal	NA	NA	NA	Solid	NA	Yes	Yes	No
Rahman (2012) [59]	Smartphone	Pairs of stereo images	NA	NA	NA	Solid fruit	6	No	Yes	No
Pouladzadeh (2014) [60]	Smartphone	1 image before and after eating	NA	NA	NA	Solid, mixed and non-mixed	3000	Yes	Yes	Yes
Myers (2015) [61]	Smartphone	Single pre-meal image	NA	NA	NA	Solid	1200 00	yes	yes	no
McAllister (2015) [62]	Smartphone	Single pre-meal image	NA	NA	NA	Solid	10	No	Yes	No
Zhang (2015) [63]	Smartphone	Single pre-meal image	Both lab & free living	NA	NA	Solid, mixed	2000	Yes	Yes	No
Fang (2015) [64]	Smartphone	Single pre-meal image	NA	NA	NA	Meal	45	No	Yes	No
Sun (2015) [65]	Wearable camera	One image in every two seconds.	NA	NA	NA	Meal	NA	Yes	Yes	No
Pouladzadeh (2016) [66]	Smartphone	1 image before and after eating	NA	NA	NA	Solid	1000 0	Yes	Yes	No

Liao (2016) [67]	Depth camera	3 images of food tray: empty tray, before eating and after eating	NA	NA	NA	Solid	NA	No	Yes	Yes
Hippocrate (2016) [68]	Smartphone / cutlery	Single image from the top with the cutlery in the picture	Lab	15	NA	Meal	119	No	Yes	yes
McClung (2017) [69]	Video camera	Pre and post image from food plate	Free-living	131	5	Solid, liquid, mixed, non-mixed	3276	No	No	No
Hassannejad (2017) [70]	Smartphone video	Six frames from video	NA	NA	NA	Meal	NA	yes	yes	no

*Note: NA: Not available;

*mixed food: multiple food items mixed together; non-mixed: food items separated on the plate

Table 2-5. Systematic Review of Image Assisted Method Towards EI Estimation

Study	Data analysis method	No of Food items	Image analysis**	Compared against***	Performance /accuracy	Real-time analysis, platform
Martin (2009) [55]	Captures images of food selection, leftovers and a reference card, nutritionist analyses images to estimate EI	NA	Manual	WFR	Free-living EI: −152 ± 694 kcal/day	Phone app
Wu (2009) [56]	Recognizes foods in videos of eating extracting SIFT features and utilizing pre-trained food model, predicts energy contents from standard fast food database	101	Automatic	DB	Not reported	NA
Zhu (2010) [17]	Segments images using hierarchical segmentation technique, classifies regions into food categories using SVM classifier, estimates portion size through volume estimation of the food from which the energy contents determined	19	Automatic	NA	Food recognition : 95.8%	Phone app
Wazumi (2011) [57]	Segments images by Hough transformation, extracts rotation invariant SPIN features, records food info in a food log system	10	Manual	NA	Food recognition : 78%	NA
Kong (2012) [58]	SIFT feature based food classification, then estimation of energy content of food items through 3D model reconstruction based volume estimation	NA	Automatic	NA	Food recognition : 92%	Phone app
Rahman (2012) [59]	Captures stereo pairs of images, create a 3D model of food items to estimate volume of the food	6	Automatic	VWD+	Volume (ml) error: 7.7%	NA

Pouladzadeh (2014) [60]	Segments images, estimates food portion from SVM classifier, measures volume and calculates nutritional facts from tables.	15	Semi-automatic	DB	Food recognition : 92.21% (single food), 85% (non-mixed foods), 35-65% (mixed foods) Energy content (kcal) accuracy: 86%	NA
Myers (2015) [61]	Detects meal from a single image, predicts foods, volume, and size of foods using CNN, estimates energy content from USDA dataset	2517	Automatic	NA	NA	Real-time, phone app
McAllister (2015) [62]	Segments images, estimates area covered by foods from the regression model, computes energy content from the selected region	1	Semi-automatic	DB	Energy content (kcal) of food item portions: 89.12%	NA
Zhang (2015) [63]	Segments images using hierarchical segmentation technique, classifies regions into food categories using SVM classifier, estimates portion size of the food to determine caloric contents	15	Automatic	NA	Food recognition: 85%	NA
Fang (2015) [64]	Estimates portion size automatically using geometric contextual information from the scene.	19	Automatic	WFR	Energy (kcal) error: 6%	NA
Sun (2015) [65]	Detects regular shaped utensils, segments food items based on color, texture measures, estimates volume of each food item using a food-specific shape model, computes energy contents and nutrients using the FNDDS database	10	Automatic	VSD┼	Volume (cm^3) estimation error : 30%	NA
Pouladzadeh (2016) [66]	Segments images using graph-cut segmentation technique, classifies and recognizes food items using deep learning neural networks, measures portion size and calculates nutritional facts from tables.	30	Semi-automatic	DB	Food recognition : 99% (single food) Energy content (kcal) estimation : NA	Real time, Phone app
Liao (2016) [67]	Filters noisy depth images, estimates volume of the foods, estimates mass of food intake using specific gravity function	3	Automatic	WFR	Food intake mass (g) error : 7.5%	NA
Hippocrate (2016) [68]	Estimates the diameter and the height of the food container and derives the food volume, given the food type, estimates the mass of the food in the image	15	Automatic	WFR	Mass (g) estimation error : 6.78%	NA
McClung (2017) [69]	Captures images using camera, estimates energy content by visual estimations from two estimator	NA	Manual	NA	NA	NA
Hassannejad (2017) [70]	Segments images by semi segmentation method interactive and user-dependent create a 3D model of food items to estimate volume of the food	10	Semi-automatic	NA	Volume (mm^3) estimation accuracy: 92%	NA

** Note: Manual: Analyzed by trained individual; Semi-Automatic: Foods in the images are marked by the user before processing; Automatic: Fully automated analysis. NA, not available.
*** WFR: weighed food record; DB: Nutrition database/ Nutritional Fact labels; NA, not applicable.
┼ VWD: measure volume using water displacement; VSD: measure volume using seed displacement

The study of [61] proposed a system called 'Im2Calories' that recognizes the contents of a meal and predicts energy in restaurant meals. The system was validated using food images from restaurants where food menus were available. The system consisted of multiple stages: 1) meal detection using GoogLeNet convolutional neural network (CNN) model, 2) recognizing the restaurant at which the meal was obtained by Google's Places API 3) food item recognition using GoogLeNet (CNN) model, 4) segmentation of the food image using "DeepLab" system, 5) volume estimation using the same CNN architecture for depth and voxel representation of the food analysis 6) EI estimation from U. S. Department of Agriculture (the USDA) database.

The study of [62] proposed a semi-automatic technique to predict energy content using regression modeling. A reference block was used in the food image to estimate the food area. The participants were asked to indicate the food area portion by drawing a polygon in the image. Based on the food portion area the proposed application trained a linear regression model to predict energy content. In [63], the author proposed a system (named Snap-n-eat) that can recognize food items on the plate and automatically estimate energy content. To identify and classify food items, the system utilized features from the segmented image and developed model using a linear SVM classifier. The system estimated the portion size of the food by counting pixels of the food portion and used it to estimate the energy content of foods from a custom database. Work reported in [64] demonstrated an approach to estimate food portion size from a single-view image. The method utilized geometric contextual information to automatically estimate the volume from the image. The study of [65] proposed the eButton which takes pictures of the food on the table automatically (one picture in every two seconds). Out of all selected papers, only this study proposed passive capture of images using the wearable camera. From the image, the volume of food items was estimated in three steps: 1) detection of utensils (e. g., circular plates or bowls); 2) segmentation of food items based on color, texture and a complexity measure; and 3) estimation of volume from food-specific shape models. The food name and volume information was then used to estimate energy and nutrient content from the Food and Nutrient Database for Dietary Studies (FNDDS, a public domain database developed by the USDA).

The same research group as in [60] proposed a combination of graph cut segmentation and deep learning neural networks to classify and recognize food items in [66]. The authors further introduced a distance-based food portion estimation in which the distance between the food image and the smartphone device was calculated using the accelerometer and magnetic field sensors. To measure the distance, the authors used the reported height of the user. Next, the energy content of food items was estimated incorporating the food portion information.

In recent years, the volume estimation technique has further improved with the introduction of depth cameras. In [67] volume estimation of food intake was done by utilizing a short-range depth camera. The depth image was acquired at the distance of 30 cm. The depth image typically introduces noise. To solve the noise problem, the author proposed bandpass filtering. The volume was estimated from the width and height of pixels relating to camera distance. The estimated volume was then converted to mass using the specific gravity function of each food item. The specific gravity functions of different food items were approximated using linear, quadratic, or cubic functions. Next, the estimated mass information was used to calculate the energy content of food items.

In [68] the authors proposed a method that measures the mass of foods using eating tools (cutlery) such as a spoon, a fork, or chopsticks. Using the Hough transformation to isolate features of shapes and canny edge detection, the system estimated the diameter and the height of the food container. The user was required to manually enter the food type to get the mass estimation.

The authors in [69] proposed the use of digital food photography (DFP) in the quantification of EI in a field setting. The study installed three DFP stations to capture pre- and post-tray images. Later, multiple nutritionists estimated EI visually and compared the estimation with the weighed record and standardized servings.

Most recently, the authors in [70], presented a novel approach to estimate the food volume using image-based modeling. Six frames from a short video of meal consumption were selected for analysis. Next, the user manually marked initial segments of food items in one of the frames. Later, the final segmentation was done using a graph-cut algorithm. The food model was developed using a point-cloud based image modeling algorithm. A fiducial marker was used as size/ground reference alongside a 3D food model to estimate the volume of food items.

Despite the improvement of computer vision systems and sophisticated image analysis algorithms, an accurate and practical method has yet to be developed. Due to the diversity of food, it is challenging to recognize every food item present in the image. The variability of individual consumption leads to further complexity, making it difficult to offer fully automatic and accurate EI estimation.

2.4.2.5 Wearable Devices - Sensors

Recent advances in wearable sensor technologies have shown great potential for the estimation of EI. The development of technology has further allowed supplying rich information about food intake by means of microstructure parameters such as eating episode duration, rate of ingestion, chewing frequency, chewing rate, bite size and others [71]. These parameters can potentially help researchers to better understand eating behaviors which may be associated with adiposity. The application of wearable sensors further reduces the burden of self-reporting of EI, alleviating some of the problems of misreporting. It also offers automatic detection of food intake events with minimal individual's active participation. Several investigators have attempted to use wearable sensor devices for estimating EI in both controlled laboratory settings and free-living environments. In this systematic review, nine studies that estimated EI or EI-related measures were thoroughly reviewed. Table 2-6 gives an overview of the included studies and Table 2-7 a summary of each of the wearable sensor methods for EI estimation.

The study of [72] demonstrated a method that can accurately detect food intake, differentiate between liquids and solids, and estimate the amount consumed (ingested mass). To predict ingested food mass, the investigators proposed separate mass models for solid and liquid intakes. The mass model for solid foods was developed based on linear regression against the measured weight of food using the total number of chews and swallows during ingestion period. The liquid mass model used only swallows as predictor considering the fact that liquids do not involve chewing.

Table 2-6. Overview of Included Studies Using Wearable Sensor Based Approaches

Study	Sample size	Age Mean±SD (range)	BMI Mean±SD (range)	Notes about subjects	Study setting	Estimated value	Primary outcome measures
Sazonov (2009) [72]	20 M: 11 F: 9	NA	29.0±6.4 kg/m^2	Healthy normal	Lab	Amount consumed (g)	Detecting periods of food intake and predicting mass of ingested foods
Amft (2009) [73]	8 M: 6 F: 2	20-35 years	NA	Healthy normal	Lab	Bite weight (g)	Bite weight prediction using food type and chewing recognitions
Liu (2012) [74]	6	NA	NA	Healthy normal	Lab	Proportio n of food consumed (%)	Estimates the proportion of food consumed using inter-image colour histogram distances
Fontana (2015) [75]	30	29±12 (19-58 years)	27.9±5.5 kg/m^2	Normal, Overwei ght Obese	Lab	EI (kcal)	Models based on counts of chews and swallows to estimate EI
Alshurafa (2015) [76]	20	20-40 years	NA	Healthy normal	Lab	Food types	Distinguish between food types, such as liquid and solid, hot and cold drinks, and hard and soft foods.
Salley (2016) [77]	280 M:132 F: 148	29.7 years	25.36±5.1 8 kg/m^2	Normal, Overwei ght Obese	Lab + free-living	*EI per bite* (kcal)	Estimating EI in free-living environment using bite count and mean EI per bite
Mirtchouk (2016) [78]	6 (total 30 meals)	NA	NA	NA	Lab + free-living	Amount consumed (g)	Food mass and food type estimation
Hezarjaribi (2017) [79]	30	18-30 years	NA	Healthy normal	Lab + free-living	Energy content (kcal)	EI estimation based on voice data
Thong (2017) [80]	NA	NA	NA	Healthy normal	Lab	Energy (KJ) and carbohydr ate (g)	Quantitative prediction on food nutrient contents, such as energy and carbohydrate.

Notes: BW, body weight; NA, not available

Table 2-7. Systematic Review of Wearable Sensor Based Method Towards EI Estimation

Study	Sensor used/sensor location	Food types/No. of foods included in study	Regression method used for analysis	Estimated equation	Group model (Gr) / Individual model (Ind)	Significant variables	Mass estimation	Energy density estimation	Compared against*	Accuracy/ performance
Sazonov (2009) [72]	Miniature microphone, Piezoelectric strain sensor	Meal, solid and liquids / NA	Linear	$M_S = 0.5\,(\bar{M}_{SW}^{S} \times N_{SW} + \bar{M}_{CHEW} \times N_{CHEW})$ $M_L = \bar{M}_{SW}^{L} \times N_{SW}$ M_S : Predicted mass of solid food (g) $\bar{M}_{SW}^{S}$: Subject's average mass per swallow of solid food (g) N_{SW} : Total number of swallow for solid or liquid food intake $\bar{M}_{CHEW}$: Average mass per chew (g) N_{CHEW} : Total number of chews M_L : Predicted mass of liquid food (g) $\bar{M}_{SW}^{L}$: Subject's average mass per swallow of liquid food (g)	Gr	Average mass per swallow, Average mass per chew, Number of swallows, Number of chews	Y	N	WFR	Average accuracy of mass (g) model for solid food intake : 91.79% Average accuracy of mass (g) model for liquid food intake: 83.76%
Amft (2009) [73]	Ear-pad chewing sound sensor	Solid / 3	Multiple linear	$W_i = a_0 + \sum_{k=1}^{N_V} a_k\, v_{ik}$ W_i : Bite weight prediction (g) a : Food-specific coefficients found by a least-square fit v : microstructure variables N_V : Total number of variables	Ind	Number of chewing event, Chewing duration	N	N	WFR	Food classification accuracy : 94% Lowest mean weight (g) prediction: 19.4% Largest mean weight (g) prediction: 31%

Liu (2012) [74]	Miniature camera and microphone	Meal	NA	NA	NA	Sound features: Energy entropy, Short time energy, Spectral roll-off, spectral centroid, spectral flux, spectral average of sub-bands, Zero crossing rate, Peak gaps between energy peaks.	N	N	NA	Not specified
Fontana (2015) [75]	Throat microphone, Piezoelectric strain sensor	Meal, solid and liquids /45	Linear	$Total\ mass\ ingested$ $M_T = M_S + M_L;$ $M_S = w_s \times MPSw_S \times N_{sw}^{s} + w_c \times (MPChew \times c_f) \times N_{chew};$ $M_L = MPSw_L \,.\, N_{sw}^{L};$ $EI = \sum_{i}^{N} m_{T_i} \,.\, CD_i$ M_T : Total mass ingested (g) M_S : Mass of solid food ingested (g) M_L : Mass of liquid food ingested (g) w_s : weight parameter for mass prediction using number of swallows w_c : weight parameter for mass prediction using number of chews $MPSw_S$: subject's average mass per swallow of solid food (g) $MPChew$: subject's average mass per chew (g) N_{sw}^{s} : total number of swallows for solid food intake N_{chew} : total number of chews c_f : correction factor $MPSw_L$: subject's average mass per swallow of liquid (g) N_{sw}^{L} : total number of swallows for liquid intake	Ind	Counts of chews and swallows	Y		WFR	Best accuracy: EI (kcal) model based on chews counts Reporting error (%): 30.42 ± 23.08

				m_{T_i} : consumed mass for the distinct food type i (g) CD_i : caloric density associated to the same food type i (kcals/g) N : total number of distinct foods types consumed in the meal						
Alshuraf-a (2015) [76]	piezoelectri c sensor	Liquid, solid, hot and cold drinks, hard and soft foods.	NA	NA	NA	NA	N	N	NA	Food type classification F-measure 90%
Salley (2016) [77]	Hand gesture sensor	Meal Solid, mixed and non-mixed /1844	Linear	$Estimated\ kilocalories\ per\ bite = -0.128 \times age + 6.167 \times sex(females = 0) + 0.034 \times height + 0.035 \times weight - 12.012 \times WHR + 22.294$ age : Age in years $height$: Height in inches $weight$: Weight in lb. WHR : Waist-to-hip ratio	Gr	Age, Sex, Height, Weight, Waist to hip ratio	N	N	WFR	Mean estimation error: -71.21±562.14 kcal
Mirtcho-uk (2016) [78]	Motion (head, both wrists) and acoustic sensors (customize d earbud)	Meal Solid, mixed and non-mixed / 1489 food and 285 drink Intakes	Random forest 40 trees	NA	Gr	NA	Y	N	WFR	Food classes classification accuracy: 82.7% Amount consumed (g) estimation error: 35.4%
Hezarja-ribi (2017) [79]	Audio recording in mobile app	Not specified	NA	NA	NA	Frequency domain features (energy, fundamental frequency	N	N	DB	EI (kcal) accuracy: 92.2%
Thong (2017) [80]	near-infrared (NIR) scanner	Liquid	Support vector And Partial least square	$E[Y\|X] = f(X, \beta)$; E : Energy content (KJ) X : an $n \times m$ matrix; n number of scans; m number of absorption; Y : observed values of	NA	NA	N	N	DB	For energy, the prediction error is less 2 KJ. For carbohydrate, the prediction error is around 0.12 g.

				energy (KJ) , carbohydrates (g) in food samples; f : Linear function β : least square errors;						

Notes: WFR: weighed food record; DB: Nutrition database/ Nutritional Fact labels

The authors in [73] proposed an ear-pad chewing sound sensor to identify chewing sequences during food intake and utilized microstructure variables extracted from chewing events to predict bite weight. To develop the model, the method classified foods, detected chewing events, and extracted microstructure variables (e.g. number of chewing events, chewing duration, chewing speed). Using both microstructure variables from chewing events and identified foods, a multiple linear regression model for bite weight was developed against weighed food record. To select informative variables, stepwise regression was carried out. Three different foods were selected for the study to evaluate the bite-weight model. The investigators proposed that the bite-weight prediction may potentially assist in estimating EI but did not actually provide these data.

In [74], the authors proposed a wearable sensor system consisting of a microphone and camera. Audio features were extracted in real-time to detect the chewing activity. The camera captured a video sequence for image analysis. Both time domain and frequency domain features were extracted from audio signals and applied to an extreme learning machine (ELM) classifier to classify chewing activity. The system further recognized food content in the images based on color histogram features. The proportion of histogram distances was used to estimate the consumed amount. The current version provided information about the preference of food types and consumption speed.

The Automatic Ingestion Monitor (AIM) combined piezoelectric strain gauge sensors for monitoring jaw motion, hand-to-mouth motion sensor, and body motion with a smartphone application [81]. The sensor system has been validated in several conditions and subject groups (both adults and infants) [82]–[85]. Recent publications suggested that the sensor system on the eyeglasses can accurately detect the chewing and be used in microstructure characterization [85], [86]. The study of [75], presented an EI estimation model from counts of chews and swallows. In laboratory settings, participants were asked to consume their meals wearing the sensor system. Participants eating behavior was recorded through a video monitoring system. Features extracted from sensors and video annotation were used to build EI models. EI was then obtained by multiplying mass intake by the energy density of the food item (obtained from food composition tables). The accuracy of EI estimation was reported for each eating occasion.

The study of [76] demonstrated a necklace-like wearable system which can potentially be used to monitor food and nutrient intake. The system included an embedded piezoelectric sensor which can detect skin motion in the lower trachea during chewing. The investigators proposed an algorithm based on time-frequency decomposition, spectrogram analysis of piezoelectric sensor signals, to accurately distinguish between food types, such as liquid and solid, hot and cold drinks, and hard and soft foods. The necklace transmitted data to a smartphone, which performed the processing of the signals, classified the food type, and provided visual feedback to the user to assist the user in monitoring their eating habits over time.

The study of [87] presented the use of wearable sensors to detect hand-to-mouth gestures and bite events. A wrist-worn "bite counter" device uses a gyroscope to detect hand-to-mouth motion related to eating. Using the bite counter, in [77], the authors proposed a multiple regression model to predict an individual's mean kilocalories per bite using anthropometric variables such as height, weight, age, sex, and waist-to-hip ratio regressed against measured kilocalorie intake. The study also compared EI estimation accuracy with the estimation made by the participants with/without prior knowledge about the energy content of food items. The accuracy of EI estimation was reported for each eating occasion.

The study of [78] proposed a multi-modal sensor system comprising of in-ear audio, head and wrist motion sensors to accurately classify food type and estimate the amount consumed. The authors applied features from audio and motion sensors to capture eating activities. The study obtained data from healthy participants wearing audio and motion sensors and used video observation as ground truth. Based on the features extracted from audio and motion sensor, chewing and duration of intake were detected. To recognize food type, the method trained a random forest classifier with 40 trees using all sensors variable. To estimate the weight of the food, along with sensor features, annotation features were also incorporated. A random forest regression with 40 trees was used to create the food weight estimation model. The method achieved 82.7% food classification accuracy and 35.4% absolute reporting error in mass estimation for 40 unique foods. Finally, the estimated amount consumed was utilized to estimate energy content from food databases. The accuracy of ingested mass was reported for each food items.

The study of [79] utilized voice to estimate EI and introduced a system called 'Speech2Health'. The system included speech processing, natural language processing (NLP), and text mining techniques. The voice data was first converted into text. By utilizing standard NLP and pattern mapping algorithm, the text was matched to nutrition-specific data. Finally, a matching algorithm was proposed to find the food name in the database and compute the EI values. The accuracy of EI was reported for each food items.

Recently, the authors of the study [80] proposed a method of estimating the energy and nutrients in food using a near-infrared (NIR) scanner. The scanner provides reflected spectrum pattern from the scans of liquid beverage samples. Using the NIR spectra features, support vector regression and partial least square regression models were developed to predict the energy and carbohydrate contents of food.

The use of wearable systems in estimating EI are increasing due to their potential for capturing detailed characteristics of all ingestion events without creating a reporting burden to the participant. Many of these are accurate in short-term studies in lab settings. However, further evaluations among large populations in community-dwelling settings are needed before their full utility and accuracy can be determined.

2.5 Discussion

The major goal of this review was to systematically explore the technology-driven methodologies that estimate EI or EI-related measures in humans. The systematic review revealed that various technologies were developed for EI estimation. Detailed review of selected methods revealed that each method has pros and cons and potentially needs further improvement.

The intake-balance method uses the mathematical models of human EI and metabolism to estimate EI. The study [19] proposed that the changes in EI over a period of time in free-living conditions can be estimated with through frequent measurements of daily BW, thus avoiding the need DLW/DXA scans and reducing the cost. Therefore, the intake-balance method provides a low-cost EI estimation method. However, a limitation of the intake-balance method is that the energy content of individual eating episodes, eating behavior and dietary choices are not measured. On top of that, these methods cannot objectively measure or estimate diet quality or eating behavior.

Technological advancement and the availability of smartphones made it possible for the general population to easily capture images of eating episodes. The technology-driven image-assisted methods further allowed portion size/volume estimation by a manual review by a trained nutritionist or automatic image recognition algorithm. Food type recognition and amount consumed estimation can also be automatically done using image processing algorithms [17], [56]–[59], [63], [65]–[67]. By utilizing the objective information from images, several image-assisted methods have shown great promise in EI estimation. Despite the advantages, the image-assisted methods are not free from issues that affect the performance. The energy content of a food may be difficult to estimate from an image: the type of food may be hard to identify, especially in mixed foods; foods of different energy content may have a similar appearance (e.g. diet vs. regular beverages); foods may be consumed from containers that obscure the food. A technological issue that affects most of the image-assisted approaches [17], [56], [58], [59], [63], [66] is that these require a fiducial marker to be placed in the image. The user has to carry the marker and remember to place it in the image capture area as they eat, which might not be feasible. Another key challenge for the implementation of image-based methods is the collection of data and complex analysis of data. Several methods use computationally complex algorithms to recognize foods and require high processing resources which are not generally available in clinical, rural, or third world settings. It is very important to recognize the food items that are actually consumed during the meal versus accurately estimate the leftover portions in order to measure EI. Apart from that, uneaten food items should be identified as people may simply be looking at food rather than eating it (e.g. in a restaurant setting, or eating in a group setting and looking at other peoples' food).

Development of wearable sensors for EI estimation has gained substantial attention in recent years. The wearable sensors are capable of capturing information about eating microstructure, which, potentially, may improve the accuracy of EI estimation by reducing the participant burden and need for self-report of every eating episode. Another potential advantage of wearable sensors is as a lower user burden compared with self-report methods.

However, the performance of wearable sensors is dependent on user's compliance with wearing the device and the social acceptability of the sensor.

Overall, our findings suggest several important aspects for consideration. These include how the studies reported the measure of accuracy, under what conditions the methods evaluated the performance, statistical analyses involved in methods, and the use of nutrition expertise in engineering literature.

We observed that the included studies either estimate EI or EI-related measures such as amount consumed/mass, portion size, volume, cumulative intake and food recognition. The measures of accuracy for EI estimation in [55], [60], [62], [64], [75], [77], [79], [80], are presented in terms of daily EI, meal/eating episodes EI and food item EI. On the other hand, the methods proposed in [17], [57]–[59], [63], [65]–[68], [70], [72], [73], [76], [78] reported accuracy of the estimation of EI-related measures such as amount consumed, volume, energy density, food recognition. Lastly, the methods presented in [19], [48], [56], [61], [69], [74] did not specify accuracy. The accuracy of any method is largely dependent on the study population, size of the dataset, set of foods present in the dataset, the study environment (laboratory or free-living), and the standard of comparison against which accuracy is evaluated. The best accuracy for EI estimation was reported in [64] as only 6% reporting error for food items EI. However, the study only included 19 types of food items in a controlled setting. Under free-living environment conditions, one can expect a large variety of foods, so the method needs to be validated under such conditions. Another study in [62] obtained 89% accuracy in estimating the energy content of just on one food item. The average accuracy of the method in [79] in energy content estimation for 30 participants was 92 % under lab conditions. However, the food types and number of food items used for the study were not specified. For food classification, the best accuracy was claimed in [66] where the authors obtained 99% accuracy for 30 food items. In the case of volume estimation, the study [70] obtained the highest accuracy of 92% considering 10 food items. Therefore, it is imperative to test the robustness of methods using real-life conditions where the system will encounter a variety of food items presented in different settings and using different containers/serveware.

Most of the studies reviewed were restricted to laboratory settings only. However, ensuring the robustness of the EI estimation method demands consistent performance in the free-living environment. The laboratory setting often lacks realism due to unnatural eating behavior [88] or potential behavior change, limited food items, mixed meals, eating in groups, etc. Depending on the study setting, an individual can consume meals fast/slow, skip meals, and/or take pauses between different parts of the meal to talk, etc. Therefore, laboratory experiments might not reflect real life in the free-living environment. Due to the controlled laboratory setting, many methods obtain a high degree of accuracy. Indeed, if a study is to be conducted in the laboratory, the gold standard of weighed food records which are inexpensive should be utilized.

In this review, the methods presented in [55], [56], [63], [77]–[79] conducted their study both in the laboratory and free-living conditions. In [55], the free-living EI was -152 ± 694 kcal/day compared to weighed food record, whereas in [77], the same method estimated -71.21 ± 562.14 kcal/meal. The study [64] claimed the best accuracy of EI estimation of 94%, however, the study setting was not specified. It can be summarized that realistic assessment needs to be done in the free-living environment.

The use of statistical analysis is also important when reporting accuracy. To realize the degree of error, the statistical analysis needs to be applied to explore how good the estimation is compared to relevant reference/gold standard. In a study [55], the validity of the method was tested by comparing weighed EI to estimated EI using Bland-Altman analysis. Similarly, the study [75] analyzed Bland-Altman plots when evaluating the accuracy of the method with respect to weighed food record. A paired samples *t-test* was used in the study [77] to compare mean EI estimation between the proposed method and participants' self-report. However, the remainder of the other studies lack or contain only cursory statistical analysis. In future, authors must substantiate their results using properly designed statistical analysis.

A common observation is that very few studies list a nutritionist/dietitian as one of the co-authors. Quite frequently, this leads to using of incorrect terminology, lack of proper evaluation methods and overstated claims. Of the reviewed studies, only 10 studies [17], [48], [49], [59], [64], [65], [69], [74], [75], [83] of the 29 (34%) listed had at least one nutritionist as a co-author. One of the recommendations based on this review is that the authors should seek collaboration within the community and strive to achieve sound engineering and nutrition science in publications.

While some of the methods show great promise under certain circumstances, objective, accurate and cost-efficient methods for estimating EI are yet to be developed. Such methods must provide 1) good accuracy in free-living conditions with large sample sizes in multiple studies and populations; 2) accurate recognition of a variety of food items, analysis of ingested mass, nutrient and energy contents of consumed foods with minimal effort; 3) characterization of food intake with minimal effort from the user; and 4) appropriate and rigorous statistical analysis.

2.6 Conclusions

This systematic review explored the technology-driven EI estimation methodologies proposed in the past two decades. Currently, there have been a range of methodologies estimating EI and related aspects such as food type recognition, amount consumed estimation, and portion size estimation. Each of the methods has its advantages and disadvantages. Due to the challenges and limitations existent in current methods, new methods are greatly needed for accurate EI estimation. In addition to developing new methods, improvement of existing energy estimation methods is also needed. Future works should be focused on exploring combining different approaches to accurately quantify EI.

2.7 References

[1] K. D. Hall, S. B. Heymsfield, J. W. Kemnitz, S. Klein, D. A. Schoeller, and J. R. Speakman, "Energy balance and its components: implications for body weight regulation123," *Am. J. Clin. Nutr.*, vol. 95, no. 4, pp. 989–994, Apr. 2012.

[2] D. A. Schoeller and D. Thomas, "Energy Balance and Body Composition," vol. 111, pp. 13–18, 2015.

[3] O. Müller and M. Krawinkel, "Malnutrition and health in developing countries," *CMAJ Can. Med. Assoc. J.*, vol. 173, no. 3, pp. 279–286, Aug. 2005.

[4] "Products - Health E Stats - Prevalence of Underweight Among Adults 2011-2012." [Online]. Available: https://www.cdc.gov/nchs/data/hestat/underweight_adult_11_12/underweight_adult_11_12.htm. [Accessed: 21-Nov-2017].

[5] "WHO | Obesity and overweight," *WHO*. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs311/en/. [Accessed: 16-Nov-2017].

[6] M. Ng *et al.*, "Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013," *The Lancet*, vol. 384, no. 9945, pp. 766–781, Aug. 2014.

[7] E. A. Finkelstein, J. G. Trogdon, J. W. Cohen, and W. Dietz, "Annual medical spending attributable to obesity: payer-and service-specific estimates," *Health Aff. Proj. Hope*, vol. 28, no. 5, pp. w822-831, Oct. 2009.

[8] T. Lehnert, D. Sonntag, A. Konnopka, S. Riedel-Heller, and H.-H. König, "Economic costs of overweight and obesity," *Best Pract. Res. Clin. Endocrinol. Metab.*, vol. 27, no. 2, pp. 105–115, Apr. 2013.

[9] J. G. Trogdon, E. A. Finkelstein, T. Hylands, P. S. Dellea, and S. J. Kamal-Bahl, "Indirect costs of obesity: a review of the current literature," *Obes. Rev.*, vol. 9, no. 5, pp. 489–500, Sep. 2008.

[10] K. M. Flegal, D. Kruszon-Moran, M. D. Carroll, C. D. Fryar, and C. L. Ogden, "Trends in Obesity Among Adults in the United States, 2005 to 2014," *JAMA*, vol. 315, no. 21, pp. 2284–2291, Jun. 2016.
[11] D. Sánchez-Carracedo, D. Neumark-Sztainer, and G. López-Guimerà, "Integrated prevention of obesity and eating disorders: barriers, developments and opportunities," *Public Health Nutr.*, vol. 15, no. 12, pp. 2295–2309, Dec. 2012.
[12] L. A. Gilmore, E. Ravussin, G. A. Bray, H. Han, and L. M. Redman, "An objective estimate of energy intake during weight gain using the intake-balance method," *Am. J. Clin. Nutr.*, vol. 100, no. 3, pp. 806–812, Sep. 2014.
[13] S. Jeran, A. Steinbrecher, and T. Pischon, "Prediction of activity-related energy expenditure using accelerometer-derived physical activity under free-living conditions: a systematic review," *Int. J. Obes.*, vol. 40, no. 8, pp. 1187–1197, Aug. 2016.
[14] A. H. C. Goris and K. R. Westerterp, "Underreporting of Habitual Food Intake Is Explained by Undereating in Highly Motivated Lean Women," *J. Nutr.*, vol. 129, no. 4, pp. 878–882, Apr. 1999.
[15] A. H. Goris, M. S. Westerterp-Plantenga, and K. R. Westerterp, "Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake," *Am. J. Clin. Nutr.*, vol. 71, no. 1, pp. 130–134, 2000.
[16] G. Johansson, A. Wikman, A. M. Ahrén, G. Hallmans, and I. Johansson, "Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living," *Public Health Nutr.*, vol. 4, no. 4, pp. 919–927, Aug. 2001.
[17] F. Zhu *et al.*, "The Use of Mobile Devices in Aiding Dietary Assessment and Evaluation," *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 4, pp. 756–766, Aug. 2010.
[18] J. A. Higgins *et al.*, "Validation of photographic food records in children: are pictures really worth a thousand words?," *Eur. J. Clin. Nutr.*, vol. 63, no. 8, pp. 1025–1033, 2009.
[19] K. D. Hall and C. C. Chow, "Estimating changes in free-living energy intake and its confidence interval," *Am. J. Clin. Nutr.*, vol. 94, no. 1, pp. 66–74, Jul. 2011.
[20] M. I. Goran, "Energy Metabolism and Obesity," *Med. Clin. North Am.*, vol. 84, no. 2, pp. 347–362, Mar. 2000.
[21] J. L. Hargrove, "History of the Calorie in Nutrition," *J. Nutr.*, vol. 136, no. 12, pp. 2957–2961, Dec. 2006.
[22] G. K. Grunwald, H. M. Seagle, J. C. Peters, and J. O. Hill, "Quantifying and separating the effects of macronutrient composition and non-macronutrients on energy density," *Br. J. Nutr.*, vol. 86, no. 2, pp. 265–276, Aug. 2001.
[23] R. P. Shook, G. A. Hand, and S. N. Blair, "Top 10 research questions related to energy balance," *Res. Q. Exerc. Sport*, vol. 85, no. 1, pp. 49–58, Mar. 2014.
[24] B. J. Rolls, E. A. Bell, and M. L. Thorwart, "Water incorporated into a food but not served with a food decreases energy intake in lean women," *Am. J. Clin. Nutr.*, vol. 70, no. 4, pp. 448–455, Oct. 1999.
[25] J. A. Ello-Martin, J. H. Ledikwe, and B. J. Rolls, "The influence of food portion size and energy density on energy intake: implications for weight management," *Am. J. Clin. Nutr.*, vol. 82, no. 1 Suppl, pp. 236S-241S, Jul. 2005.
[26] J. M. Lattimer and M. D. Haub, "Effects of Dietary Fiber and Its Components on Metabolic Health," *Nutrients*, vol. 2, no. 12, pp. 1266–1289, Dec. 2010.

[27] G. Livesey *et al.*, “Determination of digestible energy values and fermentabilities of dietary fibre supplements: a European interlaboratory study in vivo,” *Br. J. Nutr.*, vol. 74, no. 3, pp. 289–302, Sep. 1995.
[28] “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.” [Online]. Available: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000097. [Accessed: 14-Nov-2017].
[29] E. Sazonov, M. Farooq, and E. Melanson, “Assessment of Ingestion by Chewing and Swallowing Sensors,” *Advances in the Assessment of Dietary Intake.*, 15-Aug-2017. [Online]. Available: https://www.taylorfrancis.com/. [Accessed: 09-Jul-2018].
[30] C. A. Vereecken *et al.*, “Development and evaluation of a self-administered computerized 24-h dietary recall method for adolescents in Europe,” *Int. J. Obes.*, vol. 32, no. SUPPL. 5, pp. S26–S34, 2008.
[31] D. Lansky and K. D. Brownell, “Estimates of food quantity and calories: errors in self-report among obese patients,” *Am. J. Clin. Nutr.*, vol. 35, no. 4, pp. 727–732, Apr. 1982.
[32] “Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool.” [Online]. Available: https://epi.grants.cancer.gov/asa24/. [Accessed: 21-Nov-2017].
[33] C. A. Blanton, A. J. Moshfegh, D. J. Baer, and M. J. Kretsch, “The USDA Automated Multiple-Pass Method Accurately Estimates Group Total Energy and Nutrient Intake,” *J. Nutr.*, vol. 136, no. 10, pp. 2594–2599, Oct. 2006.
[34] “Fitbit Official Site for Activity Trackers & More.” [Online]. Available: https://www.fitbit.com/home. [Accessed: 21-Nov-2017].
[35] “Lose It! - Weight loss that fits.” [Online]. Available: https://www.loseit.com/. [Accessed: 21-Nov-2017].
[36] M. B. E. Livingstone and A. E. Black, “Markers of the validity of reported energy intake,” *J. Nutr.*, vol. 133 Suppl 3, pp. 895S-920S, Mar. 2003.
[37] L. S. Muhlheim, D. B. Allison, S. Heshka, and S. B. Heymsfield, “Do unsuccessful dieters intentionally underreport food intake?,” *Int. J. Eat. Disord.*, vol. 24, no. 3, pp. 259–266, Nov. 1998.
[38] L. J. Martin, W. Su, P. J. Jones, G. A. Lockwood, D. L. Tritchler, and N. F. Boyd, “Comparison of energy intakes determined by food records and doubly labeled water in women participating in a dietary-intervention trial,” *Am. J. Clin. Nutr.*, vol. 63, no. 4, pp. 483–490, Apr. 1996.
[39] S. W. Lichtman *et al.*, “Discrepancy between self-reported and actual caloric intake and exercise in obese subjects,” *N. Engl. J. Med.*, vol. 327, no. 27, pp. 1893–1898, Dec. 1992.
[40] W. V. Rumpler, M. Kramer, D. G. Rhodes, A. J. Moshfegh, and D. R. Paul, “Identifying sources of reporting error using measured food intake,” *Eur. J. Clin. Nutr.*, vol. 62, no. 4, pp. 544–552, Apr. 2008.
[41] J. A. Tooze, A. F. Subar, F. E. Thompson, R. Troiano, A. Schatzkin, and V. Kipnis, “Psychosocial predictors of energy underreporting in a large doubly labeled water study,” *Am. J. Clin. Nutr.*, vol. 79, no. 5, pp. 795–804, May 2004.
[42] L. S. Freedman *et al.*, “Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake,” *Am. J. Epidemiol.*, vol. 180, no. 2, pp. 172–188, Jul. 2014.
[43] K. R. Westerterp, L. Wouters, and W. D. van Marken Lichtenbelt, “The Maastricht protocol for the measurement of body composition and energy expenditure with labeled water,” *Obes. Res.*, vol. 3 Suppl 1, pp. 49–57, Mar. 1995.

[44] H. A. Haugen, L.-N. Chan, and F. Li, “Indirect calorimetry: a practical guide for clinicians,” *Nutr. Clin. Pract. Off. Publ. Am. Soc. Parenter. Enter. Nutr.*, vol. 22, no. 4, pp. 377–388, Aug. 2007.
[45] D. A. Schoeller and J. M. Hnilicka, “Reliability of the doubly labeled water method for the measurement of total daily energy expenditure in free-living subjects,” *J. Nutr.*, vol. 126, no. 1, pp. 348S-354S, Jan. 1996.
[46] A. E. Black and T. J. Cole, “Within- and between-subject variation in energy expenditure measured by the doubly-labelled water technique: implications for validating reported dietary energy intake,” *Eur. J. Clin. Nutr.*, vol. 54, no. 5, pp. 386–394, May 2000.
[47] M. A. Laskey, “Dual-energy X-ray absorptiometry and body composition,” *Nutrition*, vol. 12, no. 1, pp. 45–51, Jan. 1996.
[48] D. M. Thomas, C. K. Martin, S. Heymsfield, L. M. Redman, D. A. Schoeller, and J. A. Levine, “A Simple Model Predicting Individual Weight Change in Humans,” *J. Biol. Dyn.*, vol. 5, no. 6, pp. 579–599, Nov. 2011.
[49] D. M. Thomas *et al.*, “A new universal dynamic model to describe eating rate and cumulative intake curves,” *Am. J. Clin. Nutr.*, vol. 105, no. 2, pp. 323–331, Feb. 2017.
[50] 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, Sep. 1982.
[51] S. Kikunaga, T. Tin, G. Ishibashi, D.-H. Wang, and S. Kira, “The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey,” *J. Nutr. Sci. Vitaminol. (Tokyo)*, vol. 53, no. 2, pp. 109–116, Apr. 2007.
[52] D.-H. Wang, M. Kogashiwa, and S. Kira, “Development of a new instrument for evaluating individuals’ dietary intakes,” *J. Am. Diet. Assoc.*, vol. 106, no. 10, pp. 1588–1593, Oct. 2006.
[53] D.-H. Wang, M. Kogashiwa, S. Ohta, and S. Kira, “Validity and reliability of a dietary assessment method: the application of a digital camera with a mobile phone card attachment,” *J. Nutr. Sci. Vitaminol. (Tokyo)*, vol. 48, no. 6, pp. 498–504, Dec. 2002.
[54] C. K. Martin, T. Nicklas, B. Gunturk, J. B. Correa, H. R. Allen, and C. Champagne, “Measuring food intake with digital photography,” *J. Hum. Nutr. Diet. Off. J. Br. Diet. Assoc.*, vol. 27 Suppl 1, pp. 72–81, Jan. 2014.
[55] C. K. Martin, H. Han, S. M. Coulon, H. R. Allen, C. M. Champagne, and 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.*, vol. 101, no. 03, pp. 446–456, 2009.
[56] W. Wu and J. Yang, “Fast food recognition from videos of eating for calorie estimation,” in *2009 IEEE International Conference on Multimedia and Expo*, 2009, pp. 1210–1213.
[57] M. Wazumi, X. H. Han, D. Ai, and Y. W. Chen, “Auto-recognition of food images using SPIN feature for Food-Log system,” in *2011 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT)*, 2011, pp. 874–877.
[58] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” *Pervasive Mob. Comput.*, vol. 8, no. 1, pp. 147–163, Feb. 2012.
[59] M. H. Rahman *et al.*, “Food Volume Estimation in a Mobile Phone Based Dietary Assessment System,” in *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, 2012, pp. 988–995.

[60] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, “Measuring Calorie and Nutrition From Food Image,” *IEEE Trans. Instrum. Meas.*, vol. 63, no. 8, pp. 1947–1956, Aug. 2014.
[61] A. Myers *et al.*, “Im2Calories: Towards an Automated Mobile Vision Food Diary,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1233–1241.
[62] P. McAllister, H. Zheng, R. Bond, and A. Moorhead, “Semi-automated system for predicting calories in photographs of meals,” in *2015 IEEE International Conference on Engineering, Technology and Innovation/ International Technology Management Conference (ICE/ITMC)*, 2015, pp. 1–6.
[63] W. Zhang, Q. Yu, B. Siddiquie, A. Divakaran, and H. Sawhney, “‘Snap-n-Eat’: Food Recognition and Nutrition Estimation on a Smartphone,” *J. Diabetes Sci. Technol.*, vol. 9, no. 3, pp. 525–533, May 2015.
[64] S. Fang, C. Liu, F. Zhu, E. J. Delp, and C. J. Boushey, “Single-View Food Portion Estimation Based on Geometric Models,” in *2015 IEEE International Symposium on Multimedia (ISM)*, 2015, pp. 385–390.
[65] M. Sun *et al.*, “eButton: A Wearable Computer for Health Monitoring and Personal Assistance,” *Proc. Des. Autom. Conf.*, vol. 2014, pp. 1–6, 2014.
[66] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine, and S. Shirmohammadi, “Food calorie measurement using deep learning neural network,” in *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, 2016, pp. 1–6.
[67] H. C. Liao, Z. Y. Lim, and H. W. Lin, “Food intake estimation method using short-range depth camera,” in *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, 2016, pp. 198–204.
[68] E. A. Akpro Hippocrate, H. Suwa, Y. Arakawa, and K. Yasumoto, “Food Weight Estimation Using Smartphone and Cutlery,” in *Proceedings of the First Workshop on IoT-enabled Healthcare and Wellness Technologies and Systems*, New York, NY, USA, 2016, pp. 9–14.
[69] H. L. McClung *et al.*, “Digital food photography technology improves efficiency and feasibility of dietary intake assessments in large populations eating ad libitum in collective dining facilities,” *Appetite*, vol. 116, no. Supplement C, pp. 389–394, Sep. 2017.
[70] H. Hassannejad, G. Matrella, P. Ciampolini, I. D. Munari, M. Mordonini, and S. Cagnoni, “A New Approach to Image-Based Estimation of Food Volume,” *Algorithms*, vol. 10, no. 2, p. 66, Jun. 2017.
[71] A. Doulah *et al.*, “Meal Microstructure Characterization from Sensor-Based Food Intake Detection,” *Front. Nutr.*, vol. 4, 2017.
[72] E. Sazonov *et al.*, “Toward Objective Monitoring of Ingestive Behavior in Free-living Population,” *Obesity*, vol. 17, no. 10, pp. 1971–1975, May 2009.
[73] O. Amft, M. Kusserow, and G. Troster, “Bite Weight Prediction From Acoustic Recognition of Chewing,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 6, pp. 1663–1672, Jun. 2009.
[74] J. Liu *et al.*, “An Intelligent Food-Intake Monitoring System Using Wearable Sensors,” in *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2012, pp. 154–160.
[75] J. M. Fontana *et al.*, “Energy intake estimation from counts of chews and swallows,” *Appetite*, vol. 85, pp. 14–21, Feb. 2015.

[76] N. Alshurafa *et al.*, “Recognition of Nutrition Intake Using Time-Frequency Decomposition in a Wearable Necklace Using a Piezoelectric Sensor,” *IEEE Sens. J.*, vol. 15, no. 7, pp. 3909–3916, Jul. 2015.
[77] J. N. Salley, A. W. Hoover, M. L. Wilson, and E. R. Muth, “Comparison between Human and Bite-Based Methods of Estimating Caloric Intake,” *J. Acad. Nutr. Diet.*, vol. 116, no. 10, pp. 1568–1577, Oct. 2016.
[78] M. Mirtchouk, C. Merck, and S. Kleinberg, “Automated Estimation of Food Type and Amount Consumed from Body-worn Audio and Motion Sensors,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2016, pp. 451–462.
[79] N. Hezarjaribi, S. Mazrouee, and H. Ghasemzadeh, “Speech2Health: A Mobile Framework for Monitoring Dietary Composition from Spoken Data,” *IEEE J. Biomed. Health Inform.*, vol. PP, no. 99, pp. 1–1, 2017.
[80] Y. J. Thong, T. Nguyen, Q. Zhang, M. Karunanithi, and L. Yu, “Predicting food nutrition facts using pocket-size near-infrared sensor,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 742–745.
[81] J. M. Fontana, M. Farooq, and E. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1772–1779, Jun. 2014.
[82] E. S. Sazonov and J. M. Fontana, “A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing,” *IEEE Sens. J.*, vol. 12, no. 5, pp. 1340–1348, May 2012.
[83] E. Sazonov *et al.*, “Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior,” *Physiol. Meas.*, vol. 29, no. 5, pp. 525–541, May 2008.
[84] M. Farooq and E. Sazonov, “Automatic Measurement of Chew Count and Chewing Rate during Food Intake,” *Electronics*, vol. 5, no. 4, p. 62, Sep. 2016.
[85] M. Farooq and E. Sazonov, “Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses with Piezoelectric Sensor,” *IEEE J. Biomed. Health Inform.*, vol. PP, no. 99, pp. 1–1, 2016.
[86] M. Farooq, M. A. McCrory, and E. Sazonov, “Reduction of energy intake using just-in-time feedback from a wearable sensor system,” *Obesity*, p. n/a-n/a, Feb. 2017.
[87] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting periods of eating during free-living by tracking wrist motion,” *IEEE J. Biomed. Health Inform.*, vol. 18, no. 4, pp. 1253–1260, Jul. 2014.
[88] A. Doulah, X. Yang, J. Parton, J. Higgins, M. McCrory, and E. Sazonov, “The importance of field experiments in testing of sensors for dietary assessment and eating behavior monitoring,” presented at the The 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC’18), Honolulu, HI, 2018.

CHAPTER 3
MEAL MICROSTRUCTURE CHARACTERIZATION FROM SENSOR-BASED FOOD INTAKE DETECTION

Published: A. Doulah, M. Farooq, X. Yang, J. Parton, M. McCrory, J. Higgins, E. Sazonov, “Meal microstructure characterization from sensor-based food intake detection” Frontiers in Nutrition, 4, 31.

Abstract

To avoid the pitfalls of self-reported dietary intake, wearable sensors can be used. Many food ingestion sensors offer the ability to automatically detect food intake using time resolutions that range from 23 ms to 8 min. There is no defined standard time resolution to accurately measure ingestive behavior or a meal microstructure. This paper aims to estimate the time resolution needed to accurately represent the microstructure of meals such as duration of eating episode, the duration of actual ingestion, and number of eating events. Twelve participants wore the automatic ingestion monitor (AIM) and kept a standard diet diary to report their food intake in free-living conditions for 24 h. As a reference, participants were also asked to mark food intake with a push button sampled every 0.1 s. The duration of eating episodes, duration of ingestion, and number of eating events were computed from the food diary, AIM, and the push button resampled at different time resolutions (0.1–30s). ANOVA and multiple comparison tests showed that the duration of eating episodes estimated from the diary differed significantly from that estimated by the AIM and the push button (p-value <0.001). There were no significant differences in the number of eating events for push button resolutions of 0.1, 1, and 5 s, but there were significant differences in resolutions

of 10–30s (p-value <0.05). The results suggest that the desired time resolution of sensor-based food intake detection should be ≤5 s to accurately detect meal microstructure. Furthermore, the AIM provides more accurate measurement of the eating episode duration than the diet diary.

3.1 Introduction

An accurate understanding of dietary habits necessitates tracking of the dynamic process of each eating episode, known as meal microstructure (1-3). Meal microstructure includes factors such as eating episode duration (the duration from the start of the meal to the end including pauses), duration of actual ingestion (time spent eating in a given eating episode), the number of eating events (a bite, potentially followed by a sequence of chews and swallows), rate of ingestion, chewing frequency, chewing efficiency and bite size (4). The meal microstructure is directly related to the ingestive behavior of individuals. Therefore, the study of meal microstructure may potentially yield new insights in the treatment of obesity and comorbid conditions. An accurate method of monitoring food intake is necessary to capture meal microstructure and provide a better understanding of eating behaviors. Further, such methods could potentially facilitate novel methods to reduce caloric intake and/or provide more effective self-assessment and feedback tools for those on a calorie restricted diet (5).

A wealth of preclinical data exists illustrating the importance of meal microstructure to caloric intake and weight control in animal models [6]–[8] but detailed human studies are rare. Among the few human studies that do exist, there are demonstrated differences in meal microstructure in obese versus lean people, men versus women, and among individuals in various states of health [9]–[11]. Furthermore, larger food portion sizes have been shown to increase bite size and eating rate, leading to higher caloric intake in overweight women [12]. In another study, eating the same meal slowly vs quickly increased energy expenditure via the thermic effect of food (TEF), and improved metabolic parameters [13]. However, there are currently no free-living data from clinical studies which inform how these changes might affect caloric intake and weight control.

The scarcity of human data with regard to meal microstructure may be related to the difficulty of obtaining accurate data, especially in a community-dwelling situation, in which dietary intake and meal microstructure fluctuate more than in a controlled laboratory setting. Traditional methods of assessing dietary intake such as food frequency questionnaires, diet records, and 24-hr dietary recall rely on self-report by participants (14–16). The self-reporting errors may be up to 50% of estimated intake, resulting in inaccurate assessment which may impact medical diagnosis or dietary interventions (17). The primary causes for inaccuracy in self-report include under or over reporting of all food items consumed and poor estimation of portion consumed. In addition, there is a possibility of a change in the eating behavior of individuals when they know they are being observed (18). Furthermore, traditional self-report methods do not provide important information about meal pattern/microstructure such as the number of bites or eating rate (19).

In an effort to monitor food intake behaviors, wearable sensor systems which integrate different sensor modalities have been proposed. Most of these methods use sensors that measure behavioral manifestations of eating, such as hand-to-mouth gestures, bites, chews, and swallows (20). Various approaches have been used, including an oral strain gauge sensor to measure tongue pressure and flexing during chewing (21), an ear-pad sound sensor to capture air-conducted vibrations while chewing (22), miniature microphones in the outer ear canal to capture chewing (23), an acoustic sensor worn around the neck to detect sounds made by the user's mouth and throat while eating (24), a combination of a 3D gyroscope and 3 proximity sensors worn in an earpiece to measure ear canal deformations (25), a watch-like sensor (Bite Counter) to track wrist motion during hand-to-mouth gestures (26), a textile capacitive sensor worn as a neckband which detected swallowing and physical activity (27), a 3-axis accelerometer worn as a smartwatch to detect hand-to-mouth gestures (28), a piezoelectric sensor-based microphone to assess chewing and swallowing sounds (29) and others. Signal processing and pattern recognition algorithms interpret the sensor signals to recognize food intake, often after segmenting the sensor signals into time intervals, or epochs, of fixed duration. The food intake recognition algorithm processes the fragment of the sensor signal within a given time interval and assigns it a label "food intake" or "not food intake." Thus, the duration with which the sensor signal is segmented in turn determines the time resolution of food intake detection. Table 3-1 lists the studies mentioned above with respective time resolutions.

Table 3-1. Overview of Food Intake Detection Related Literature with Sensor Time Resolution

Study	**Sensor/Device**	**Signal**	**Time resolution**
Pabler *et al.*[23]	Miniature microphone	Body sound	23 ms
Amft *et al.*[22]	Acoustic sensor	Chewing sound	125 ms
Yatani *et al.*[24]	Acoustic sensor	Body sound	186 ms
Dong *et al.*[26]	Inertial sensor	Arm movement	1s
Rahman *et al.*[29]	Piezoelectric sensor	Chewing, swallowing	1s – 5s
Sazonov *et al.*[31]	Acoustic sensor	Swallowing sounds	1.5s
Farooq *et al.*[35]	Piezoelectric strain sensor and accelerometer	Chewing and physical activity	3s
Bedri *et al.*[25]	3D gyroscope, proximity	Ear canal deformations	5 s
Stellar *et al.* [21]	Oral strain gauge	Tongue Pressure and flexing during chewing	Chart speed 5 s/inch
Thomaz *et al.*[28]	Inertial sensor	Arm movement	6s

Sazonov *et al.*[32]	Piezoelectric strain gauge	Chewing	15s, 30s, and 60s
Sazonov *et al.*[30]	Sensor	Chewing and swallowing frequency	30s
Fontana *et al.*[33]	Piezoelectric strain gauge	Jaw motion	30s
Farooq *et al.*[34]	Electroglottograph	Swallow	30s
Cheng *et al.*[27]	Textile capacitive sensor	Swallow, swallow frequency and physical activity	1.5 min – 8 min

Our group has been developing systems to characterize food intake behavior by non-invasive monitoring of swallowing and chewing [30], [31]–[33]. In an earlier study [30], we proposed detection of food intake based on chewing and swallowing frequency. The instantaneous swallowing frequency was averaged over a sliding window of 30 s. In another study, the automatic detection of food intake was based on swallowing sounds using a high fidelity microphone placed over the laryngopharynx [31]. The sound data were segmented into a series of overlapping epochs (375 ms, 750 ms, 1.5 s, 3 s) with the 1.5 s epoch demonstrating the highest recognition accuracy. We then explored swallowing detection with an electroglottograph using epochs 30s in duration [34]. We have also proposed non-invasive monitoring of chewing using a piezoelectric strain gauge sensor [32]. Three different time resolutions (15 s, 30 s, and 60 s) were evaluated to determine an appropriate window size for the detection of food intake with the 30 s resolution being most accurate. Recently, we presented and validated a novel wearable sensor system (Automatic Ingestion Monitor, AIM) for detecting food intake in community-dwelling conditions [33] by monitoring jaw motion. In that study, we used 30s windows to recognize food intake. More recently, we modified the AIM to monitor the deformation of the temporalis muscle during food intake [35]. In that study, the food intake was recognized with 99.85% accuracy and a time resolution of 3s.

Thus, we have established that the AIM detects food intake with a high degree of accuracy, but we have not yet determined which time resolution will provide the greatest accuracy in estimating the meal microstructure.

The purpose of this study was two-fold: (1) to characterize meal microstructure, including duration of eating episodes, duration of actual ingestion, and number of eating events in community-dwelling healthy young adults using the AIM, and (2) to determine the optimal time resolution for doing so. The recommendations on time resolution may be used in future studies to accurately evaluate the practical capabilities of existing and emerging methods of food intake detection.

3.2 Methods

3.2.1 Data Collection

Experimental data were collected from a total of 12 participants (6 male, 6 female) with mean age 26.7 y (SD ± 3.7) and mean body mass index (BMI) 24.4 kg/m^2 (SD ± 3.8). No individuals reported any medical condition that would affect normal food intake. All participants read and signed an informed consent document before the start of the experiment. The study was approved by the Internal Review Board at The University of Alabama. Each participant was asked to wear the sensor system (AIM) for a 24-hr period where food intake was ad libitum. During the experiment, participants were able to perform daily living activities without restrictions.

The AIM consisted of data collection module, worn on a lanyard around the neck and had interface for 3 different sensors:

a) *Jaw motion sensor* - to detect the characteristic motion of the jaw during chewing (32,36). This sensor was attached directly below the ear using medical adhesive.

b) *Hand gesture sensor* - to detect hand-to-mouth gestures associated to bites. It consisted of a radio frequency transmitter worn on the inner side of the dominant arm and an RF receiver on the data collection module operating in radio-frequency identification band of 125 kHz.

c) *Tri-axial accelerometer* - to detect body acceleration. This sensor was located in the data collection module.

d) *Push button* - as the reference method for reporting food intake. The accuracy of the push button report was tested in a controlled lab study (37) against video observation. The participants were asked to hold the push button in the non-dominant hand.

Sensor signals were acquired by the data collection module at a 1 kHz sampling frequency. All sensor signals were quantized with 12-bit resolution and transmitted via onboard Bluetooth to an Android smart phone. Apart from wearing the AIM, participants were also asked to keep a paper food diary noting the start and end times of each eating episode, what foods and beverages were consumed.

Participants were instructed to press the push button at the start of every bite of solid or semi-solid food or the start of a chewing episode. The button was held down for the duration of chewing and released at the end of the chewing episode. The button was also used to report beverage intake by pressing and holding the button for the duration of the beverage ingestion episode.

3.2.2 Food Intake Detection

Three different methods were used in this study to monitor food intake and meal microstructure: the food diaries, AIM and push button. In the diary-based method, the food intake information was directly obtained from the completed food diaries. To detect the food intake using the AIM, a feature extraction algorithm and classification model developed in [33] was used. The model automatically recognized food intake with 30s of resolution from AIM sensor signals.

3.2.3 Meal microstructure analysis

The microstructure parameters of each eating episode were assessed from the food intake indicated by the methods described in Section 2.2. An example of a single eating episode assessed by each of these methods is shown in Fig. 3-1.

The microstructure parameters extracted from the diary, AIM and push button were:

Number of eating events (N), defined as the number of active ingestion segments in an eating episode (meal). The number of eating events for food diary (N_D) was always 1. The number of eating events for button and AIM were represented as N_B and N_A respectively.

Eating episode duration (D_{EE}), the duration between the start and stop times of eating episode, including segments without food intake. In the case of the food diary, the difference between the reported start and stop times was defined as the eating episode duration (D_{EED}). The computation of duration from the AIM and the push button (D_{EEA} and D_{EEB}) used the algorithm described below.

Duration of actual ingestion (D_I), the duration of actual eating (subtracting any non-eating segments) within an eating episode. In the case of the food diary, the duration of actual ingestion (D_{ID}) was the same as eating episode duration (D_{EED}). The duration of each i-th atomic eating event for push button and AIM were expressed as α_i and β_i as shown in Fig. 3-1 and durations of actual ingestion were computed as, $D_{IB} = \sum_{i=1}^{N_B} \alpha_i$ and $D_{IS} = \sum_{i=1}^{N_A} \beta_i$.

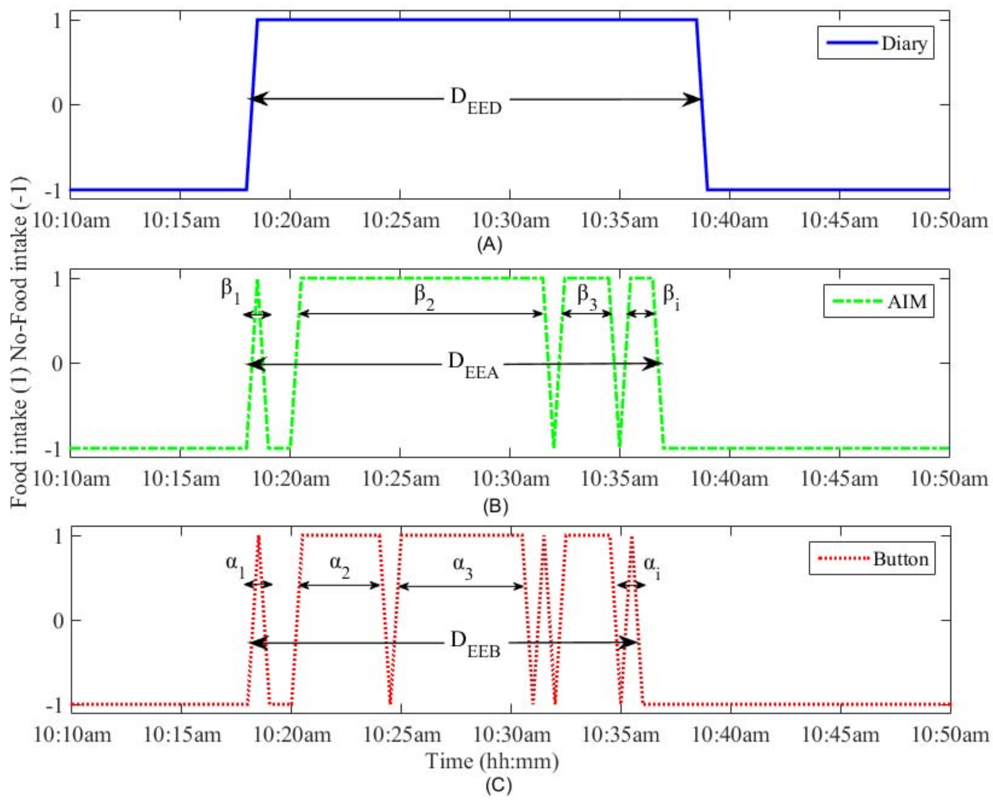

Figure 3-1. An eating episode reported by **(A)** the food diary, **(B)** AIM and **(C)** push button. The α_i and β_i represent the duration of individual eating event for push button and AIM respectively. The D_{EED}, D_{EEA}, D_{EEB} represent the eating episode duration from food diary, AIM and push button respectively.

The computation of D_{EEA} and D_{EEB} required determination of the boundaries of the eating episode. As shown in Fig. 3-1, eating episodes may have had pauses and/or breaks within the meal that needed to be "smoothed" for the estimation of duration. To smooth the signal, a function called 'kernel' is employed to compute the average of the neighboring data points. In this work, a smoothing kernel with the shape of a Gaussian (normal distribution, σ = standard deviation) curve was used on the AIM and push button. The food intake detection by AIM was performed on 30s intervals. To determine the suitable width of the kernel, the σ values from 1 to 6 (detection intervals) were tested. The optimal width was found to be at σ =

5 (150s). Then, the start and end points of each eating episode were determined by the intersection of the original and Gaussian-smoothed signals as shown in Fig. 3-2. The duration between A and B was defined as the eating episode duration. A script was written in MATLAB (Mathworks Inc., Natick, MA) to compute all durations.

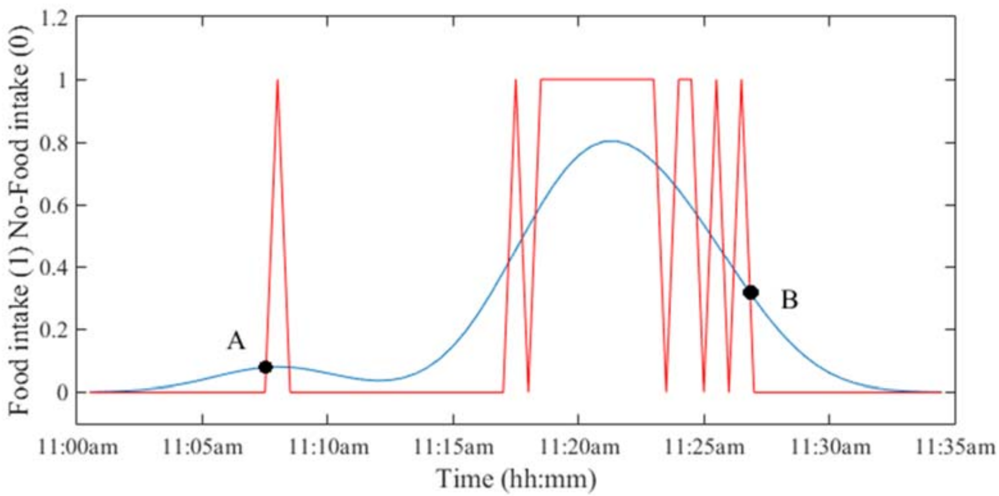

Figure 3-2. Determination of the duration of eating episode by using a Gaussian kernel function. The intersection of the original and smoothed signals provide the start time (point A) and end time (point B) of the eating episode. The duration of the eating episode is computed as the difference between time at points B and A.

3.2.4 Analysis of the time resolution

The signal from the push button reported food intake with the best time resolution (0.1s) was used for the analysis of the optimal resolution to capture meal microstructure parameters. Different time resolutions can be used to best characterize different aspects of meal microstructure. For example, chewing events can be captured in short time resolution of 3s. On the other hand, the detection of swallowing may need a window of as long as 30s. We investigated a range of time resolutions in capturing the microstructure parameters. To test a range of time resolutions representative of methods in Table 3-1, the push button signal was resampled with progressively longer window sizes (1s-30s, representing the range of detection intervals reported in the recent literature) using a resampling algorithm [38]. The microstructure parameters were then computed from the resampled signal and tested for equality using the statistical analysis described below.

3.2.5 Statistical Analysis

Statistical analysis was performed with SAS 9.0 (SAS Institute, Cary NC, USA) and Matlab 2015 (Mathworks Inc., Natick, MA). Differences in D_{EE} and D_I computed from the AIM, diary, and push button at time resolutions 0.1s-30s were analyzed with a linear mixed model with the participant as a random factor. If the mixed model showed a significant difference among methods, Tukey-Kramer post hoc multiple comparisons analysis were performed to determine which methods differed from each other. Data for the N_B at different time resolutions were analyzed by one-way repeated ANOVA to determine whether different time resolution yielded differing results. Since the parametric method did not pass the residual diagnostic criteria, a nonparametric Friedman test method was adopted for repeated ANOVA. If the ANOVA results were significant, Tukey-Kramer post hoc multiple comparisons test was performed to determine which time resolutions differed from each other. To assess the relative bias (mean difference) and random error (1.96 SD of the difference) between methods, the Bland and Altman plots [39] were investigated. Statistical significance was assumed at p-value < 0.05.

3.3 Results

In the study, out of the 12 participants, 4 participants did not provide timely information for the start and end of the eating episodes in the food diaries. For the remaining 8 participants, 23 eating episodes had complete time information (based on the dairies), whereas 6 eating episodes had partial time information (e.g. only start time and no end time) and therefore only 23 eating episodes were included in the analysis. The mixed model showed that there were statistically significant differences in meal microstructure between the food diary, the AIM, and the push button. Multiple comparison analysis showed that the D_{EE} and D_I from the food diary differed significantly from the AIM ($p< 0.001$) and the push button (p-value < 0.001), but the AIM and push-button results did not differ from each other (Fig. 3-3A). Participants' self-reported meal durations from the food diary were significantly over-reported in comparison to the AIM and the push button (Fig. 3-3B). The Bland Altman analysis in Fig. 3-4, shows good agreement between the AIM and push-button methods but poor agreement between the diary and other methods. With regard to eating episode durations in Fig. 3-4(A-C), the limits of agreement between the AIM and the push button were narrower compared to the limits of agreement between the diary and the push button and the diary with the AIM. A similar narrow range of limit agreement was found for actual ingestion durations shown in Fig. 3-4(F).

The narrow degree of dispersions in the distribution of D_{IB} and D_{IA} compared to D_{EEB} and D_{EEA} respectively indicate that the D_I is significantly smaller than the measured D_{EE} from boundaries (Fig. 3-5).

Results of the non-parametric Friedman test shows that at least two N computed from different resolutions of the push button method were significantly different (p-value < 0.05). Post hoc Tukey-Kramer test demonstrates that the N_B showed no significant differences for the high resolutions (0.1s, 1s, 5s) and exhibited differences for low resolutions (10s-30s; Fig. 3-6A). The results also indicate that the low resolutions (e.g. 15s) exhibit no significant differences for low resolutions (10s-30s) but significant differences for high resolutions (0.1s, 1s, 5s). Fig. 3-6B demonstrates the box-plot distributions of N_B at different resolutions. For low resolutions, the N_B begin to exhibit small mean with compact distributions indicating the loss of meal microstructure information.

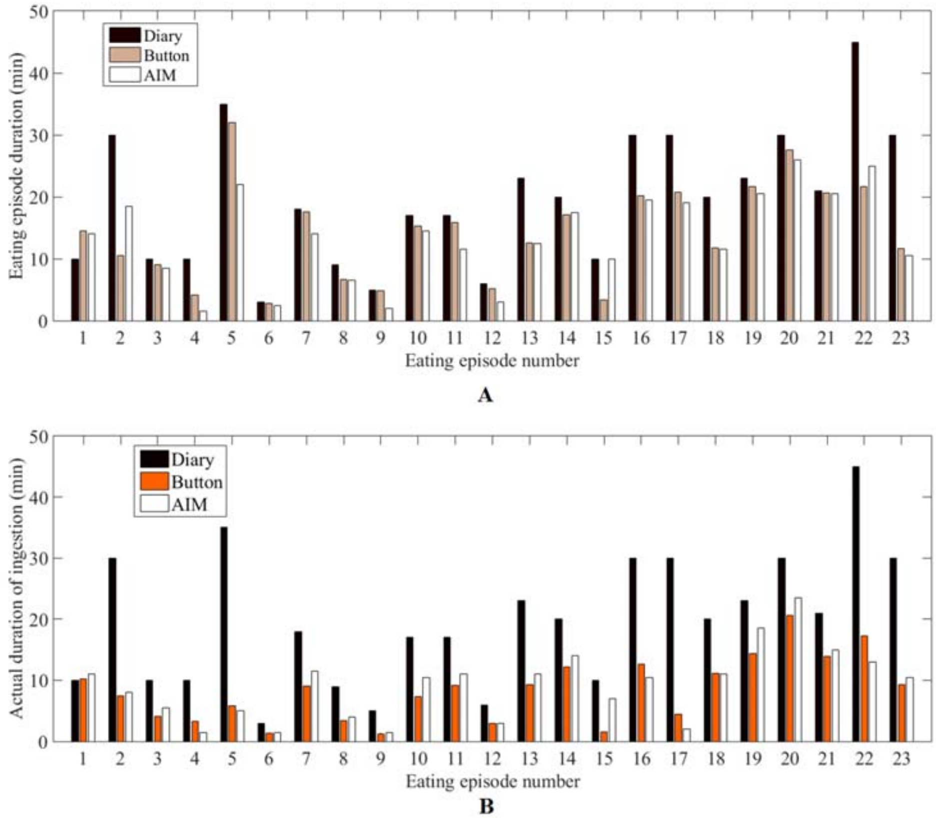

Figure 3-3. Duration across all eating episodes. **(A)** Eating episode duration. **(B)** Duration of actual ingestion.

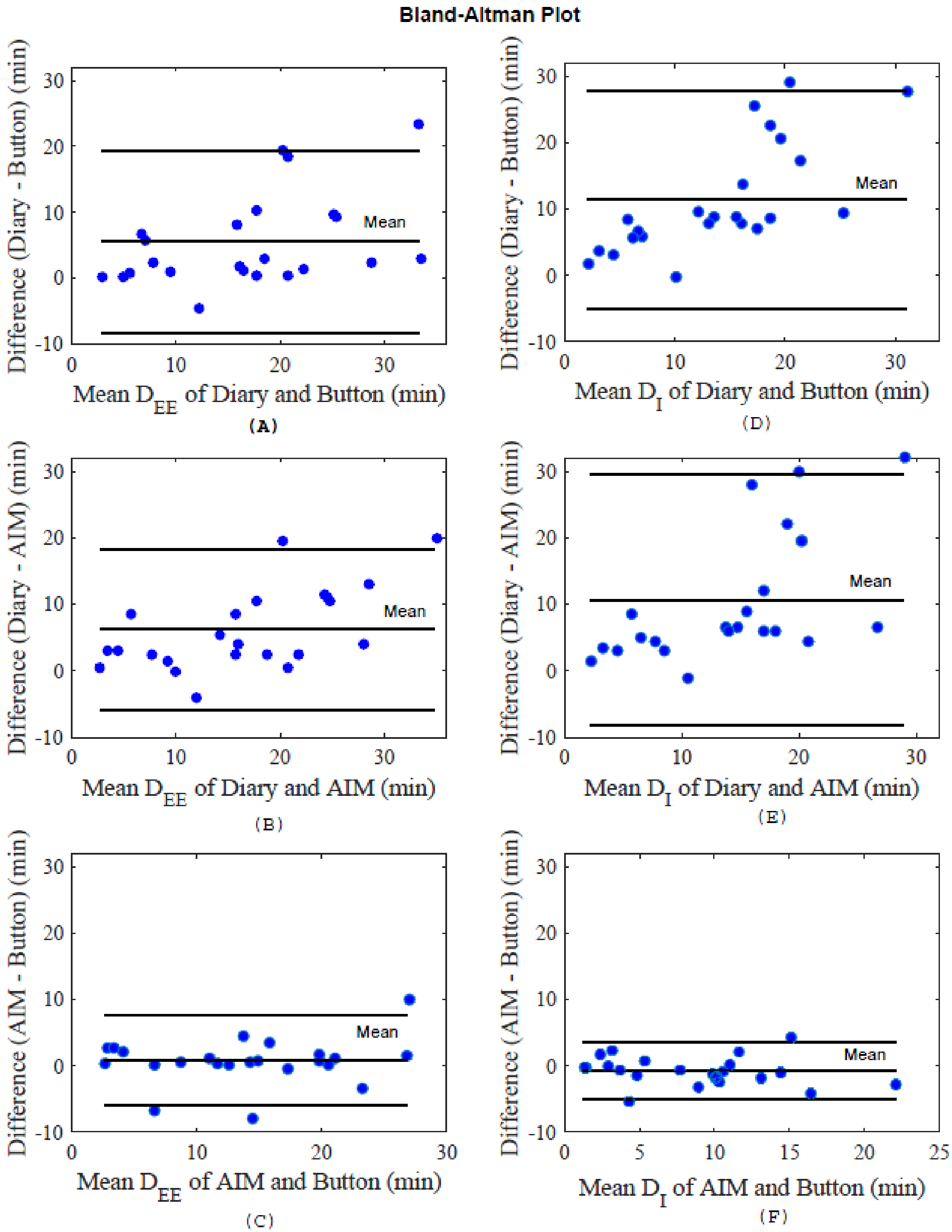

Figure 3-4. Bland-Altman plots for number of eating episodes **(A)** The eating episode duration (D_{EE}) of diary and button **(B)** Duration of ingestion (D_I) of diary and button **(C)** D_{EE} of diary and AIM **(D)** D_I of diary and AIM **(E)** D_{EE} of AIM and button **(F)** D_I of AIM and button. The blue dots represent eating episodes.

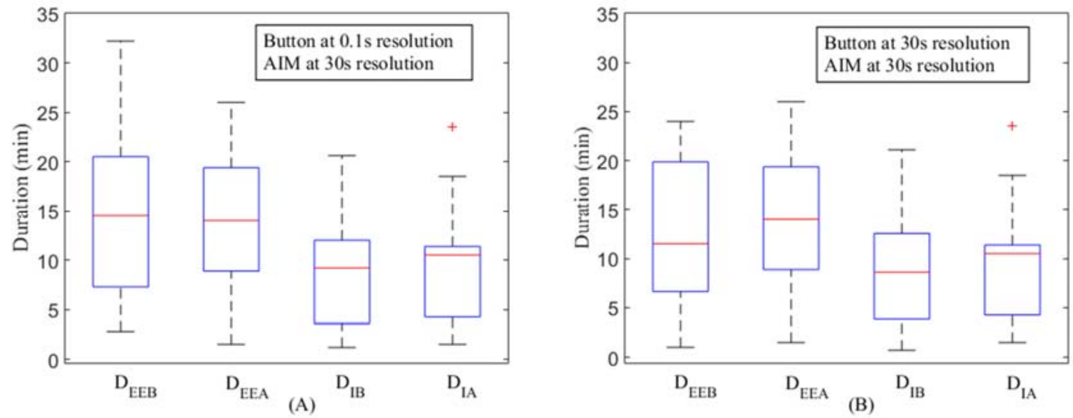

Figure 3-5. Box plots for measured durations of eating episodes (measured by push button - D_{EEB}, and AIM - D_{EEA}) and ingestion (button - D_{IB}, AIM - D_{IA}) at various time resolutions. **(A)** push button at 0.1s resolution and AIM at 30s resolution **(B)** push button at 30s and AIM at 30s resolution. The red line indicates the median. Upper and lower whiskers show the minimum and maximum changes within 25th and 75th percentile, respectively. Lower and upper horizontal blue lines (on the box) indicate 1st and 3rd quartile. Data points outside of the box are labeled as "outliers" and shown with a red cross.

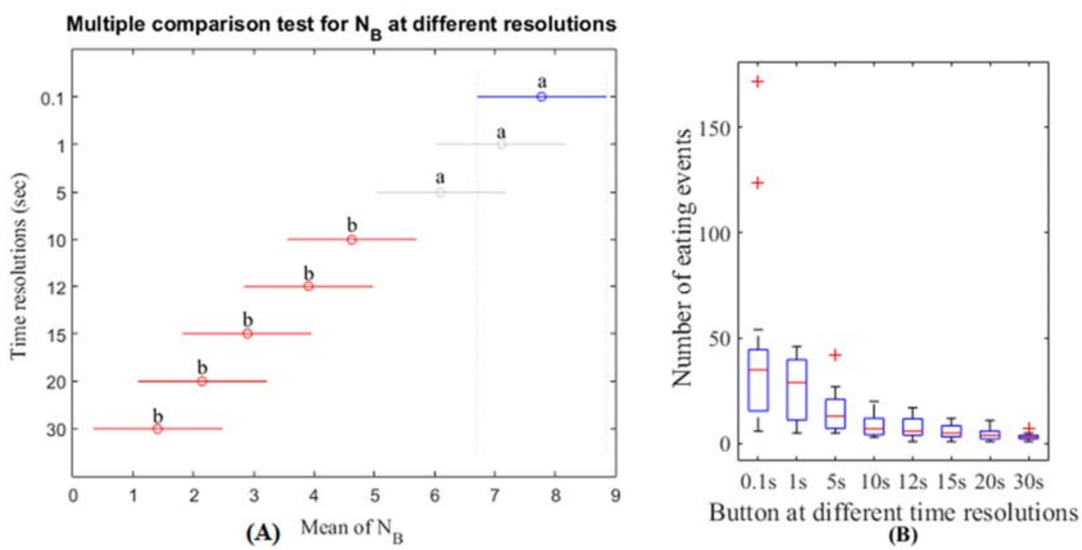

Figure 3-6. **(A)** Post hoc Tukey-Kramer test for the number of eating events from push button at different resolutions; The mean of N_B for each time resolution is represented by the symbol 'o'. The letter 'a' at the line indicates that the N_B are not significantly different whereas the letter 'b' indicates the N_B are significantly different. **(B)** Distribution of N_B at different resolutions; Data points outside of the box are labeled as "outliers" and shown with a red cross.

3.4 Discussion

In this study, we compared the assessment of meal microstructure parameters between food diary, a wearable sensor (AIM) method, and a reference push button. In an effort to find a standard time resolution of food intake detection, we also provided an analysis of time resolution that may be used to accurately evaluate meal microstructure parameters. The major findings suggested that compared with the push button (reference method), the AIM sensor provided more accurate meal microstructure information relative to the food diary. Furthermore, we found a sensor time resolution of 5 seconds was adequate to evaluate the meal microstructure parameters. These findings imply that to characterize meal microstructure, the AIM sensor-based dietary assessment method would be preferred over traditional food diary. Moreover, a sensor time-resolution of < 5 seconds should be used in future studies to best characterize meal microstructure using the sensor technologies.

Meal duration has significant effects on the total energy consumed [40], thus, accurate estimation of the meal/eating episode duration and duration of actual ingestion are very important for understanding eating behavior. We observed that the food diary over-estimated the eating occasion duration for most of the eating episodes compared with the AIM and reference method. One potential reason for the difference is the inherent reporting error of self-reported food diaries. On the other hand, the AIM does not depend on self-report, therefore, one would expect the AIM to be more accurate compared to the food diary. Bland-Altman analysis showed that the estimated durations from AIM had the lowest relative bias and the narrowest limit of agreement with the push button compared to the food diary. These findings demonstrate that the AIM could provide more accurate information about the eating episode duration and actual ingestion duration in comparison to the food diary. In addition, AIM could potentially offer information about the within meal behavior by examining meal microstructure. Therefore the results suggest that unlike the food diary, the AIM can potentially detect the process of food ingestion and measure microstructure parameters without creating a reporting burden for the user. The AIM can potentially offer a further characterization of the ingested foods analyzing the microstructure parameter which is not possible from food diary. An illustrative example is duration of actual ingestion vs. duration of the eating episode. Even the best modern electronic diary is not capable of assessing the actual time spent eating and time spent in other activities during a meal. Use of sensor technology allows to measure these microstructure parameters and potentially use them as a metric in comparing food intake in different individuals.

Our findings suggest that the duration of actual ingestion was significantly lower than the eating episode duration because an eating episode typically consisted of several eating events and intra-meal pauses. These pauses can be short or long depending on individual eating habits or surroundings and could lead to substantial differences between eating episode duration and actual ingestion duration. The distributions of $D_{IB,}$ and D_{IA} were significantly compact and lower than the distributions of D_{EEB} and D_{EEA} (Fig. 3-5). It is also evident that there was no significant difference between the median lines among respective time resolutions which implied that AIM could accurately estimate the duration of ingestion even with longer, 30s resolution. To provide a comparison with the same resolution, push-button signals was resampled and durations were computed (Fig. 3-5B). The D_{EEB} at 30s resolution exhibited compact quartile ranges compared to D_{EEB} at 0.1s. The median lines of both D_{EEB} and D_{IA} at 30s started indicating small difference with the median lines of D_{EEA} and D_{IA} at 30s. A potential reason could be the resampling of push button to 30s that reduced the number of eating events with each eating episode.

The use of a range of time resolutions of the sensor (1s – 30s) was important because it represents the state of the art in sensor detection of food intake reported in recent literature. A part of reason is that the selection of time resolution for a given sensor is dependent on the physical phenomena being captured. For instance, the short duration may better capture the microstructure properties of food intake, but the time resolution may be limited by the nature of the physiological process used for detection of food intake. When using swallows to detect food intake, the ingestion is manifested as an increase in swallowing frequency from approximately 2 swallows per minute to >4 swallows per minute, thus limiting the time resolution to approximately 30 s [30]. Chewing has a frequency of 0.94-2Hz and therefore may employ a time resolution (otherwise known as detection window) as short as 3 s [35]. Chewing sounds occur mostly in the range 1 kHz−2 kHz [41] and therefore, detection windows as short as 23 ms [23] were reported for the sounds. The study of [42] reported that consecutive bites are separated by 5-15s for foods with different levels of palatability. In general, longer detection windows include more physiological events of interest (chews, swallows, etc.) and, therefore, may provide a higher accuracy compared to shorter detection windows. Thus, there is a potential tradeoff between optimal time resolution and the accuracy of food intake detection, which in turn has an impact on the accuracy of representing the meal microstructure. Results demonstrate that the time resolutions of 10s-30s indicated small N_B with compact distributions and resolutions of 0.1s-5s indicated comparatively spread distributions with large N_B. Therefore, it can be inferred that the time resolutions of 0.1s-5s describe the meal microstructure accurately.

In the AIM sensor module, the jaw motion sensor signals were analyzed by means of chew detection to monitor food intake. The signals were divided into non-overlapping segments of 30s due to the historical reasons of the technology development. However, the analysis of N_B suggested that the desired time resolution of sensor-based food intake detection should be ≤5s to preserve the meal microstructure. Such window duration is potentially supported the frequency range of chewing (1.25-2.5 Hz) and use of shorter detection windows for detection of chewing by the AIM will be explored in the future.

The major strengths of this study were the performance comparison between the food diary and AIM, and finding out a potential time resolution to capture meal microstructure by sensor-based methods. A limitation of this study was utilizing the push button as the reference method for assessing the microstructure of a meal. There is a possibility that using a handheld pushbutton might change the eating behavior of the participants as one hand is occupied by the button. While the button may have imposed some changes in hand use, many foods is consumed with one hand only. This is especially true for quick ingestive events, such as grabbing and consuming a small food item. Therefore, we believe that the microstructure captured in this experiment is representative of real-world eating. The participants may also press or release the button accidentally and misreported the intake. While such errors are possible, hand button is arguably is one of the best ways to assess eating microstructure in free living. As our experience shows, the quality of video annotation of eating microstructure degrades greatly in conditions close to free-living due to the difficulties in interpretation of complex ingestive behaviors (e.g. eating while talking). Great hand dexterity of humans (for example, demonstrated in complex button press combination in computer work and gaming) combined with self-perception ("feeling") of the ingestion process allows for accurate representation even of transient events such as swallowing [37]. As a potential alternative to handheld button, a foot pedal may potentially be utilized to record the ground truth so that the participant can use both hands while consuming the food. Further studies will need to involve more participants and a larger number of eating episodes. The future work should also investigate more complex microstructure parameters such as eating rate.

3.5 Conclusion

Precise characterization of microstructural properties of a meal is a key to capturing of accurate eating patterns using sensor-based methods. Results suggested that the duration of the eating episodes estimated from food diaries was significantly different from the duration estimated by the AIM and pushbutton and furthermore, that the AIM is more accurate that food diaries. Based on this work, the desired time resolution of sensor-based food intake detection should be ≤5s to adequately document the meal microstructure parameters.

3.6 Reference

[1] Chapelot D. "2 - Quantifying satiation and satiety," in *Satiation, Satiety and the Control of Food Intake* Woodhead Publishing Series in Food Science, Technology and Nutrition. (Woodhead Publishing), 12–39. Available at: http://www.sciencedirect.com/science/article/pii/B9780857095435500025 [Accessed July 11, 2016]

[2] Bellisle F, Le Magnen J. The structure of meals in humans: Eating and drinking patterns in lean and obese subjects. *Physiol Behav* (1981) **27**:649–658. doi:10.1016/0031-9384(81)90237-7

[3] Bellisle F, Lucas F, Amrani R, Le Magnen J. Deprivation, palatability and the micro-structure of meals in human subjects. *Appetite* (1984) **5**:85–94.

[4] Guss JL, Kissileff HR. Microstructural analyses of human ingestive patterns: from description to mechanistic hypotheses. *Neurosci Biobehav Rev* (2000) **24**:261–268. doi:10.1016/S0149-7634(99)00079-2

[5] Farooq M, McCrory MA, Sazonov E. Reduction of Energy Intake using Just-In-Time Feedback from a Wearable Sensor System. *Obesity* (2017) **accepted**:

[6] Chenu O, Vuillerme N, Bucki M, Diot B, Cannard F, Payan Y. TexiCare: An innovative embedded device for pressure ulcer prevention. Preliminary results with a paraplegic volunteer. *J Tissue Viability* (2013) **22**:83–90. doi:10.1016/j.jtv.2013.05.002

[7] Janhunen SK, van der Zwaal EM, La Fleur SE, Adan RAH. Inverse Agonism at α2A Adrenoceptors Augments the Hypophagic Effect of Sibutramine in Rats. *Obesity* (2011) **19**:1979–1986. doi:10.1038/oby.2011.51

[8] Glass MJ, Grace MK, Cleary JP, Billington CJ, Levine AS. Naloxone's effect on meal microstructure of sucrose and cornstarch diets. *Am J Physiol - Regul Integr Comp Physiol* (2001) **281**:R1605–R1612.

[9] Spiegel TA, Kaplan JM, Tomassini A, Stellar E. Bite size, ingestion rate, and meal size in lean and obese women. *Appetite* (1993) **21**:131–145. doi:10.1016/0195-6663(93)90005-5

[10] Spiegel TA. Rate of intake, bites, and chews-the interpretation of lean-obese differences. *Neurosci Biobehav Rev* (2000) **24**:229–237.

[11] Elfhag K, Barkeling B, Carlsson AM, Rössner S. Microstructure of eating behavior associated with Rorschach characteristics in obesity. *J Pers Assess* (2003) **81**:40–50. doi:10.1207/S15327752JPA8101_04

[12] Almiron-Roig E, Tsiountsioura M, Lewis HB, Wu J, Solis-Trapala I, Jebb SA. Large portion sizes increase bite size and eating rate in overweight women. *Physiol Behav* (2015) **139**:297–302. doi:10.1016/j.physbeh.2014.11.041

[13] Reddy NL, Peng C, Carreira MC, Halder L, Hattersley J, Piya MK, Tripathi G, Randeva HS, Casanueva FF, McTernan PG, et al. Enhanced thermic effect of food, postprandial NEFA suppression and raised adiponectin in obese women who eat slowly. *Clin Endocrinol (Oxf)* (2015) **82**:831–837. doi:10.1111/cen.12652

[14] Day N, McKeown N, Wong M, Welch A, Bingham S. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. *Int J Epidemiol* (2001) **30**:309–317. doi:10.1093/ije/30.2.309

[15] Muhlheim LS, Allison DB, Heshka S, Heymsfield SB. Do unsuccessful dieters intentionally underreport food intake? *Int J Eat Disord* (1998) **24**:259–266. doi:10.1002/(SICI)1098-108X(199811)24:3<259::AID-EAT3>3.0.CO;2-L

[16] Thompson FE, Subar AF. "Dietary assessment methodology," in *Nutrition in the Prevention and Treatment of Disease* (Academic Press, San Diego, CA).

[17] Goris AH, Westerterp-Plantenga MS, Westerterp KR. Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake. *Am J Clin Nutr* (2000) **71**:130–134.

[18] Pettitt C, Liu J, Kwasnicki RM, Yang G-Z, Preston T, Frost G. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. *Br J Nutr* (2016) **115**:160–167. doi:10.1017/S0007114515004262

[19] Gemming L, Ni Mhurchu C. Dietary under-reporting: what foods and which meals are typically under-reported? *Eur J Clin Nutr* (2016) **70**:640–641. doi:10.1038/ejcn.2015.204

[20] Lopez-Meyer P, Makeyev O, Schuckers S, Melanson EL, Neuman MR, Sazonov E. Detection of Food Intake from Swallowing Sequences by Supervised and Unsupervised Methods. *Ann Biomed Eng* (2010) **38**:2766–2774. doi:10.1007/s10439-010-0019-1

[21] Stellar E, Shrager EE. Chews and swallows and the microstructure of eating. *Am J Clin Nutr* (1985) **42**:973–982.

[22] Amft O, Kusserow M, Troster G. Bite Weight Prediction From Acoustic Recognition of Chewing. *IEEE Trans Biomed Eng* (2009) **56**:1663–1672. doi:10.1109/TBME.2009.2015873

[23] Päßler S, Wolff M, Fischer W-J. Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food. *Physiol Meas* (2012) **33**:1073–1093. doi:10.1088/0967-3334/33/6/1073

[24] Yatani K, Truong KN. BodyScope: A Wearable Acoustic Sensor for Activity Recognition. in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* UbiComp '12. (New York, NY, USA: ACM), 341–350. doi:10.1145/2370216.2370269

[25] Bedri A, Verlekar A, Thomaz E, Avva V, Starner T. A Wearable System for Detecting Eating Activities with Proximity Sensors in the Outer Ear. in *Proceedings of the 2015 ACM International Symposium on Wearable Computers* ISWC '15. (New York, NY, USA: ACM), 91–92. doi:10.1145/2802083.2808411

[26] Dong Y, Scisco J, Wilson M, Muth E, Hoover A. Detecting periods of eating during free-living by tracking wrist motion. *IEEE J Biomed Health Inform* (2014) **18**:1253–1260. doi:10.1109/JBHI.2013.2282471

[27] Cheng J, Zhou B, Kunze K, Rheinländer CC, Wille S, Wehn N, Weppner J, Lukowicz P. Activity Recognition and Nutrition Monitoring in Every Day Situations with a Textile Capacitive Neckband. in *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication* UbiComp '13 Adjunct. (New York, NY, USA: ACM), 155–158. doi:10.1145/2494091.2494143

[28] Thomaz E, Essa I, Abowd GD. A Practical Approach for Recognizing Eating Moments with Wrist-mounted Inertial Sensing. in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* UbiComp '15. (New York, NY, USA: ACM), 1029–1040. doi:10.1145/2750858.2807545

[29] Rahman T, Adams AT, Zhang M, Cherry E, Zhou B, Peng H, Choudhury T. BodyBeat: A Mobile System for Sensing Non-speech Body Sounds. in *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services* MobiSys '14. (New York, NY, USA: ACM), 2–13. doi:10.1145/2594368.2594386

[30] Sazonov E, Schuckers SAC, Lopez-Meyer P, Makeyev O, Melanson EL, Neuman MR, Hill JO. Toward Objective Monitoring of Ingestive Behavior in Free-living Population. *Obesity* (2009) **17**:1971–1975. doi:10.1038/oby.2009.153

[31] Sazonov E, Makeyev O, Lopez-Meyer P, Schuckers S, Melanson E, Neuman M. Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior. *IEEE Trans Biomed Eng* (2010) **57**:626–633. doi:10.1109/TBME.2009.2033037

[32] Sazonov E, Fontana JM. A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing. *IEEE Sens J* (2012) **12**:1340–1348. doi:10.1109/JSEN.2011.2172411

[33] Fontana JM, Farooq M, Sazonov E. Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior. *IEEE Trans Biomed Eng* (2014) **61**:1772–1779. doi:10.1109/TBME.2014.2306773

[34] Farooq M, Fontana JM, Sazonov ES. A Novel Approach for Food Intake Detection Using Electroglotography. *Submitted* (2013)

[35] Farooq M, Sazonov E. A Novel Wearable Device for Food Intake and Physical Activity Recognition. *Sensors* (2016) **16**:1067. doi:10.3390/s16071067

[36] Fontana JM, Sazonov ES. A robust classification scheme for detection of food intake through non-invasive monitoring of chewing. in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4891–4894. doi:10.1109/EMBC.2012.6347090

[37] Sazonov E, Schuckers S, Lopez-Meyer P, Makeyev O, Sazonova N, Melanson EL, Neuman M. Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiol Meas* (2008) **29**:525–541. doi:10.1088/0967-3334/29/5/001

[38] Yaroslavsky L. Boundary effect free and adaptive discrete signal sinc-interpolation algorithms for signal and image resampling. *Appl Opt* (2003) **42**:4166–4175.

[39] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* (1999) **8**:135–160. doi:10.1177/096228029900800204

[40] Pliner P, Bell R, Hirsch ES, Kinchla M. Meal duration mediates the effect of "social facilitation" on eating in humans. *Appetite* (2006) **46**:189–198. doi:10.1016/j.appet.2005.12.003

[41] Brochetti D, Penfield MP, Burchfield SB. Speech Analysis Techniques: A Potential Model for the Study of Mastication Sounds. *J Texture Stud* (1992) **23**:111–138. doi:10.1111/j.1745-4603.1992.tb00515.x

[42] Päßler S, Wolff M, Fischer W-J. Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food. *Physiol Meas* (2012) **33**:1073–1093. doi:10.1088/0967-3334/33/6/1073

[43] Bellisle F, Guy-Grand B, Le Magnen J. Chewing and swallowing as indices of the stimulation to eat during meals in humans: effects revealed by the edogram method and video recordings. *Neurosci Biobehav Rev* (2000) **24**:223–228. doi:10.1016/S0149-7634(99)00075-5.

CHAPTER 4
CLUSTERING OF FOOD INTAKE IMAGES INTO FOOD AND NON-FOOD CATEGORIES

Published: A. Doulah, E. Sazonov, "Clustering of Food Intake Images into Food and Non-Food Categories" In: Rojas I., Ortuño F. (eds) Bioinformatics and Biomedical Engineering. IWBBIO 2017. Lecture Notes in Computer Science, 2017, vol 10208. Springer, Cham.

Abstract

For dietary assessment, food images could be useful to identify foods, portion sizes and estimate calorie in meals, either by human nutritionists or through image recognition. Images captured by a wearable camera during eating may include both food and non-food images. To avoid reviewing each image, only informative food images should be included in the analysis and non-food image should be discarded. This study proposed a methodology for clustering of food images into food and non-food groups based on histogram matching, without explicit recognition of the image content. Data was collected from 7 participants wearing an eyeglasses camera. A total of 10 meals were recorded at the sampling rate of 5 s, yielding a total of 1077 images. Each image was labeled by a human rater as a food or non-food image. Histogram matching with Bhattacharyya distance was applied to form a similarity matrix for extracted images from each meal. Both k-means and affinity propagation (AP) algorithms were investigated to cluster the images. Results show that the overall average food image clustering accuracy in respect to human annotation for 10 videos was 0.93 ± 0.04 for AP and 0.90 ± 0.03 for k-means. Similarly, overall average non-food image clustering accuracy was 0.81 ± 0.06 for AP and 0.70 ± 0.09 k- means method.

4.1 Introduction

The dietary assessment methods aim to provide an accurate estimation of foods consumed, portion size and calorie intake. To alleviate problems associated with self-reported food intake, various methods of dietary monitoring have been proposed ranging from digital diaries with food images to wearable on body sensors. The technological advances in computer vision allowed the introduction of image/video analysis for diet assessment [1]. Typically, a food imagery based method acquires images before, during and after a meal using a wearable camera or a mobile device. The food images can be analyzed to identify different foods, their portion sizes and nutritional content. In early studies, image analysis was performed by human nutritionists. The study of [2] proposed an approach called the Remote Food Photography Method where the participants submit food images to a nutritionist for the analysis of energy intake estimation. Recently, by utilizing crowdsourcing technique, an application called PlateMate [3] estimated nutrition information from food images provided by users. The crowdsourcing used non- professional nutritionists to assess energy and nutrient content of food from digital images.

Recently, automatic image recognition gained popularity in identifying and quantifying food images. Typically, a set of features from the images are extracted and fed to a classifier to recognize various food items. Several methods have been proposed to recognize foods in the image via supervised and unsupervised machine learning techniques. A food recognition application was developed in [4] for the classification of fast-food images into four classes. The study extracted color (normalized RBG values), texture (local entropy, standard deviation, range) and context-based features from images and fed to the artificial neural network for food recognition. The study of [5] applied a support vector machine (SVM) classifier for the recognition of nineteen classes of foods utilizing a set of color (color components and pixel intensities) and texture (responses from Gabor filter) features. In [6], a simple Bayesian probabilistic classifier was used to match images of food items to offline food database containing images of homemade foods, fast-food and fruits. The authors used scale invariant feature transform (SIFT) features and clustered the features into visual words before feeding the classifier. The study of [7] proposed a pairwise classification method that uses the user's speech input to improve the food recognition process. The recognition method incorporated the use of color neighborhood and texton histogram features, Adaboost feature selection and SVM classifier. In [8], the authors created a database of fast-food images and videos in an effort to make a benchmark for food recognition. Based on color histograms and SIFT bag features, the proposed methods were evaluated for seven fast-food classes. In [9], a multiple kernel learning based food recognition method was proposed utilizing the bag of scale invariant feature transform, Gabor filter responses, and color histograms features. Recently, the study of [10] proposed automatic food recognition based on bag of features with support vector machine classifier. A mobile user interface was proposed in [11] that utilizes a client-server image recognition and portion estimation software to estimate calorie content. Some major potential hurdles in the automatic approach are to determine exact portion size and accurate food recognition. Apart from that, the accuracy of automatic recognition of a great variety of foods is still suboptimal. The image capture by mobile devices such as smartphones requires active participation by the person being monitored [2]. Such images are typically taken before and after a meal and are likely to contain clear images of the food and the leftovers. Image captured by wearable cameras does not require the active participation of the person being monitored and thus potentially reduces the user burden [12]. However, wearable cameras capture many images that may contain non-food items. Therefore, a method that could separate food images and non-food images is needed. In this paper, an approach is proposed so that only informative food images can be included in the analysis and non-food images can be discarded. We propose to use the histogram of color textures features directly to find similarities between food images and non-food images. A pairwise matching algorithm compared the meal images and created a similarity matrix for the image collection. Two different unsupervised clustering techniques were investigated to identify groups of food image/non-food images within a meal. The performance demonstrated the potential of image grouping without recognition.

4.2 Subject and Methods

Fig. 4-1 shows the flow diagram of the methodology of the proposed system. The following subsections describe each step in detail.

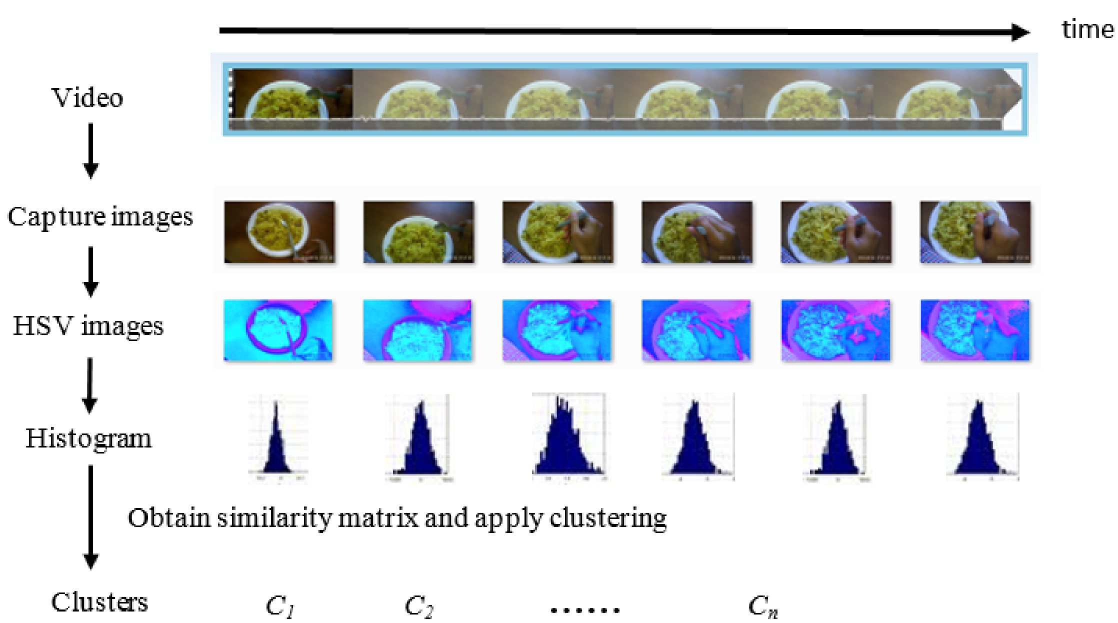

Figure 4-1. Flow of processing.

4.2.1 Data collection and annotation

The first step involved the recording of meal videos with a wearable camera. The camera was attached to the eyeglasses so that the images would contain most of the food items in food plate/bowl. A total of 7 subjects were recruited for the study. During the visit, subjects were asked to wear the camera during food consumption at breakfast, lunch or dinner. A total of 10 meal videos were recorded at 30fps and each meal duration was about 10-15mins. The second step involved image extraction from the recorded videos. A total of 1077 images were obtained from all 10 meal videos at 150 frames (five seconds) interval. Fig. 4-2 shows examples of selected meal images obtained from each of the captured videos. The images were manually labeled as food or non-food images by a human rater. The food images (class-1) contained the full amount or partial amount of food items. The images that contained no food items, surroundings only, faces of individuals, snapshots of personal gadgets (mobile/laptop) were labeled as non-food images (class-2).

4.2.2 Histogram feature and similarity matching

Color features of image typically constitute as one of the best descriptors by providing valuable sources of information. Among color descriptors, color histograms are the most common and widely used in various object recognition applications. In this work, the RGB images in each meal were first transformed into HSV color space images. The color feature for each image was represented using HSV color histogram. A total of 512 bins (8x8x8) with 8 levels for each color was used. Based on the histogram features of images, a similarity matrix was formed for the image collection. The histogram matching was done utilizing Bhattacharyya distance as,

$$d(h_a, h_b) = \sqrt{1 - \frac{1}{\sqrt{\bar{h}_a \bar{h}_b N^2}} \sum_{g \in G} \sqrt{h_a(g).h_b(g)}} \qquad (4.1)$$

Where h_a and h_b are two histograms, N and g are number of bins and pixel intensity values respectively. Low scores of d indicated good matches and high scores indicated bad matches. A value of 0 was perfect match while 1 indicated a perfect mismatch.

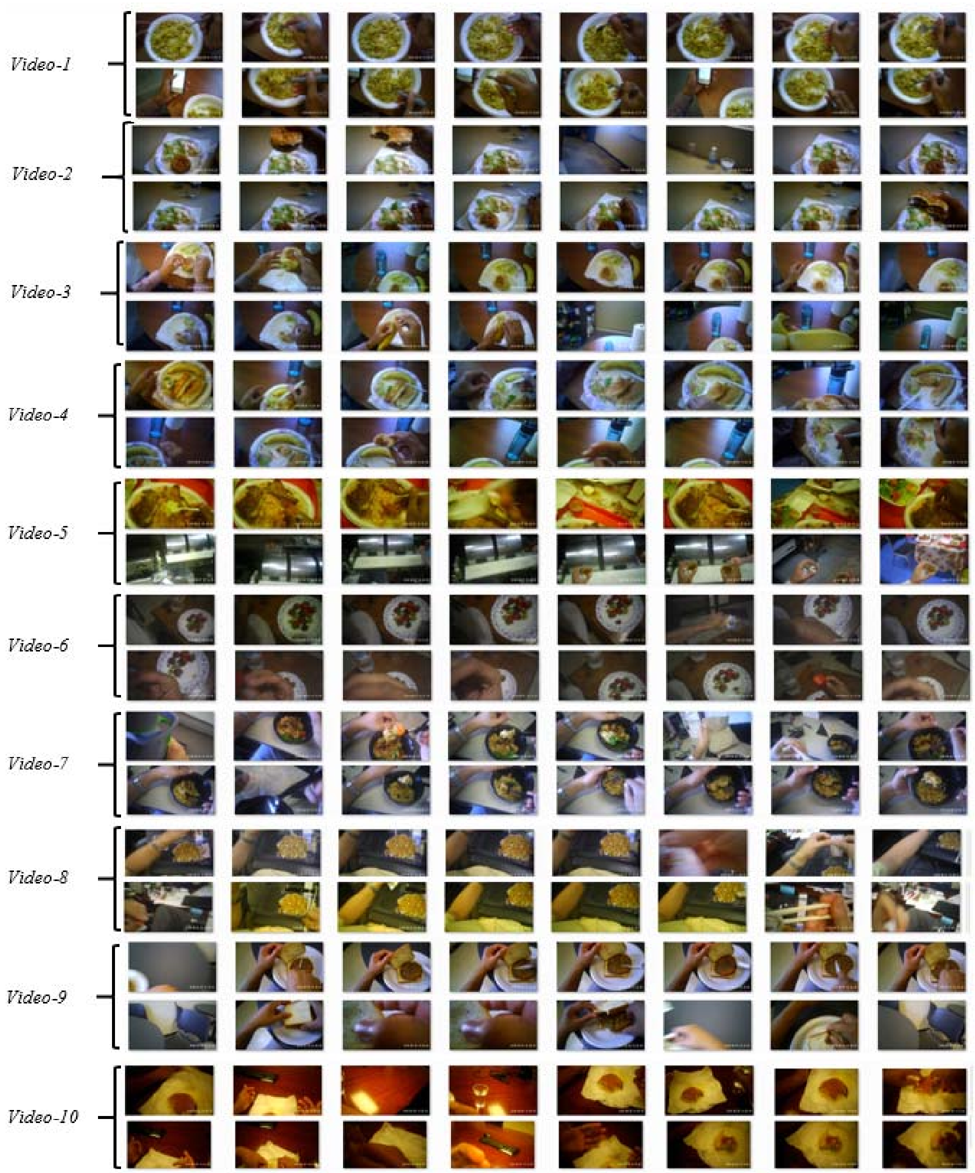

Figure 4-2. Example meal images from each of the videos.

4.2.3 Clustering Algorithms

Clustering technique classifies the objects into different groups, partitioning of a data set into clusters of common trait utilizing distance measurement. Two different unsupervised clustering techniques were investigated to cluster groups of food image/non-food images within a meal: K-means clustering and affinity propagation. K-means method is an unsupervised clustering method that classifies the input data objects into multiple classes on the basis of their inherent distance from each other. On the other hand, Affinity propagation (AP) clustering algorithm proposed in [13], takes as input a collection of real-valued similarities between data points and clusters by passing messages between data points. AP identifies exemplars among data points and forms clusters of data points around these exemplars. It operates by simultaneously considering all data point as potential exemplars and exchanging messages between data points until a good set of exemplars and clusters emerges matches. A value of 0 was perfect match while 1 indicated a perfect mismatch.

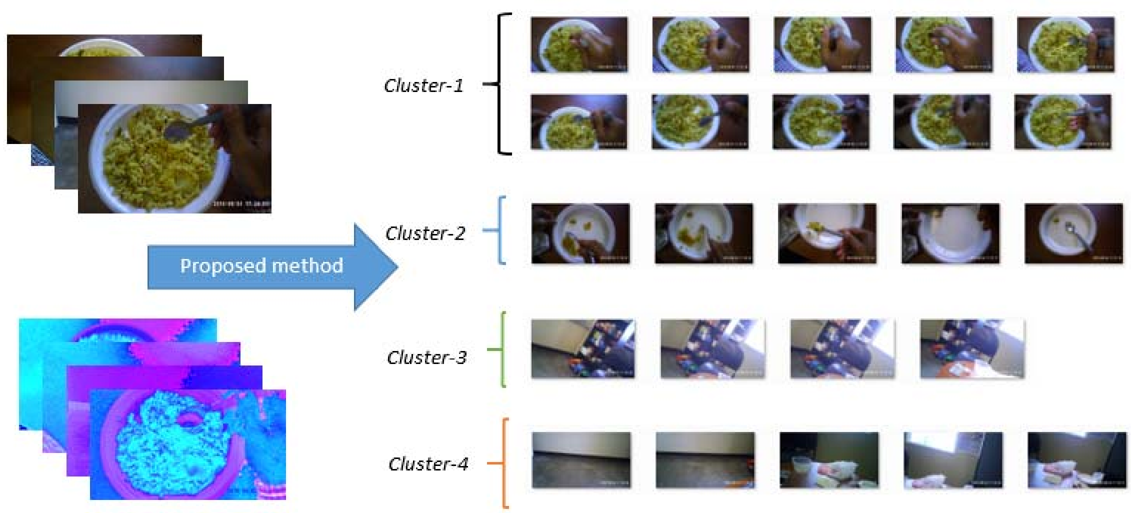

Figure 4-3. Clustering by the proposed method.

4.2.4 Performance evaluation

To provide an objective performance evaluation of the proposed approach, the goodness of image clustering was measured via purity value [14] and overall food image/non-food images clustering accuracy in respect to the human rater.

For a given set of n images labeled as l classes, the images are grouped into m clusters C_j, $j = 1, 2, \dots, m$, the purity for C_j is defined as,

$$P(C_j) = \frac{1}{|C_j|} \max_l |C_{j,l}| \tag{3.2}$$

where $C_{j,l}$ consists of images in C_j that belong to class l, and $|C_j|$ represents the size of the set. Each cluster may contain images of different classes. Purity gives the ratio of the dominant class size in the cluster to the cluster size itself. The larger value of purity means that the cluster is a "purer" subset of the dominant class.

The overall food image clustering performance was defined as the percentage of food images clustered together out of originally labeled total number of food images.

$$Acc\_FIC = \frac{\sum_1^x P(C_{j,1}).|C_{j,1}|}{\sum_1^j |C_{j,1}|} \tag{3.3}$$

Similarly, the overall non-food image clustering performance was defined as the percentage of non-food images clustered together out of an originally labeled total number of non-food images.

$$Acc\_NFIC = \frac{\sum_1^y P(C_{j,2}).|C_{j,2}|}{\sum_1^j |C_{j,2}|} \tag{3.4}$$

where x and y are the number of clusters recognized as pure food image clusters and non-food image clusters respectively.

4.3 Results

The experimental results from two clustering algorithms are provided in Table I in terms of purity value of all clusters and overall clustering performance compared to manually labeled images. Fig. 4-3 illustrates the performance of the proposed approach for one meal video. The RGB meal images are transformed to HSV before applying to clusters. Top few images for each cluster are shown. Fig. 4-4 demonstrates the comparative performance of both k-means and AP clustering method for all 10 videos. The overall average food image clustering accuracy for 10 videos were 0.93 ± 0.04 for AP and 0.90 ± 0.03 k-means. Similarly, the overall average non-food image clustering accuracy are 0.81 ± 0.06 for AP and 0.70 ± 0.09 k-means method. The mean purity of food image clusters for AP is 0.96 and for k-means is 0.93. Also, the mean purity of non-food image clustered for both AP and k-means are 0.91 and 0.87 respectively.

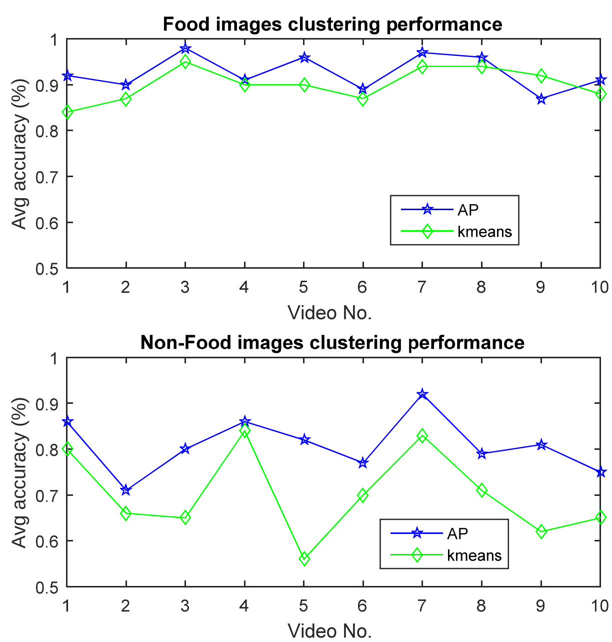

Figure 4-4. Performance of clustering.

4.4 Discussion

Dietary assessment through food imagery utilizes different image processing techniques. These techniques need to include complex food recognition algorithm which might not be readily available. On the other hand, a variety of food items may limit the accuracy of food recognition. Therefore, the method would require human intervention to estimate caloric intake. To assist human nutritionist, clustering meal images could become very useful to reduce the number of food images to be reviewed. The results from the proposed approach suggest that based on the similarity between the images, clustering can be done with significant accuracy, without attempting to recognize the food items in the images. Clustering, therefore, makes it possible to summarize a large collection of images and select just food images to analyze.

The proposed method used HSV color space rather than RGB. HSV images not only include color information but also keep some invariance in intensity and color changes. Since all of the meal videos were recorded in different lighting conditions, HSV image would be more suitable than RGB images.

The experimental results show that the AP clustering algorithm performs better than the k-means for clustering meal images. For a few cases, k-means performed comparatively better, however, AP exhibited consistency in clustering similar images together in either case of food images and non-food images. The performance of k-means in clustering of non-food images is significantly less compared to AP. In terms of non-food image clustering, Fig. 4-4 illustrates a consistent trend for AP but no consistent trend was found in k-means. A potential reason for getting good performance from AP clustering algorithm is that the AP considers all data points as a potential exemplar at initial iteration and then progress with a good match.

Table 4-1. Performance Evaluation

No	Total Images	Clustering Method	***P_{C1}* (class)**	***P_{C2}* (class)**	***P_{C3}* (class)**	***P_{C4}* (class)**	***P_{C5}* (class)**	***P_{C6}* (class)**	% of food images correctly clustered together	% of non -food images correctly clustered together
1	164	***AP***	0.97(2)	0.79(2)	1.00(1)	1.00(2)	0.98(1)	0.95(1)	0.92	0.86
		k-means	0.97(2)	1.00(1)	0.79(2)	0.47(2)	0.97(1)	0.96(1)	0.84	0.80
2	89	***AP***	0.93(1)	0.82(2)	0.90(1)	0.80(2)	1.00(1)	1.00(2)	0.90	0.71
		k-means	0.92(1)	0.74(2)	0.90(1)	1.00(1)	0.80(1)	1.00(2)	0.87	0.66
3	88	***AP***	1.00(1)	1.00(1)	1.00(1)	0.71(2)	1.00(1)	1.00(2)	0.98	0.80
		k-means	0.50(2)	1.00(1)	1.00(1)	0.71(2)	1.00(1)	1.00(1)	0.95	0.65
4	132	***AP***	0.79(2)	1.00(2)	1.00(2)	1.00(1)	0.88(1)	0.97(1)	0.91	0.86
		k-means	0.91(2)	1.00(2)	0.79(2)	1.00(1)	0.88(1)	0.97(1)	0.90	0.84
5	113	***AP***	1.00(1)	1.00(1)	0.97(1)	1.00(2)	1.00(2)	0.77(2)	0.96	0.82
		k-means	0.91(1)	1.00(1)	0.91(1)	0.83(2)	1.00(2)	0.84(2)	0.90	0.56
6	123	***AP***	0.93(1)	0.82(1)	1.00(1)	1.00(2)	1.00(1)	0.73(2)	0.89	0.77
		k-means	0.81(1)	0.82(1)	1.00(1)	0.73(2)	1.00(1)	1.00(2)	0.87	0.70
7	88	***AP***	0.97(1)	1.00(1)	1.00(2)	0.90(1)	1.00(1)	1.00(2)	0.97	0.92
		k-means	0.91(1)	1.00(1)	1.00(2)	0.90(1)	1.00(1)	1.00(2)	0.94	0.83
8	140	***AP***	1.00(1)	1.00(2)	1.00(1)	0.95(1)	1.00(1)	1.00(2)	0.96	0.79
		k-means	1.00(1)	1.00(2)	0.86(1)	0.95(1)	1.00(1)	1.00(2)	0.94	0.71
9	69	***AP***	0.57(2)	1.00(1)	1.00(1)	0.75(2)	1.00(2)	1.00(2)	0.87	0.81
		k-means	0.67(1)	1.00(1)	1.00(1)	0.75(2)	1.00(2)	1.00(2)	0.92	0.62
10	71	***AP***	0.89(1)	1.00(1)	0.93(1)	0.90(1)	0.82(1)	1.00(2)	0.91	0.75
		k-means	0.89(1)	0.85(1)	0.93(1)	0.90(1)	0.82(1)	1.00(2)	0.88	0.65
Summary		***AP***				***Average ± Std***			**0.93 ± 0.04**	**0.81 ± 0.06**
		k-means				***Average ± Std***			**0.90 ± 0.03**	**0.70 ± 0.09**

For the goodness of clustering analysis, purity value alone may not provide an overall estimate of the system performance even though they provide the quality of image clusters. Because a cluster could contain a small number of class images with high purity value. Therefore, one needs to consider both the purity value of image clusters and a total number of images present in the cluster. For this purpose, overall accuracy measures with respect to all class images were introduced.

It is also observed from both of the clustering algorithms that they create multiple food/non-food image clusters which are expected. During a full meal, the food items are being consumed over time, and therefore yields a different portion of foods on the plate constituting different images. Likewise, in the case of non-food images, the subject might take look into different objects and surroundings. Another possible reason could be the surrounding lighting conditions.

The proposed method offers potential use to cluster food and non-food images for a small number of participants. Therefore, further experiments are needed to test the performance of the method on a wider population under diverse lighting conditions.

From the standpoint of generating clusters within the food images, further work needs to be done to separate between full amount of food images and a partial amount of foods present in the image. Another potential future direction could be to analyze the time progression of meal from the clusters. As a future work, cluster centers could be identified to rank representative images in a meal.

4.5 Conclusion

In this paper, a clustering of food image and non-food images method is proposed without food recognition. A pairwise matching algorithm from the histogram of meal images was investigated. Two different clustering approaches were evaluated and compared. The affinity propagation algorithm outperformed the k-means clustering method in grouping meal images. The proposed method demonstrates the potential use of minimizing the number of images that need to be analyzed for the estimation of caloric intake. Future works could be done in clustering food images and find out the best few representative images.

4.6 References

[1] Fontana, J., Sazonov, E.: Detection and characterization of food intake by wearable sensors. In: Wearable Sensors: Fundamentals, Implementation and Applications. pp. 591–616. Academic Press (2014).

[2] Martin, C.K., Han, H., Coulon, S.M., Allen, H.R., Champagne, C.M., Anton, S.D.: A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. Br. J. Nutr. 101, 446–456 (2009).

[3] Noronha, J., Hysen, E., Zhang, H., Gajos, K.Z.: Platemate: Crowdsourcing Nutritional Analysis from Food Photographs. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. pp. 1–12. ACM, New York, NY, USA (2011).

[4] Shroff, G., Smailagic, A., Siewiorek, D.P.: Wearable Context-aware Food Recognition for Calorie Monitoring. In: Proceedings of the 2008 12th IEEE International Symposium on Wearable Computers. pp. 119–120. IEEE Computer Society, Washington, DC, USA (2008).

[5] Zhu, F., Bosch, M., Woo, I., Kim, S., Boushey, C.J., Ebert, D.S., Delp, E.J.: The Use of Mobile Devices in Aiding Dietary Assessment and Evaluation. IEEE J. Sel. Top. Signal Process. 4, 756–766 (2010).

[6] Kong, F., Tan, J.: DietCam: Automatic Dietary Assessment with Mobile Camera Phones. Pervasive Mob Comput. 8, 147–163 (2012).

[7] Puri, M., Zhu, Z., Yu, Q., Divakaran, A., Sawhney, H.: Recognition and volume estimation of food intake using a mobile device. In: 2009 Workshop on Applications of Computer Vision (WACV). pp. 1–8 (2009).

[8] Chen, M., Dhingra, K., Wu, W., Yang, L., Sukthankar, R., Yang, J.: PFID: Pittsburgh fast-food image dataset. In: 2009 16th IEEE International Conference on Image Processing (ICIP). pp. 289–292 (2009).

[9] Joutou, T., Yanai, K.: A food image recognition system with Multiple Kernel Learning. In: 2009 16th IEEE International Conference on Image Processing (ICIP). pp. 285–288 (2009).

[10] Anthimopoulos, M.M., Gianola, L., Scarnato, L., Diem, P., Mougiakakou, S.G.: A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model. IEEE J. Biomed. Health Inform. 18, 1261–1271 (2014).

[11] Ahmad, Z., Khanna, N., Kerr, D.A., Boushey, C.J., Delp, E.J.: A mobile phone user interface for image-based dietary assessment. Presented at the (2014).

[12] Liu, J., Johns, E., Atallah, L., Pettitt, C., Lo, B., Frost, G., Yang, G.-Z.: An Intelligent Food-Intake Monitoring System Using Wearable Sensors. In: 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN). pp. 154–160 (2012).

[13] Frey, B.J., Dueck, D.: Clustering by Passing Messages Between Data Points. Science. 315, 972–976 (2007).

[14] Chen, Y., Wang, J.Z., Krovetz, R.: Content-based Image Retrieval by Clustering. In: Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval. pp. 193–200. ACM, New York, NY, USA (2003).

CHAPTER 5
"AUTOMATIC INGESTION MONITOR VERSION 2" – A NOVEL WEARABLE DEVICE FOR AUTOMATIC FOOD INTAKE DETECTION, CHEW COUNT ESTIMATION AND PASSIVE CAPTURE OF FOOD IMAGES

The full validation part of this study is not complete yet and the manuscript will be submitted after the study.

Abstract

In the context of dietary assessment, monitoring of food intake utilizing food image capture and/or wearable sensors has received increased attention. "Active" capture methods use smartphones and rely on self-report by the user to take an image of each eating episode. "Passive" methods use wearable cameras that continuously capture images. Passive capture results in large datasets with most images not related to food consumption, and also presents privacy concerns. In this paper, we propose a novel wearable sensor system (Automatic Ingestion Monitor, AIM-2) designed to capture images only during detected eating episodes. The AIM-2 also contains a novel sensor to estimate chew counts. The device was validated on a dataset collected from 5 subjects wearing the AIM-2 for 24h in pseudo-free-living and 24h in a free-living environment. Privacy concerns were assessed by a questionnaire on a scale 1-7. The AIM-2 was able to detect food intake over 3-second epochs with a (mean and standard deviation) F1-score of 90.5 ± 3.16%, and estimate chew counts with an accuracy of 90.6% ± 3.56%. The accuracy of eating episode detection was 87%, suggesting that AIM-2 captured most of the ingested meals. Out of a total of 17,252 images captured in the study,

1389 images belonged to detected eating episodes (198±38 images per eating episode). Image capture only during food intake reduced privacy concerns from 6.6±0.5 (extremely concerned, continuous capture) to 1.8±0.7 (not concerned, food only capture). Results suggest that the proposed system can provide accurate detection of food intake and estimation of chew counts, reduce the number of images for analysis and potentially alleviate privacy concerns of the users.

5.1 Introduction

The emergence of chronic diseases such as obesity, overweight, and eating disorders has prompted extensive research efforts to understand and study the contributing factors towards the development of these diseases [1], [2]. One potential possibility to realize the causes and treatments of the disease is to accurately monitor the food intake of the individual. Traditional methods of monitoring individual food intake through self-report such as food records [3], food frequency questionnaires [4], 24-hours food recall [5], have proven to be highly inaccurate due to participant's inherent misreport [6]. Therefore, it is challenging to use inaccurate and incomplete data from self-reports to objectively monitor individuals' food intake and dynamic process of food consumption known as meal microstructure [7].

Accurate and objective assessment of food intake must provide a proper characterization of food intake, recognition of consumed food items, estimation portion size, and energy content [8]. Over the past decade, to tackle the problems associated with self-report, various methods of automatic detection and monitoring of food intake have been proposed, ranging from image-based methods to wearable on-body sensors. In recent years, image-based methods have been investigated for dietary assessment utilizing either smartphones or wearable cameras. These methods rely on the captured image as the primary source of information. The image capturing can be categorized into two methods – active and passive.

Active methods assume use of the hand-held camera (such as smartphone camera) to capture the images of food by the user. Typically, these methods acquire images of food items before, during and after a meal. In active methods, the analysis of acquired images is typically performed in two ways: manual annotation of the images and automatic image recognition and processing. During manual annotation, food images are analyzed by an expert nutritionist to identify different types of foods, their portion size, and the energy content. The study of [9] proposed an approach called the Remote Food Photography Method where the participants submit food images to a nutritionist for the analysis. Recently, an application called PlateMate [10] analyzed food images by utilizing a crowdsourcing technique. On the other hand, existing active methods that rely on automatic image recognition and processing to segment food images, recognize foods, estimate portion size/volume and compute energy content. Several methods have been introduced to identify foods in the image via supervised and unsupervised machine learning techniques. The authors of [11] proposed a method consisting of the segmentation of food items in the image, food recognition, automatic portion estimation, and extraction of energy information that was consumed during a meal. In [12] the authors proposed a camera phone based automatic food intake monitoring system (DietCam). Apart from color and texture features from images, the author utilized statistical features considering the noise produced by camera sensors. Then a Bayesian probabilistic approach was adopted to classify food items. The authors in [13] proposed a system where the estimation of the volume of food items was performed on captured images of a smartphone. A pair of images were captured from the left and right side of the food item placed on the plate. The method included three major steps towards food image recognition: disparity & depth estimation, 3D reconstruction, and volume estimation. The authors in [14] proposed a system called 'Im2Calories' that recognized the contents of food items consumed in the meal and predicted energy content. The system proposed to recognize food items based GoogLeNet convolutional neural network (CNN) model. The method did not detect the food intake instances rather assumed the user consumed their meals at a restaurant. Major advantages of active methods include reduction of reporting burden compared to traditional self-report and inclusion of food images (initial state and leftovers) that help nutrition analysis of the eating episodes. Apart from these advantages, they also provide the detailed information about the timing, location, and duration of eating episodes. While providing advantages, the active methods also involve some limitations. The image capture of all eating episodes requires active participation from the participants. Some of the methods are required to place fiducial markers (known as dimensional and color references) in the food image to assist manual review/ computer [12], [13]. Furthermore, to enhance the image analysis, some methods require the images to be captured at a certain angle between 45° and 60° [13].

The image can also be acquired by a "passive" method using wearable devices that would capture food and non-food images continuously and automatically without the active participation of the user. The passive methods minimize the burden of active capture, through the use of a wearable camera. Due to capturing both food and non-food images, passive methods form large image data set and introduce privacy concerns. The large image data sets make the manual image review difficult. To lessen the burden of image review, deep-learning based image processing are investigated to automatically differentiate between food and non-food images. The author in [15] proposed a chest-worn electronic device; 'eButton' that automatically captured images of consumed food in every two seconds interval. The volume of food items was then recognized by segmenting items based on color, texture and a complexity measure. The same research group utilized the ‘eButton’ device to capture images of real-world activities including both food and non-food images [16]. The study acquired around 30,000 images and manually reviewed all images to label as food and non-food images. Later, they proposed and validated a food detection algorithm using CNNs. In [17], the authors investigated the possibility to improve the accuracy of 24-hours dietary recall using a wearable chest-worn camera (Microsoft SenseCam). After completing the 24-hours recall, the participants reviewed the images and suggested changes to self-reported intakes. The method was then validated against the doubly labelled water (DLW) technique. Despite advantages, the passive devices, also carries limitations. The passive capture from the wearable camera may result in large image data sets for analysis. The manual review of such image sets is very time-consuming. Another limitation is that the automatic image capturing device may take unwanted images that can raise privacy concerns and potential ethical issues. The wearer inadvertently might collect inappropriate images, such as adjusting clothing in front of a mirror, using the bathroom, browsing internet/social networking sites, reading confidential documents, unwanted images of surroundings – getting undressed/ faces of family members without consent [18]. Therefore, methods relying on wearable cameras should deal with these situations. To minimize the number of images for review and reduce privacy concerns, off-line food intake detection can potentially be done and only food images can be reviewed.

Several studies proposed food intake monitoring systems that utilize wearable sensors to detect food intake events. In a pilot study [19], the authors proposed a wearable sensor system consisting of a microphone and a miniature camera. Audio features were extracted in real-time to detect the chewing activity. The detection of chewing activity further triggered the camera to capture a video sequence. From the video sequence, a number of images were extracted for analysis. In the proposed approach, there exists a possibility that the false chewing activity from surrounding noise can trigger the camera. The continuous video recording may decrease the battery life immensely. The same research group utilized the proposed device in another pilot study [20] where the participants were instructed to turn the camera on to begin recording the food intake episodes. The study [21] investigated a smartwatch camera to take images of the food consumption. The authors explored the temporal variation of the accelerometer and gyroscope data to trigger the camera. The findings suggested that hand-to-mouth gestural trajectory might be used to trigger smartwatch camera.

Various wearable sensors have been proposed to detect different stages of eating, i.e. bites, chewing and swallowing of food. Several hand/wrist-worn wearable device have been introduced related to bite and gesture detection including accelerometers, gyroscopes and smart watches [22]–[24]. The act of chewing detection has been done via chewing sounds [25]–[29], muscle movement during chewing using strain sensors [30]–[33], or EMG and force sensors [34]–[36]. Researchers also proposed sensors to monitor swallowing using microphones in the ear or surface electromyography on throat [37]–[39]. Our research group previously proposed a piezoelectric strain gauge based chewing sensor system "Automatic Ingestion Monitor (AIM)" for monitoring food intake in the free-living environment [40]. The sensor system has been validated in several conditions and subject groups (both adults and infants) [7], [38], [41]–[43]. We have also demonstrated that the chew counts are highly correlated to the ingested mass and type of food [44]–[47] and thus sensor-measured chew counts may be used a measure of portion size. The chewing sensor in these studies was applied to the skin using medical adhesive tape, which inconvenienced the participants wearing the sensor.

In this paper, we propose a novel wearable sensor system (Automatic Ingestion Monitor version 2, AIM-2) that can automatically detect food intake, estimate chew counts and passively capture images of food intake with a non-adhesive sensor system. The major contributions of the proposed system are: (1) AIM-2 uses non-adhesive sensors for accurate detection of food intake, with validation performed both in laboratory and free-living environments; (2) AIM-2 enables accurate estimation of chew counts by a non-adhesive sensor, potentially assisting in portion size estimation; (3) AIM-2 provides a potential solution to significantly reduce the number of images captured by this “passive” device; and (4) This study demonstrates that capture of images only during detected food intake significantly reduces privacy concerns of the users.

5.2 Material and Methods

5.2.1 Sensor System

In this study, a novel wearable sensor system (Automatic Ingestion Monitor, AIM-2) was used. AIM-2 consists of a sensor module which housed a miniature 5 Megapixel camera with 120- degree wide-angle gaze-aligned lens, a low-power 3D accelerometer (ADXL362 from Analog Devices, Norwood, MA, USA) and a novel bending sensor (SpectraSymbol 2.2" flex sensor). The sensor module was connected to the frame of eyeglasses by a heat-shrink tube (Fig. 5-1) in a location such that the maximum curvature of the sensor touched the skin over the temporalis muscle where the strongest muscle contraction during chewing was palpated. The camera continuously captured images at a rate of one image per 15-second interval. In this study, the image capture was not sensor-driven so that the sensor-based food intake detection could be validated against image-based food intake detection. The flex sensor was used for estimation of chew counts. To protect the sensor from perspiration, commercially available polyurethane coating was used over the top layer of the flex sensor. Data from the accelerometer and flex sensor were sampled at 128 Hz. All collected sensor signals and captured images were stored on an SD card and processed off-line in MATLAB (Mathworks Inc., Natick, MA, USA) for algorithm development and validation.

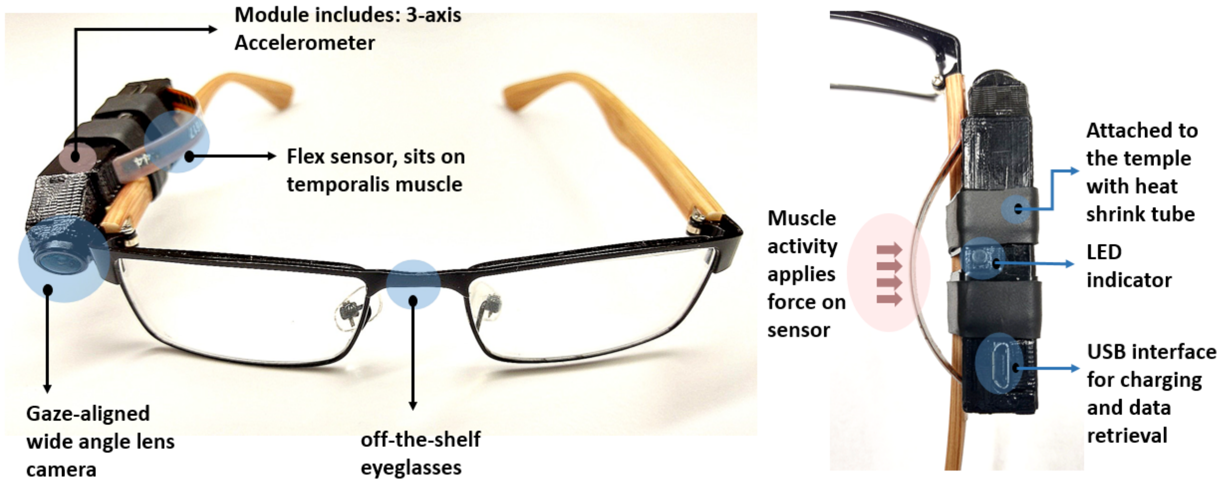

Figure 5-1. AIM-2. The sensor module is attached to the temple of off-the-shelf wearable eyeglasses with heat shrink tubes.

5.2.2 Data Collection Protocol

Five volunteers were recruited for this study (4 male and 1 female, mean± SD age of 29.4 ± 3.13 years, range 26-34 years, mean body mass index (BMI) 23.65 ± 2.49 kg/m^2, range 20.8 to 25.5 kg/m^2). The University of Alabama's Institutional Review Board approved the study. Participants did not have any medical conditions that would impact their chewing or eating. Participants came to the laboratory for four visits over x days. The laboratory was instrumented with two high-definition cameras (Contour Roam2 LLC, USA, and GW-

2061IP, GW Security, Inc. CA, USA) to record the experiment. Prior to the experiment, the participants made their first visits to the laboratory. They went through the screening process and were given a summary of the experimental protocol. Each participant signed an informed consent form. After consent was acquired, the sensor system (AIM-2) was affixed to eyeglasses of the participants (if the participants had the corrective lens, they were instructed to bring their spare eyeglasses) or non-prescription eyeglasses (if the participants had no corrective lens, they chose from the eyeglasses available at the lab). The eyeglasses mounted with a sensor system for each participant were kept in the lab until the next visit. During the visit, the participants were also trained to report dietary intake and activity using the mobile applications Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool [48] and aTimeLogger [49], respectively. On the first day of the experimental protocol, the participants came to the laboratory three times (visits 2-4). Several participants were invited to the laboratory at the same time to simulate social eating. For the second visit, the participants were instructed to arrive at the laboratory in the morning (between 7:00 AM - 9:00 AM) after an overnight fast. Upon the arrival, participants were given the AIM-2 device and were reminded about the instructions for the experiment. Each participant consumed a full meal (breakfast) purchased from the University food court. The participants chose the type and quantity of food items that they wanted to consume. Research assistants weighed each food item and used a nutrition software "Food Processor" [50] to enter the weight and other details about the food items. After selection of foods for the meal, participants were required to consume the meal at the laboratory. There was no restriction on the time required to finish the meal.

During the food consumption in the lab, participants used the foot pedal connected to a USB data logger to mark ingestion of solid and liquid food items. They were asked to press the pedal the moment the food was placed in the mouth (a bite), and hold the pedal until the food was fully consumed. For beverages, they were asked to press and hold the button from the moment they brought the liquid to mouth until the last swallow. When the participants had finished eating, the research assistant weighed leftovers and updated the amount consumed in the nutrition software. To include a representative set of daily activities, participants performed a few activities in addition to eating (e.g. sitting, sitting and talking, walking on the treadmill, walking and talking and shelving). Then, the participants left the laboratory and continued carrying out their usual daily activities in the free-living environment. Visits 3 and 4 were made on the same day to consume lunch and dinner following a similar protocol. Since the eating episodes were consumed at the laboratory and other daily activities were not restricted, the first day of study was considered to be conducted in the pseudo-free-living environment.

After the completion of day-1, the participants continued to free-living for 24 hours. During this period, the participants had virtually no restrictions on the types of the activities performed. They continued their normal daily routine with three exceptions: to take the AIM-2 off during 1) taking the shower 2) any water activities that may damage the device, and 3) sleeping. The participants were also asked to self-report all of their laboratory and eating episodes using the ASA24 mobile app (used a food diary). They were also asked to self-report major activities from the set of sleeping, eating, sedentary, and physically active through the aTimeLogger application. After completing the free-living part of the study, the participants reported to the laboratory to return the AIM-2.

Table 5-1. Experimental protocol

Day	Activity	Description	Duration
Prior to experiment	**First visit**		
		Initial interview, the signing of the informed consent, completing a food frequency questionnaire, AIM-2 sensor preparation	20 min
		Training on mobile apps	15 min
		Training on portion size estimation	20 min
		Testing on portion size estimation	15 min
1	**2nd visit to the lab**		
	1	Sit silent: sit in a comfortable position	5 min
	2	Sit while talking: read a document aloud	5 min
	3	Eat a meal: eat a meal (eating episode -1)	20 - 40 min
	4	Walking while silent: walk on a treadmill at a self-selected comfortable speed	5 min
	5	Walking while talking: walk on a treadmill at a self-selected comfortable speed and talk with the research assistant	5 min
	6	Sit silent: sit in a comfortable position	5 min
	7	Activities of daily living: shelving /stacking items	5 min
1	**3rd and 4th visits to the lab**	Carry out usual daily activities, with the exception of 2 eating episodes to be consumed at the laboratory	1 day
2	**Free-living**	Carry out usual daily activities and eating the usual diet.	1 day

5.2.3 Sensor Signal Processing and Annotation

Before extracting features from sensor signals, the raw sensor data from the accelerometer and flex sensor were pre-processed. A high-pass filter with a cutoff frequency of 0.1 Hz was applied to remove the DC component from the accelerometer signal. To adjust inter-subject variations, the signals were then normalized [40]. The flex sensor signal was demeaned. As the typical chewing frequency is in the range of 0.94 to 2 Hz, therefore, a low-pass filter with a cutoff of 3Hz was used to remove unwanted noise from the flex sensor [51]. Next, the signals were divided into non-overlapping 3s fixed time segments called 'epochs'. In our previous study [52], we have shown that 3s epoch duration is sufficient to accurately preserve information of meal microstructure [7]. Fig. 5-2 shows the accelerometer signals, flex sensor signals and corresponding foot pedal signals during an eating episode. Fig. 5-3 illustrates a zoomed version of two chewing sequences of an eating episode captured by the flex sensor.

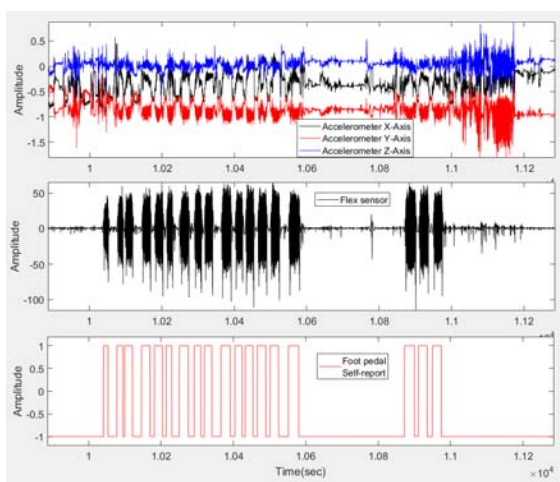

Figure 5-2. Signals from 3-axes of the accelerometer are shown on the top segment. The middle segment demonstrates the flex sensor signal. The foot pedal signal marked by the participants is shown on the bottom segment. Here '1' indicates food intake and '-1' indicates no food intake.

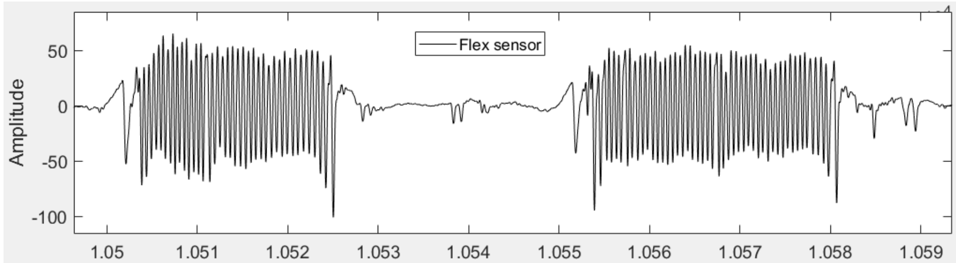

Figure 5-3. Two chewing sequences captured by the flex sensor.

To obtain chew counts, three trained human raters annotated the video data. Fig.5-4 illustrates a screenshot of the recorded video of a participant while eating. The raters identified each chewing sequences in the eating episode and counted the number of chews using a tally counter. The rater paused the video recording after each chewing sequence, recorded the chew counts in an Excel sheet and resumed the recording to annotate the next chewing sequence. The rater continued the same process until the last chewing sequence of the eating episode. To check the reliability of the raters, a set of three meals were given to each of raters to annotate.

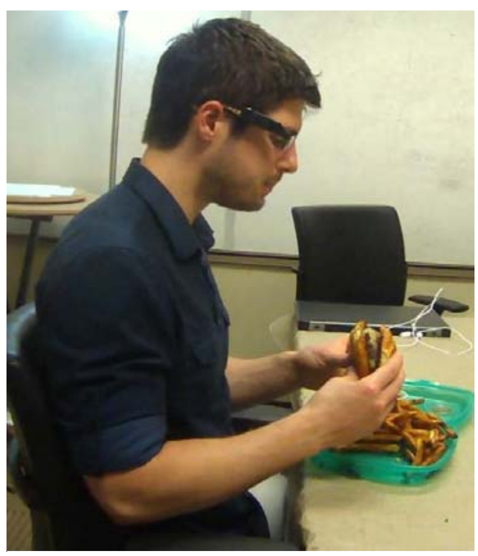

Figure 5-4. A screenshot of the recorded video of a participant while eating.

The reliability of the annotation procedure was evaluated by the intra-class correlation coefficient (ICC). For a sample size of 3 meals, three raters achieved 0.96 ICC for total chew counts. The resultant chew counts data was utilized for the validation of the chew count algorithm.

5.2.4 Feature Extraction

A set of 38 time and frequency domain features reported in our previous studies [31], [53] was utilized in this work. The accelerometer sensor signal contains chewing information in the frequency range 1.25-2.5 Hz, physical activity in the 2.5-10 Hz range and speech in the range 10-30 Hz [53]. For each of 3 axes of the accelerometer and net acceleration, 38 features were computed on 3s epoch data. For flex sensor signal, the same 38 features were computed, plus five additional correlation-based features. The entire feature list is listed Table 5-2. Aggregating all features, 195 features were computed for each epoch.

5.2.5 Feature Selection and Classification

To reduce the computational burden and redundancy in the computed features, a two-stage feature selection procedure was used. First, a ranking of computed features was carried out. To rank the feature based on mutual information, an algorithm called minimum Redundancy and Maximum Relevance (mRMR) was applied on the computed features [54]. Secondly, a Forward Feature Selection (FFS) was applied to the top-ranked 40 features selected by mRMR to get an optimal feature set. To avoid introducing bias into classification models that might result in overfitting, feature selection was done on an independent dataset (collected from two pilot subjects, who are not a part of the population for this study).

Foot pedal signals were used as a source of labels for the development of signal processing and pattern recognition algorithms. If more than 50% of an epoch belonged to food intake, the epoch i was assigned a label Ep_i = '1' (food intake), otherwise the

Table 5-2. Feature sets computed from both accelerometer sensor and flex sensor epochs

Features	Signal	Total
38 time and frequency domain features : Mean absolute value (MAV), root mean square (RMS), maximum value (Max), median value (Med), ratio of MAV to RMS, ratio of Max to RMS, ratio of MAV to Max, ratio of Med to RMS, signal entropy, number of zero crossings (ZC), mean time between ZC, number of peaks (NP), average range, mean time between peaks, ratio of NP to ZC, ratio of ZC to NP, wavelength, number of slope sign changes, frequency spectrum energy (spec_energy), energy spectrum in chewing range (chew_energy), entropy of spectrum chewing range (chew_entr), ratio of chew_energy to spec_energy, energy spectrum in walking range (walk_energy), entropy of spectrum walking range (walk_entr), ratio of walk_energy to spec_energy, energy spectrum in talking range (talk_energy), entropy of spectrum talking range (talk_entr), ratio of talk_energy to spec_energy, ratio of chew_energy to walk_energy, ratio of chew_entr to walk_entr, ratio of chew_energy to talk_energy, ratio of chew_entr to talk_entr, ratio of walk_energy to talk_energy, ratio of walk_entr to talk_entr, fractal dimension, peak frequency in chewing range, peak frequency in walking range, peak frequency in talking range.	Acc x-axis Acc y-axis Acc z-axis Net-acceleration Flex sensor	38*5 = 190
5 correlation related features : correlation coefficient, the 1st autocorrelation function coefficient, fundamental frequency, pitch period, value of autocorrelation at zero lag	Flex sensor	5
	Total	195

label Ep_i = '-1' (no food intake). A group classification model based on Support Vector Machine (SVM) was trained to identify food-intake and non-food intake epochs. Linear SVM is a supervised learning technique that has demonstrated excellent results for classification problem. The primary advantages of SVM are good generalization and speed [55]. Training of the model was performed using Classification Learner tool in MATLAB 2017 (from Mathworks Inc.). The performance of the linear SVM model was evaluated in two forms: 1)

Per-epoch classification accuracy of laboratory meals and 2) Per-epoch classification accuracy of free-living meals.

To evaluate per-epoch classification accuracy of laboratory meals, a leave-one-out cross-validation procedure was used. This allowed training the SVM model with data from 4 participants and testing the trained SVM model with data from the remaining participant. The procedure was repeated 5 times such that each participant was used for testing once. The accuracy for each iteration was assessed as F1 score:

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively.

To evaluate per-epoch classification accuracy of free-living meals, the trained linear SVM classification model was used. Unlike the food pedal reference (food intake/no-food intake epochs) in laboratory meals, the reference for free-living meals was not readily available. For the sake of accuracy evaluation in free-living, the boundaries of self-reported eating episodes from aTimeLogger mobile app were used. The same metrics (i.e. precision, recall and F1-score) were computed to report the performance.

5.2.6 Image Capture and Manual Review

In this validation study, the AIM-2 captured images every 15 seconds. Due to continuous image capture, the AIM-2 might capture unwanted images and therefore the participants were given the opportunity to review and delete them before the research assistant viewed them.

For validation purposes, the complete set of captured images was reviewed. The first goal of the review was to identify any missing eating episode by food intake detection algorithm. The image review provided the total number of image-detected eating episodes. Processing of the sensor data provided the total number of sensor-detected eating episodes. Following our previous research [7], a smoothing kernel with the shape of a Gaussian curve was applied on the food intake epochs. Based on the Gaussian-smoothed signals, the number of sensor-detected eating episodes was determined. Then, the accuracy of eating episode detection was computed as the ratio of the number of sensor-detected eating episodes to the number of image-detected eating episodes.

The second goal was to eliminate any false positives in food intake detection. All sensor-detected eating episodes and corresponding images were reviewed. If no images of food item were detected in a sensor-detected eating episode, then it was counted as false positive.

The third goal was to quantify the reduction in the number of images captured by the "passive" device. For each participant-day, the total number of images, the total number of food images in sensor-detected eating episodes and the percentage of food images of all captured images were computed.

5.2.7 Chew counts estimation

To obtain the chew counts from sensor-detected eating episodes, the novel flex sensor was utilized. Both eating events and non-eating movements of the jaw cause the temporal muscle to contract which results in variations (peaks and valleys) in the flex sensor signal. Using the peak detection algorithm, the presence of these peaks and valleys can be detected. The number of peaks and valleys can be used for estimation of the number of chews during eating events. To avoid peaks caused by movements other than chewing, a threshold was set to determine whether the peaks and valleys belong to chewing event. For threshold detection, a histogram-based approach, previously proposed by our research group in [42] was adopted. Fig. 5-5 illustrates a flow diagram of chew counting algorithms.

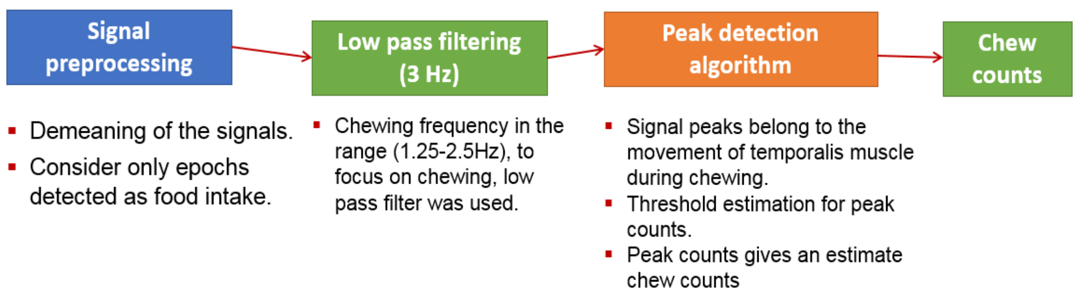

Figure 5-5. Flow diagram of chew counting algorithm.

The chew counts algorithm was applied only on the epochs which were labeled as food intake by the classifier. For performance evaluation of the chew counting algorithm, the estimated chew counts were compared to the video annotated chew counts and error was computed for each participant (total chews in three laboratory meals). Both mean errors ($Error\_CC$) and mean absolute errors ($AbsError\_CC$) were reported as:

$$Error\_CC = \left[\frac{(Act_{CC} - Est_{CC}) * 100}{Act_{CC}}\right], \tag{2}$$

$$AbsError\_CC = \left|\left[\frac{(Act_{CC} - Est_{CC}) * 100}{Act_{CC}}\right]\right| \tag{3}$$

where $\boldsymbol{Act_{CC}}$ is the total number of chew counts in three laboratory meals annotated from video and $\boldsymbol{Est_{CC}}$ is the total number of chew counts in three laboratory meals estimated from the chew count algorithm.

5.2.8 Privacy Concerns Survey

Since the participants did not have any control over the passive image capture from AIM-2, there exist possible privacy concerns. To assess the privacy concerns of participants, a survey [56] was conducted after the experiment. The participants were presented with the following questions and asked to answer using a scale of 1 to 7, where ‘1-2’ means ‘not concerned’; ‘3-5’ means ‘somewhat concerned’; ‘6’ means ‘concerned’; ‘7’ means ‘extremely concerned’:

- How concerned are you about your privacy if the device captures images continuously throughout the day?
- How concerned are you about your privacy if the device captures images only during eating events?

The responses were recorded. To analyze the responses, the mean and standard deviation of the recorded score for both cases were computed. In addition, the percentage of participants in each of four privacy concern categories was computed. All analyses were done using Microsoft Excel 2016 (from Microsoft Inc.).

5.2.9 Statistical Analysis

To compare the total number of chews estimated by the system with the manually annotated chew counts, Bland-Altman plot and limit of the agreement were examined. All statistical analyses were done using MATLAB 2017 (from Mathworks Inc.).

5.3 Results

After feature selection, nine features out of all features were selected for training the classifier. Results of the leave one- out cross-validation procedure are shown in Tables 5-3 and 5-4 (laboratory environment and free-living environment, respectively). The F1-score metrics of food intake detection for 3 sec epochs were 0.91 ± 0.03 for lab meals and 0.86 ± 0.08 for free-living.

The AIM-2 chew courting algorithm was able to achieve average chew estimation accuracy of 90.59 ± 3.56%. Table 5-5 lists total actual chew counts from video and algorithm estimated chew counts. Fig. 5-6 shows the Bland-Altman plots between the mean of actual chew counts from video and estimated chew counts.

Table 5-3 Precision, Recall (Sensitivity), and F1-Score for 3s epochs for Laboratory Experiments

Participant #	Precision (%)	Recall (%)	F1-score (%)
1	0.99	0.89	0.94
2	0.99	0.86	0.92
3	0.91	0.89	0.90
4	0.88	0.88	0.88
5	0.97	0.81	0.86
Mean ± Std	**0.95 ± 0.04**	**0.87 ± 0.03**	**0.91 ± 0.03**

Table 5-4 Precision, Recall (Sensitivity), and F1-Score for 3s epochs for Free-Living Experiments

Participant #	Precision (%)	Recall (%)	F1-score (%)
1	0.9278	0.9530	0.9402
2	0.7618	0.8043	0.7825
Mean ± Std	**0.8448 ± 0.0830**	**0.8787 ± 0.0789**	**0.8614 ± 0.0789**

Table 5-5 Actual chew counts from video and algorithm estimated chew counts.

Participant #	**Actual chew counts from video**	**Estimated chew counts**	***Error_CC (%)***	***AbsError_CC (%)***
1	990	934	5.66 %	5.66 %
2	1870	1641	12.25 %	12.25 %
3	1914	1731	3.34 %	3.34 %
4	2167	1956	9.56 %	9.56 %
5	640	577	9.74 %	9.74 %
		Mean:	9.41%	9.41%
		STD:	3.56%	3.56%

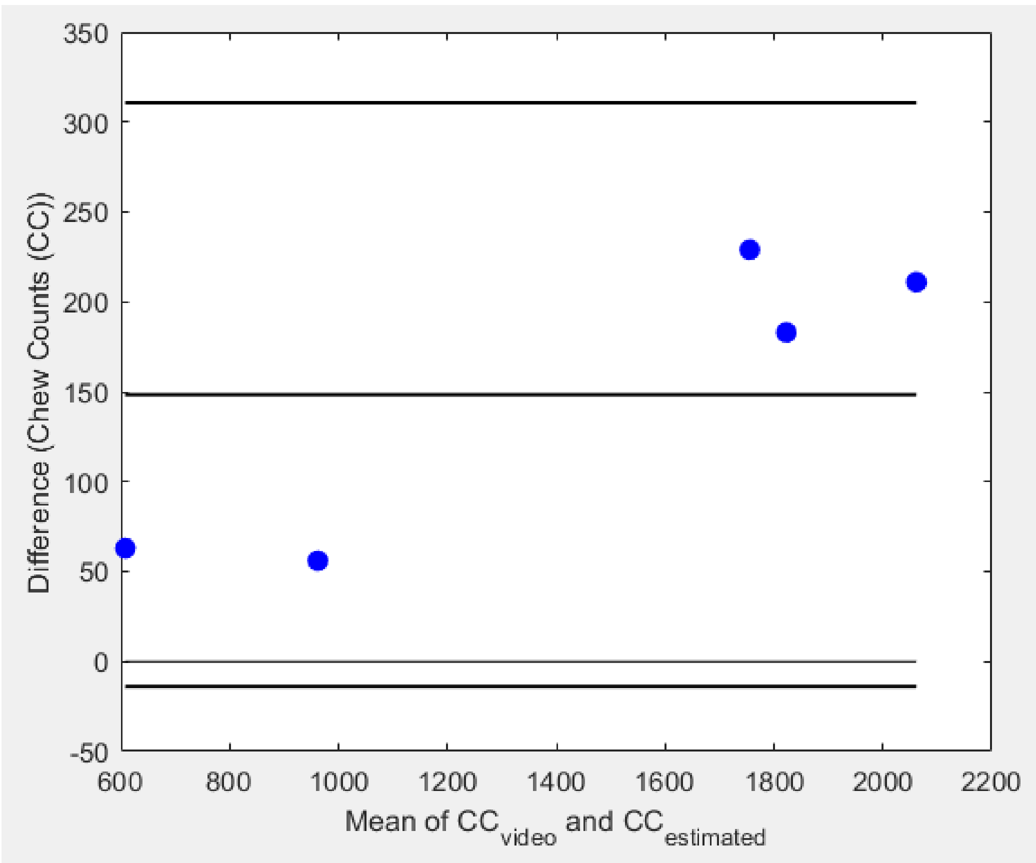

Figure 5-6. Bland-Altman plots between the mean of actual chew counts from video and estimated chew counts.

The mean accuracy of sensor-detected eating episodes was 87%, suggesting that the AIM-2 captured most of the ingested eating episodes. Manual review of all captured images revealed that three drinking episodes out of 23 sensor-detected episodes were missed.

The manual review of images detected three false positive in sensor-detected eating episodes.

While the AIM-2 was worn, 17,252 images were captured, with 1389 images of food and beverage consumption during sensor-detected events. Fig. 5-7 shows selected food and non-food images captured by AIM-2.

Table 5-6 demonstrates image statistics captured by AIM-2. For each participant day, on average 8.05% of all images were food images. This result indicates that the number of images for analysis can be significantly reduced. The amount of food images also bolsters the fact that the system may provide a potential solution to privacy concerns.

All participants responded to the privacy concerns questionnaire. Results show that the image capture only during food intake reduced privacy concerns from 6.6±0.5 (very concerned, continuous capture) to 1.8±0.7 (not concerned, food only capture). Fig. 5-8 demonstrates the distribution of scores of privacy concerns. The results indicated that the majority of participants were extremely concerned about privacy if the device captures the image continuously. Moreover, the majority of participants were not concerned about privacy if the device captures images only during food intake.

Figure 5-7. Examples of AIM-2 images, (top) selected non-food images (bottom) selected food images from several image-detected eating episodes.

Table 5-6 Number of images captured by AIM-2

Participant day #	**Total hours data collected (hours)**	**Total number of Images**	**Food Images**	**% of food images**
1	7	2640	194	7.35 %
2	6	2186	175	8.01 %
3	7	2594	212	8.17 %
4	6	2293	180	7.85 %
5	6	2285	262	11.47 %
6	7	2597	145	5.58 %
7	7	2657	221	8.32 %
Total	46	17252	1389	8.05 %

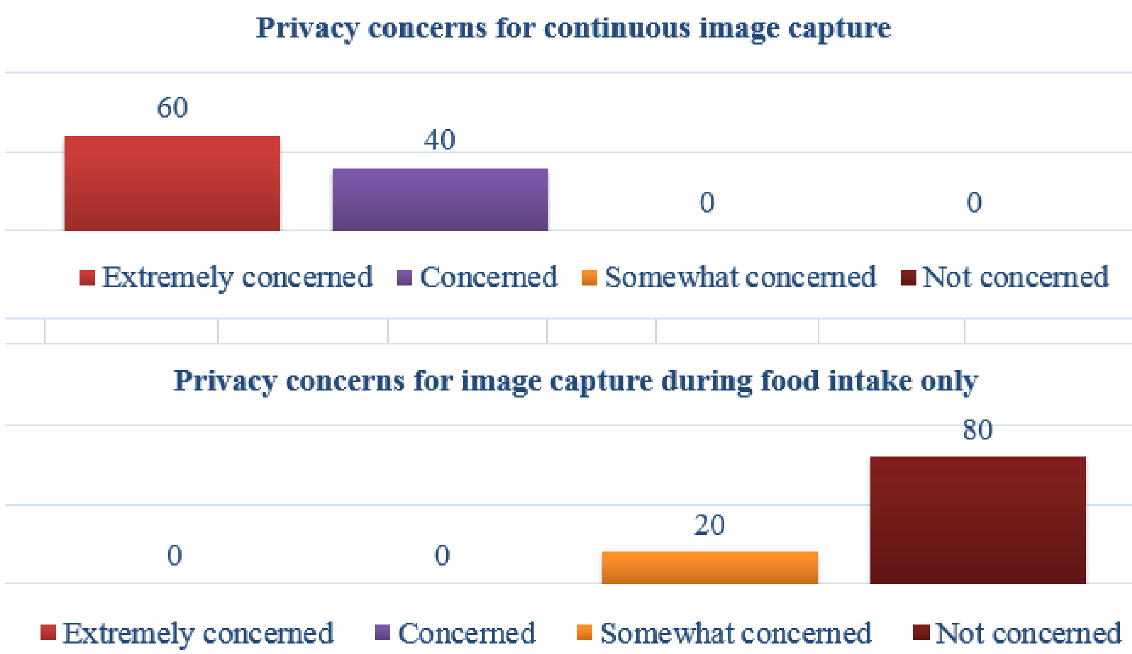

Figure 5-8. Level of privacy concerns for image capture.

5.4 Discussion

With the advancement of technology, research has demonstrated increased attention in the development of image-based food intake monitoring devices. The main objective of the present work was to develop a wearable sensor system with a non-adhesive sensor that automatically detects food intake, estimates chew counts using a new chewing sensor, and passively captures food images. Another goal was to develop a method that can help reduce the number of images for analysis and alleviate some of the privacy concerns.

This work demonstrated the detection of food intake in both laboratory and free-living settings. The validation of food intake detection only on laboratory setting is not the end goal as the laboratory setting often lacks realism due to unnatural eating behavior [57] or potential behavior change, limited food items, mixed meals, eating in groups, etc. To test the robustness of detection methods, it is imperative to use real-life conditions where the system will encounter a variety of daily activities that might affect the eating behavior.

The AIM-2 detected 90% and 86% of food intake epochs (3sec duration) in the laboratory and free-living environment respectively. In the free-living part of the study, the participants reported the start and end times of each eating episode using the aTimeLogger activity log application. The accuracy of the free-living environment was evaluated with the boundaries (start and end time) of self-reported eating episodes. Unlike the eating episodes consumed at the laboratory, the participants did not report each chewing sequence in free-living. The fine detail of food intake events was not available to compare the detected food intake labels. Since the self-report labeled all epochs including eating and non-eating (e.g. pauses) within the eating episode boundary as food intake, the accuracy report might have been affected. Regarding the sensor-detected eating episodes, the system recognized 20 out of 23 image-detected eating episodes, missing three drinking episodes that were detected during all image review. As the system primarily relies on chewing detection, few purely drinking episodes were not detected.

Results of chew count estimation suggest that the flex sensor of the system can provide an objective and accurate estimation of chewing behavior. Our previous studies proposed chew count estimation from a strain-gauge sensor attached to the skin's surface using medical adhesive or medical tape. Few limitations of such configuration could be individual's comfort, sensitive skin and risk of sensor detachment due to wet skin or perspiration. The current work proposes a novel sensor that can provide an estimation of chew counts without being adhesively attached to the body.

An image-based method typically acquires images either from handheld devices (e.g. smartphones, digital camera, tablet) or wearable cameras. To provide objective information about food consumption, the images can be analyzed by nutritionists. The studies that utilize the wearable camera to capture images continuously, may form a large dataset for nutritionist's review and entail privacy concerns. This work explored an approach that significantly reduces the number of images to be analyzed by considering only the images where the sensors detected food intake event. Table 5-4 shows that during a typical day on an average only ~9% images are food images.

The captured images may also be processed by automatic image recognition algorithms, rather than manually reviewed by nutritionists. Several methods use computationally complex algorithms (e.g. deep neural network) for food recognition and require high processing resources to process thousands of images captured during the whole day. These methods are not currently sufficiently accurate for field deployment. The system presented in this work proposed a method of summarizing a large collection of images and selecting just food intake images for further processing, which may speed up the performance.

Unlike the methods, that use the two images of the meal before and after to analyze food intake, the proposed system would provide images for the full duration of the meal. There is a possibility that one can place a food item on top of another which may cause some food items to be concealed. Therefore, a series of food images would provide the meal progression and help in detecting all food items.

Regarding privacy concerns, the results from privacy concerns questionnaire bolster the fact that the majority of participants would feel more secure if the device capture images only during food intake.

The proposed system explores the potential to implement sensor-guided image capture. The system also suggests that one can expect potential improvement in battery life if the camera captures images only during food intake. Apart from battery life, a small battery can potentially be utilized to cover the whole day and provide sufficient food images for dietary monitoring. The device with a small battery would also improve wearability.

Another advantage of the proposed method is that the image review of the food intake events allowed for further validation of ingested eating episodes. The visual review identified three false positives in sensor-detected eating episodes.

Overall, the presented wearable sensor system was able to detect all eating events and reduce the number of images significantly for analysis. Also, the system showed around ~90% accuracy in chew count estimation using a new chewing sensor. However, the work is not free from limitations and needs further work to improve. One limitation is that the system detects food intake using chewing information only. The liquid intake detection is still needed to be incorporated in the system. Another limitation is that the study was conducted in a small population. Further improvement can be done in the device by validating the method on more subjects in a free-living environment.

5.5 Conclusion

In this paper, we propose a novel wearable sensor system that can detect eating instances and demonstrate the number of images can be reduced significantly by analyzing the images during food intake event only. In leave-one-subject-out cross-validation experiments, an average of 90.6% accuracy in food intake detection and 90.59% of chew count estimation were obtained. Detection of most of the eating episodes and the number of food images compared to all captured images provide the promise to develop sensor-triggered image capture system. Further works are needed to both implement the system and validate on more participants in a free-living environment.

5.6 References

[1] D. E. Wilfley, M. B. Schwartz, E. B. Spurrell, and C. G. Fairburn, "Using the eating disorder examination to identify the specific psychopathology of binge eating disorder," *Int J Eat Disord*, vol. 27, no. 3, pp. 259–269, Apr. 2000.

[2] C. G. Fairburn, "Eating Disorders," in *eLS*, John Wiley & Sons, Ltd, 2001.

[3] J. M. De Castro, "Methodology, Correlational Analysis, and Interpretation of Diet Diary Records of the Food and Fluid Intake of Free-living Humans," *Appetite*, vol. 23, no. 2, pp. 179–192, Oct. 1994.

[4] N. Day, N. McKeown, M. Wong, A. Welch, and S. Bingham, "Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium," *International Journal of Epidemiology*, vol. 30, no. 2, pp. 309–317, Apr. 2001.

[5] S. S. Jonnalagadda *et al.*, "Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique," *J Am Diet Assoc*, vol. 100, no. 3, pp. 303–308; quiz 309–311, Mar. 2000.

[6] D. A. Schoeller, L. G. Bandini, and W. H. Dietz, “Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method,” *Can. J. Physiol. Pharmacol.*, vol. 68, no. 7, pp. 941–949, Jul. 1990.

[7] A. Doulah *et al.*, “Meal Microstructure Characterization from Sensor-Based Food Intake Detection,” *Front. Nutr.*, vol. 4, 2017.

[8] J. Fontana and E. Sazonov, “Detection and characterization of food intake by wearable sensors,” in *Wearable Sensors: Fundamentals, Implementation and Applications*, 1st edition., Academic Press, 2014, pp. 591–616.

[9] C. K. Martin *et al.*, “Validity of the Remote Food Photography Method (RFPM) for Estimating Energy and Nutrient Intake in Near Real-Time,” *Obesity*, vol. 20, no. 4, pp. 891–899, Dec. 2011.

[10] J. Noronha, E. Hysen, H. Zhang, and K. Z. Gajos, “Platemate: Crowdsourcing Nutritional Analysis from Food Photographs,” in *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, 2011, pp. 1–12.

[11] F. Zhu, “Multilevel Image Segmentation With Application In Dietary Assessment And Evaluation,” Ph.D. Dissertation, Purdue University, 2011.

[12] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” *Pervasive and Mobile Computing*, vol. 8, no. 1, pp. 147–163, Feb. 2012.

[13] M. H. Rahman *et al.*, “Food Volume Estimation in a Mobile Phone Based Dietary Assessment System,” in *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, 2012, pp. 988–995.

[14] A. Myers *et al.*, “Im2Calories: Towards an Automated Mobile Vision Food Diary,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1233–1241.

[15] W. Jia *et al.*, “Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera,” *Public Health Nutrition*, vol. 17, no. 8, pp. 1671–1681, Aug. 2014.

[16] W. Jia *et al.*, “Automatic food detection in egocentric images using artificial intelligence technology,” *Public Health Nutr*, pp. 1–12, Mar. 2018.

[17] L. Gemming *et al.*, “Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall,” *British Journal of Nutrition*, vol. 113, no. 2, pp. 284–291, Jan. 2015.

[18] P. Kelly *et al.*, “An ethical framework for automated, wearable cameras in health behavior research,” *Am J Prev Med*, vol. 44, no. 3, pp. 314–319, Mar. 2013.

[19] J. Liu *et al.*, “An Intelligent Food-Intake Monitoring System Using Wearable Sensors,” in *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2012, pp. 154–160.

[20] C. Pettitt, J. Liu, R. M. Kwasnicki, G.-Z. Yang, T. Preston, and G. Frost, “A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate,” *British Journal of Nutrition*, vol. 115, no. 1, pp. 160–167, Jan. 2016.

[21] S. Sen, V. Subbaraju, A. Misra, R. K. Balan, and Y. Lee, “The case for smartwatch-based diet monitoring,” in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015, pp. 585–590.

[22] Y. Dong, A. Hoover, J. Scisco, and E. Muth, “A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking,” *Appl Psychophysiol Biofeedback*, vol. 37, no. 3, pp. 205–215, Sep. 2012.

[23] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting periods of eating during free-living by tracking wrist motion,” *IEEE J Biomed Health Inform*, vol. 18, no. 4, pp. 1253–1260, Jul. 2014.

[24] H. Kalantarian and M. Sarrafzadeh, “Audio-based detection and evaluation of eating behavior using the smartwatch platform,” *Computers in Biology and Medicine*, vol. 65, pp. 1–9, Oct. 2015.

[25] E. Thomaz, C. Zhang, I. Essa, and G. D. Abowd, “Inferring Meal Eating Activities in Real World Settings from Ambient Sounds: A Feasibility Study,” *IUI*, vol. 2015, pp. 427–431, 2015.

[26] O. Amft, “A wearable earpad sensor for chewing monitoring,” in *2010 IEEE Sensors*, 2010, pp. 222–227.

[27] E. Thomaz, I. Essa, and G. D. Abowd, “A Practical Approach for Recognizing Eating Moments with Wrist-mounted Inertial Sensing,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2015, pp. 1029–1040.

[28] M. Shuzo *et al.*, “Wearable Eating Habit Sensing System Using Internal Body Sound,” *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, vol. 4, no. 1, pp. 158–166, 2010.

[29] S. Päßler and W.-J. Fischer, “Acoustical method for objective food intake monitoring using a wearable sensor system,” in *2011 5th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2011*, 2011, pp. 266–269.

[30] E. 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, 2012.

[31] J. M. Fontana, M. Farooq, and E. S. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING - in preparation*, 2013.

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

[33] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner, “Detecting Mastication: A Wearable Approach,” in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, New York, NY, USA, 2015, pp. 247–250.

[34] K. Fueki, T. Sugiura, E. Yoshida, and Y. Igarashi, "Association between food mixing ability and electromyographic activity of jaw-closing muscles during chewing of a wax cube," *J Oral Rehabil*, vol. 35, no. 5, pp. 345–352, May 2008.

[35] K. Kohyama, E. Hatakeyama, T. Sasaki, H. Dan, T. Azuma, and K. Karita, "Effects of sample hardness on human chewing force: a model study using silicone rubber," *Archives of Oral Biology*, vol. 49, no. 10, pp. 805–816, Oct. 2004.

[36] V. A. Bousdras *et al.*, "A novel approach to bite force measurements in a porcine model in vivo," *International Journal of Oral and Maxillofacial Surgery*, vol. 35, no. 7, pp. 663–667, Jul. 2006.

[37] J. M. Fontana and E. S. Sazonov, "Evaluation of Chewing and Swallowing Sensors for Monitoring Ingestive Behavior," *Sensor Letters*, vol. 11, no. 3, pp. 560–565, Mar. 2013.

[38] E. Sazonov *et al.*, "Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior," *Physiol Meas*, vol. 29, no. 5, pp. 525–541, May 2008.

[39] O. Makeyev, P. Lopez-Meyer, S. Schuckers, W. Besio, and E. Sazonov, "Automatic food intake detection based on swallowing sounds," *Biomedical Signal Processing and Control*, vol. 7, no. 6, pp. 649–656, Nov. 2012.

[40] 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, Jun. 2014.

[41] 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.

[42] M. Farooq and E. Sazonov, "Automatic Measurement of Chew Count and Chewing Rate during Food Intake," *Electronics*, vol. 5, no. 4, p. 62, Sep. 2016.

[43] M. Farooq and E. Sazonov, "Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses with Piezoelectric Sensor," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, no. 99, pp. 1–1, 2016.

[44] E. Sazonov *et al.*, "Toward Objective Monitoring of Ingestive Behavior in Free-living Population," *Obesity*, vol. 17, no. 10, pp. 1971–1975, 2009.

[45] J. M. Fontana *et al.*, "Energy intake estimation from counts of chews and swallows," *Appetite*, vol. 85, pp. 14–21, Feb. 2015.

[46] P. Lopez-Meyer, S. Schuckers, O. Makeyev, J. M. Fontana, and E. Sazonov, "Automatic identification of the number of food items in a meal using clustering techniques based on the monitoring of swallowing and chewing," *Biomedical Signal Processing and Control*, vol. 7, no. 5, pp. 474–480, Sep. 2012.

[47] P. Lopez-Meyer, O. Makeyev, S. Schuckers, E. L. Melanson, M. R. Neuman, and E. Sazonov, "Detection of Food Intake from Swallowing Sequences by Supervised and Unsupervised Methods," *Ann Biomed Eng*, vol. 38, no. 8, pp. 2766–2774, Aug. 2010.

[48] A. F. Subar *et al.*, "The Automated Self-Administered 24-Hour Dietary Recall (ASA24): A Resource for Researchers, Clinicians and Educators from the National Cancer Institute," *J Acad Nutr Diet*, vol. 112, no. 8, pp. 1134–1137, Aug. 2012.

[49] "aTimeLogger 2 - mobile time tracking app | aLoggers." [Online]. Available: http://www.atimelogger.com/. [Accessed: 18-Jun-2018].

[50] "Food Processor Nutrition Analysis Software," *ESHA Research*. .

[51] J. M. C. Po, J. A. Kieser, L. M. Gallo, A. J. Tésenyi, P. Herbison, and M. Farella, "Time-frequency analysis of chewing activity in the natural environment," *J. Dent. Res.*, vol. 90, no. 10, pp. 1206–1210, Oct. 2011.

[52] M. Farooq and E. Sazonov, "A Novel Wearable Device for Food Intake and Physical Activity Recognition," *Sensors*, vol. 16, no. 7, p. 1067, Jul. 2016.

[53] M. Farooq and E. Sazonov, "Accelerometer-Based Detection of Food Intake in Free-Living Individuals," *IEEE Sensors Journal*, vol. 18, no. 9, pp. 3752–3758, May 2018.

[54] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.

[55] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[56] O. of the P. C. of Canada, "2016 Survey of Canadians on Privacy," 26-Jan-2017. [Online]. Available: https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2016/por_2016_12/. [Accessed: 09-Jun-2018].

[57] A. Doulah, X. Yang, J. Parton, J. Higgins, M. McCrory, and E. Sazonov, "The importance of field experiments in testing of sensors for dietary assessment and eating behavior monitoring," presented at the The 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'18), Honolulu, HI, 2018.

CHAPTER 6

A NOVEL METHOD OF ESTIMATING ENERGY INTAKE IN LABORATORY AND FREE-LIVING CONDITIONS USING A WEARABLE SENSOR AND FOOD IMAGERY.

The full validation part of this study is not complete yet and the manuscript will be submitted after the study.

Abstract

Accurate assessment and proper understanding of energy intake (EI) in the free-living environment may play an important role in the prevention and treatment of obesity. While, traditional self-report methods provide information about EI, a variety of factors including misreporting, forgetfulness, errors in portion size estimation can lead to low accuracy in EI measurement. To overcome these problems, researchers have proposed several image-based and sensor-based EI estimation methods. Image-based methods were proposed to estimate portion size and EI, but most of these methods require a fiducial marker in the image to provide a reference for both portion size and color, which may not be suitable for regular use. Several studies have explored wearable sensors and extracted microstructure information (e.g. counts of chews, bites, and swallows) to estimate ingested mass and EI, but they do not account for the type of food being consumed. We propose a novel method combing the EI estimates from a sensor-based regression model and food images. The method was developed and validated on data collected from 4 subjects wearing the wearable sensor system (Automatic Ingestion Monitor, AIM-2) for 24h in pseudo-free-living and 24h in a free-living environment. EI was estimated by four methods: a sensor model based on sensor-derived

chew counts, sensor-assisted image model from a manual review of the food images, the average of the sensor and image models, and self-reported food diary. Weighed food records were used as the gold standard. The respective estimates of meal EI had (mean ± standard deviation) absolute percentage errors of 25.5% ± 24.1%, 30.0% ± 26.0%, 22.4% ± 15.3% and 47.9% ± 48.7% respectively. These results suggest that the average estimation of EI from sensor and images may offer a promising alternative to overcome limitations of self-report.

6.1 Introduction

Obesity and overweight are a growing global epidemic and are major contributors to the development of chronic diseases such as type 2 diabetes, asthma, cardiovascular diseases, cancers, and musculoskeletal disorders. Obesity is often considered to be a result of a sustained energy imbalance whereby energy intake (EI) exceeds energy expenditure (EE), causing an accumulation of fat in the body [1], [2]. An accurate understanding of EI can help develop strategies to combat obesity. Despite decades of obesity research, an accurate estimation of EI in human is still a difficult challenge. The doubly labeled water (DLW) method is considered the most precise reference for estimating EI in free-living humans over a long period of time [3]. However, the method is highly expensive and requires specialized training and equipment. Additionally, the DLW method does not identify individual eating episodes.

Traditional methods of EI assessment largely rely on self-report in various forms. Among the validated self-report methods, food frequency questionnaires (FFQs), 24 h recalls, weighed food records (WFRs) and short assessment screeners, have been mostly adopted for EI measurement [4], [5]. Most of these methods are recognized to be highly inaccurate due to their reliance on participants own declaration (self-report) and associated underreporting, misreporting and not-reporting of EI [6].

As an alternative to self-report, various methods have been suggested utilizing food images and wearable on-body sensors. With the advancement in smartphone technology and wearable devices, the feasibility of capturing food images has improved significantly. The images of food items consumed in a meal can be used to track EI. Several smartphone apps (MyFitnessPal [7], LoseIt [8], MyNetDiary [9]) have been developed in tracking EI; however, they require manual data entry which is time-consuming and tedious. Consequently, the users lose their interest in using such apps for extended periods of time [10]. While these apps may be useful to individuals trying to track their diet, they may be insufficiently accurate for research applications. Several approaches utilized smartphone, hand-held, and wearable cameras to acquire food images and relied on expert nutritionists to analyze the image offline for EI estimation. The study of [11] presented an EI estimation approach named "Remote Food Photography Method" in which the participants used a smartphone to capture images of their foods and plate waste and expert nutritionists analyzed the images to obtain EI. Alternatively, several existing works used computer vision algorithms to segment food images, recognize foods, estimate portion size/volume and compute energy content of foods. In [12] a mobile food record application was proposed to estimate EI of food items. The method consisted of the segmentation of food items in the image, food recognition using support vector machine classifier, portion estimation using through 3-D reconstruction and extraction of energy information that was consumed during the meal. Another study [13] proposed a method to estimate EI in food's image by computing the volume of food portions using image segmentation followed by SVM classification. To compute the volume of the food portion, a thumb was placed beside the dish as a reference object. Recently, the study of [14] proposed a system called 'Im2Calories' that recognized the food items in a meal and predicted energy contents. The system was validated on the food images from restaurants where food menus were available. To estimate EI, the system utilized GoogLeNet convolutional neural network (CNN) model to automatically recognize food items and estimate volume based on depth and voxel representation of the food items. The authors in [15] proposed a chest-worn electronic device; 'eButton' that automatically captured images of consumed food in every two seconds interval. The volume of food items was then recognized by segmenting items based on color, texture and a complexity measure. The food name and volume information was then used to estimate energy and nutrient content from the Food and Nutrient Database for Dietary Studies (FNDDS, a public domain database developed by the U. S. Department of Agriculture). One of the major challenges in image-based EI estimation is to accurately estimate the portion sizes of food items. Current approaches to portion size estimation require the user to carry and place a fiducial marker (known dimensional and color references) next to the meal before capturing the image. This procedure might not be convenient all the time and post-processing of the food image is time-consuming and sometimes inaccurate.

Development of wearable sensors for EI estimation has gained substantial attention in recent years. They could potentially allow supplying rich information about food intake by capturing timing, duration, rate of ingestion and thereby estimate ingested mass, bite-weight, EI of ingested food items without imposing heavy burden on the user [16].

The authors in [17] proposed an ear-pad chewing sound sensor to identify chewing sequences during food intake and developed a linear regression model utilizing microstructure variables extracted from chewing events to predict bite weight. The work presented in [18], utilized a sensor system called BiteCounter [19] and proposed a multiple regression model to predict an individual's mean EI per hand-to-mouth gesture. The method employed anthropometrics variables such as height, weight, age, sex, and waist-to-hip ratio to develop the regression model. Another study [20] proposed a multi-modal sensor system comprising of in-ear microphone, head and wrist motion sensor to estimate ingested mass. The authors extracted significant features from audio and motion sensor to capture eating activities and trained a random forest regression with 40 trees to predict ingested mass. Finally, the estimated mass was utilized to estimate EI from food databases. The study of [21] proposed to use speech signal to estimate EI. The approach first converted the speeches into text utilizing speech processing. Next, a text matching algorithm was proposed using pattern mapping to find the food name in the database and estimate EI.

Our research group presented a wearable sensor system in [22] and proposed a method to predict ingested food mass using separate mass models for intake of solids and liquids. The prediction models were developed based on linear regression using the total number of chews and swallows during ingestion period. The study of [23] used the sensor system "Automatic Ingestion Monitor (AIM)" comprised of piezoelectric strain gauge sensors for monitoring jaw motion, hand-to-mouth motion sensor, and body motion sensor with a smartphone application [24], presented an EI estimation model from counts of chews and swallows. In laboratory settings, participants were asked to consume their meals wearing the sensor systems. The participants eating behavior was recorded through a video monitoring system. Features extracted from sensors and video annotation were used to build EI models. The EI was obtained by multiplying the mass intake with energy density of the food item. While wearable sensors offer potential methods for EI estimation, one major challenge of such methods is that they do not provide explicit information of consumed food items.

To address, some of the issues of both image-based and sensor-based EI estimation methods, a novel wearable sensor system (Automatic Ingestion Monitor, AIM-2) has been used in this study. The system combined a wearable camera that passively captured food images with sensors that detected food intake and quantified chewing microstructure (e.g. chewing, chewing rate, chewing durations etc). To objectively estimate EI, three approaches were explored. Previously our research group proposed EI estimation method utilizing counts of chews and swallows [23]. In this work, firstly, we explored the AIM-2 to estimate EI using a model based on sensor-derived chew counts. Secondly, a sensor-assisted image model from manual review of the food images by an expert nutritionist was investigated to identify food items, food types, portion size and estimate EI. To facilitate the process of image review, a custom-designed software written in Java was used to enable review of food images within sensor-detected episodes, thus significantly reducing the reviewing burden for the human annotator. Thirdly, the average of the sensor and image models was computed. It was hypothesized that the averaging of two estimates would minimize the error. Lastly, a self-reported food diary was used to estimate EI to compare sensor and image models with self-report. The major contributions of the proposed work are: (1) estimation of EI utilizing both sensors and images and validation in the free-living environment; 2) estimation of EI without using any fiducial markers in the image; and (3) providing detailed information about the food items consumed during any eating episode.

6.2 Methods

6.2.1 Subjects and protocol

Four healthy subjects (4 males) with a mean (±SD) age of 24.00 ± 4.32 years (range: 20-30 years) and a mean (±SD) body mass index (BMI, in kg/m2) of 24.25± 1.89 (range: 21.5-25.6) participated in this study. Participants came to the laboratory for four visits. Two high-definition cameras (Contour Roam2 LLC, USA, and GW- 2061IP, GW Security, Inc. CA, USA) were mounted on the wall of the laboratory to record the experiment. All participants made their first visit before starting the experiment, read and signed an informed consent form. The University of Alabama's Institutional Review Board approved the study. Participants were excluded from the study if they possessed any of the conditions such as: 1) temporomandibular joint (TMJ) disease; 2) dysphagia; 3) difficulties with chewing. Next, the sensor system (AIM-2) was prepared for each of the participants and attached to the temple of non-prescription eyeglasses with heat shrink tubes. Participants with corrective lens were instructed to bring their spare eyeglasses for sensor attachment. The AIM-2 consisted of a module which integrated a miniature 5 Megapixel wide-angle lens camera, a low-power 3D accelerometer (ADXL362 from Analog Devices, Norwood, MA, USA) and a bending sensor (SpectraSymbol 2.2" flex sensor). For each of the participants, the attachment of the sensor to the eyeglasses was carefully checked and retained in the lab until the next visit. Next, the participants were trained to report dietary intake and activity using two mobile applications (Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool [25] and aTimeLogger [26]). The ASA24 phone application was used in the food diary mode [27]. The participants were trained on how to enter the portion size of each food items during eating episodes. A short training on visual estimation of portion size was given to the participants using food replicas. The participants were scheduled for the next visit and left the laboratory.

Figure 6-1. AIM-2 used to monitor chewing activities and capture passive images. The sensor module is attached to the temple of off-the-shelf wearable eyeglasses with heat shrink tubes.

On the first day of the experiment, the participants came to the laboratory three times (visits 2-4). Up to two participants were invited to the laboratory at the same time to simulate social eating. For the second visit, the participants came to the laboratory in the morning (between 7:00 AM - 9:00 AM) after an overnight fast. Upon the arrival, participants were given the AIM-2 device and reminded about the instructions for the experiment. Each participant was asked to consume a full meal (breakfast) purchased from the University food court. The participants chose the type and quantity of food items that they wanted to consume. Research assistant weighed each food item and used a nutrition software "Food Processor" [28] (detail procedure is described in Section- 6.2.4) to enter the details of food items. There was no restriction on the time required to finish the meal. When the participants had finished eating, the research assistant weighed leftovers and updated the amount consumed in the nutrition software. Visits 3 and 4 were made on the same day to consume lunch and dinner following the same procedure. The daily usual activities of the participants were not restricted on the first day except they had to consume their meals at the laboratory.

In the second day of the study, the participants continued wearing the device for 24 hours. During this period, the participants had virtually no restrictions on the types of the activities performed. They continued their normal daily routine with three exceptions: to take the AIM-2 off during 1) taking the shower 2) any water activities that may damage the device, and 3) sleeping. The participants were also reminded to self-report all of their laboratory and free-living eating episodes using the ASA24 mobile app (used a food diary). After completing the free-living part of the study, the participants reported to the laboratory to return the AIM-2.

6.2.2 Study design

Four different methods were proposed to estimate EI: 1) a sensor model based on sensor-derived chew counts; 2) sensor-assisted image model from manual review of the food images 3) the average of the sensor and image models and; 4) self-reported food diaries completed by participants using ASA24. Weighed food records were completed by a research assistant to provide reference for both training and validating the sensor model. For laboratory eating episodes, comparisons were made between all four methods mentioned above with respect to weighed records that were used as the gold standard. For free-living eating episodes, daily EI was computed and comparisons were made between sensor model, image review, average EI and self-reported food diaries. In addition, the estimated energy requirements for each participant were computed. The reporting error was expressed as mean absolute percentage error. The following subsections describe each of EI measurement elaborately.

6.2.3 EI Measurement: weighed food records ($\boldsymbol{EI_{WFR}}$)

A trained research assistant recorded the food items and weighed all foods and beverages individually before and after each eating episode. Data from weighed records were entered into the food analysis program, The Food Processor Nutrition Analysis Software (version 7.6, 2000, ESHA Research, Salem, OR) to derive EI of each eating episode. The Food Processor software is a validated software currently used in the field of nutrition research. It provides broad nutritional information about food items, including weight, kcal, caffeine, vitamin, total fat, grams, the percentage of protein, carbohydrates, and sugar. To obtain nutritional information with the software, the food items and the amount of each food items consumed had to be specified. The software automatically calculated the total energy content (kcal) and provided the report immediately.

6.2.4 EI Measurement: Sensor model based on sensor-derived chew counts ($\boldsymbol{EI_{SEN}}$)

Estimation of EI for each eating episode was performed utilizing sensor information. Fig. 6-2 illustrates the procedure for EI regression model. The eating episodes were detected using the food intake detection algorithm described in [chapter-5]. To detect the eating episodes, the AIM-2 sensor signals were first processed and divided into non-overlapping 3s fixed time segments called ‘epochs’. Features were extracted from the epochs. The features were then used to train and validate classification model that detected food intake epochs. The eating episodes were identified from the food intake epochs. Next, the chew counts algorithm was applied only on the epochs which were labeled as food intake by the classifier. Once the chew counts for each eating episodes were obtained, features were extracted from chew counts and flex sensor signal. These features were used to build the EI model.

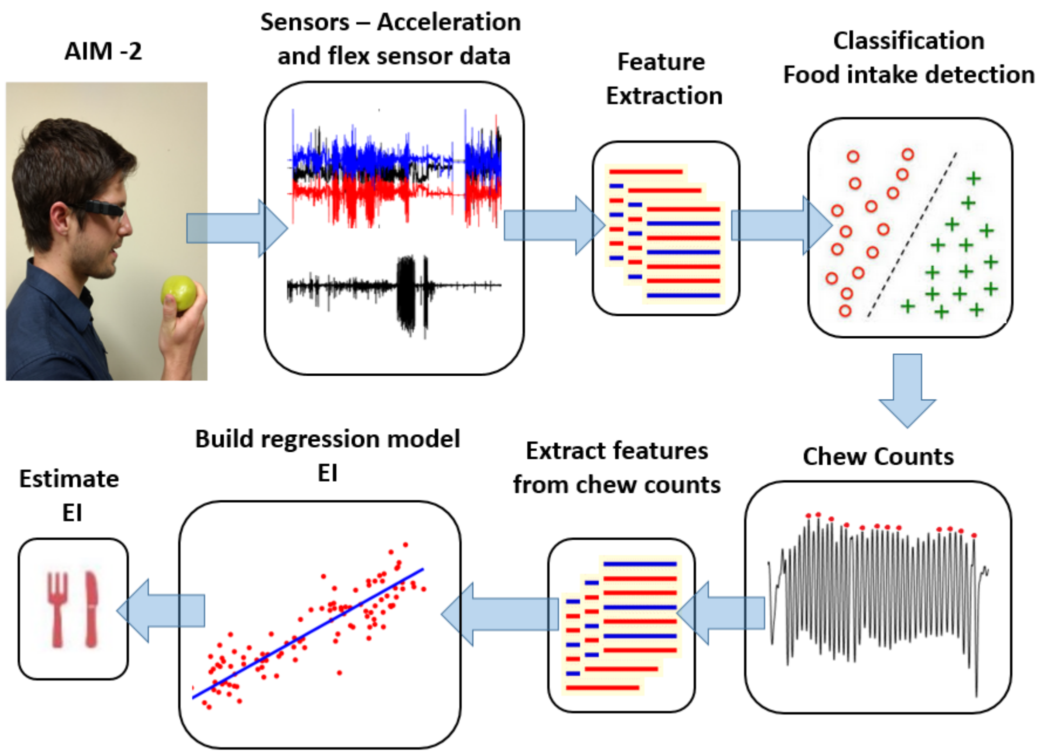

Figure 6-2. Flow diagram of EI sensor model based on sensor-derived chew counts.

Thirteen features were extracted for each eating episode from chew counts. The description of features is listed in Table 6-1.

Table 6-1 Features extracted from chew counts

Variable names	Description
numb_chews_seq_sensor	Number of chewing sequence per eating episode
total_chews	Total number of chews per eating episode
total_chews_du	Total chew duration per eating episode
avg_chews_perSeq	Average of chews on chewing sequence per eating episode
sd_chews_perSeq	Standard deviation of chews on chewing sequence per eating episode
var_chews_perSeq	Variance of chews on chewing sequence per eating episode
chewRate	total_chews/ total_chews_du
avg_chews_du_perSeq	Average chew duration on chewing sequence per eating episode
sd_chews_du_perSeq	Standard deviation chew duration on chewing sequence per eating episode
var_chews_du_perSeq	Variance chew duration on chewing sequence per eating episode

Along with chew related features, twelve flex sensor features were computed. These features were derived from epochs, which were detected as food intake by detection algorithm. The sum and average values of features were computed over the eating episode. These sensor features included the number of mean crossings, duration between mean crossings, entropy, waveform length, mean amplitude, periods from autocorrelation function, power of the signal and maximum frequency present in the signal spectrum.

$$\boldsymbol{EI_{SEN}} = a_n F_n + \cdots + a_1 F_1$$

where, $a_n, \ldots a_1$ are the regression coefficients, $F_n, \ldots F_1$ are the features extracted from chew counts and $\boldsymbol{EI_{SEN}}$ is estimated EI.

For model fitting (training), the $\boldsymbol{EI_{WFR}}$ was considered as response variable. The forward feature selection procedure was applied to select the most important features from the initial set of features. To avoid introducing bias into the regression model, feature selection was done on an independent dataset. The selected features were then used to build multiple linear regression model with their intercepts set to 0 to predict EI using leave-one-subject-out cross-validation. The model was trained with eating episodes from 3 participants, and tested with eating episodes from the remaining participant. The procedure was repeated 4 times such that each participant was used for testing once. Finally, a group regression model was developed for EI estimation in free-living eating episodes.

6.2.5 EI Measurement: Sensor-assisted image model from manual review of the food images ($\boldsymbol{EI_{IMG}}$)

Estimation of EI was performed by manually reviewing the AIM-2 images. Fig. 6-3 illustrates the procedure for EI estimation using image review. First, the sensor-detected eating episodes were used to identify the food images. A well-trained nutritionist at the

Colorado Clinical and Translational Sciences Institute's (CCTSI) Nutrition Core, then manually reviewed the images of sensor-detected eating episodes only and estimate EI for each eating episode.

A number of software packages have been developed for nutrition analysis [28]–[30]. These packages allow searching foods and brand products, entering weight, servings and portions of food items, and estimating nutritional contents. In order to review images and enter information into the software, one has to open food images and the software separately, which may be burdensome. To facilitate the review process, a novel Java-based software was designed which can simultaneously review images, recognize food items in the image and enter details to food items to obtain nutrition information (i.e. weight, kcal, carb, fat, protein etc.).

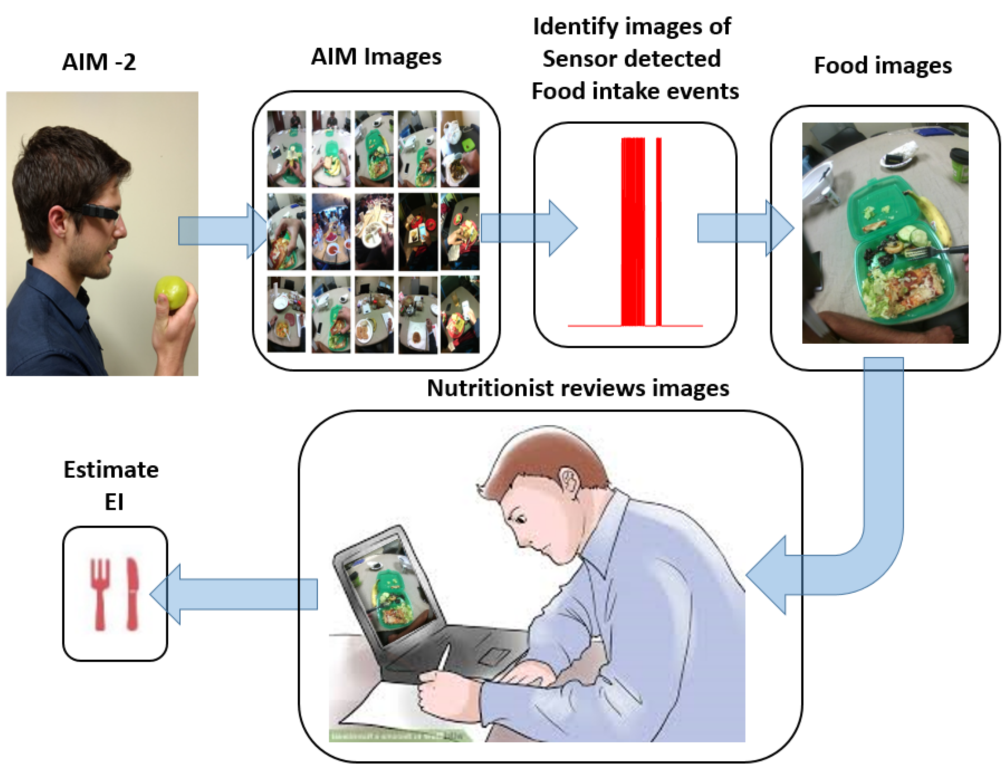

Figure 6-3. EI estimation using image review.

On top of that, the software also included an added option of viewing sensor-detected food intake events. Fig. 6-4 illustrates the interface of the software. A graphical user interface (GUI) was developed in the software that included three main panels and two bars. Three panels were image panel (displaying thumbnail image), plot panel (plotting sensor matrices such as food intake detection from flex or accelerometer sensor, chewing count etc), and meal panel (in which user can add eating episode and consumed food items). The two bars were menu bar (contained menus) and status bar (provided status of total number of images, time of selected image and other important messages). The image panel contained all image captured by AIM-2 during the day and plot panel loaded the food intake detection labels. A time-based link was established between the images and food intake detection plot such that if the user would click on the plot, the corresponding image that was captured at that time would show up at the image panel. To see the images clearly, the thumbnail images could be zoomed-in by double-clicking the image. Further zoom-in feature was also possible. To compute the EI and related parameters, two major database (FNDDS and USDA food database [31]) were included in the software. From the selected food items, portion size of each food item, the software automatically computed the energy value of each food items using those two databases. Finally, the software generated a report files including EI and ingested mass information of each of the reviewed eating episodes.

In this study, the CCTSI Nutrition Core nutritionist reviewed the images from detected food intake events. All images of an eating episode were reviewed so that any leftovers could be identified and recorded. The EI of each eating episode was computed immediately upon entering information of food items. In this way, the nutritionist reviewed all the eating episodes, manually recorded food items information and estimated EI.

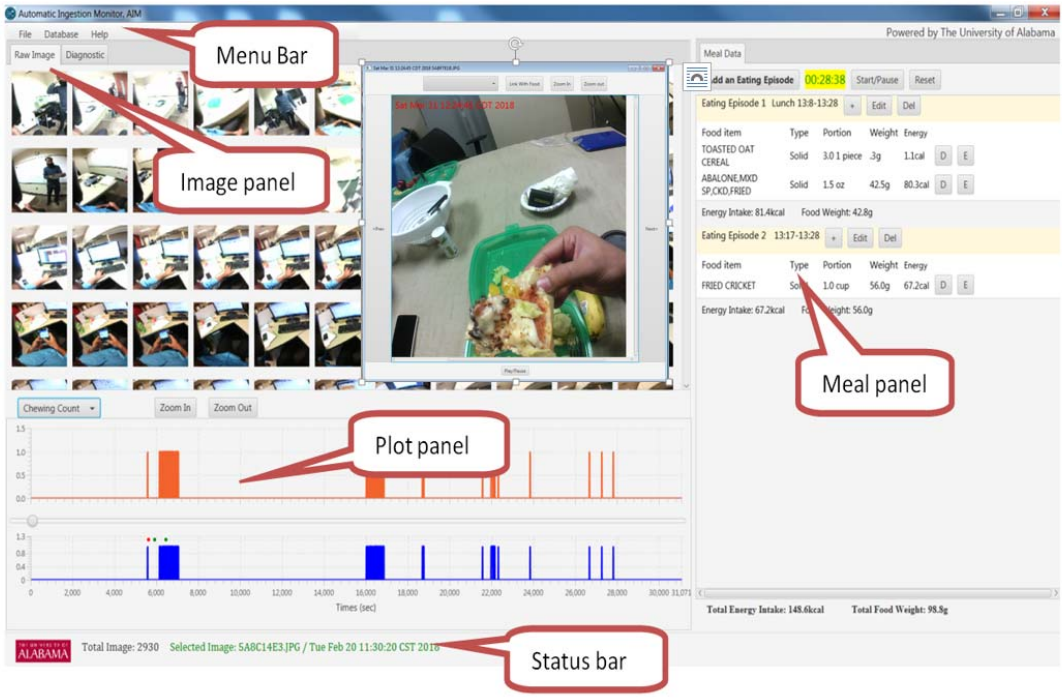

Figure 6-4 Screenshot of the main window of AIM-2 Java software.

6.2.6 EI Measurement: Average of Sensor model estimated EI and Image review estimation EI ($\boldsymbol{EI_{AVG}}$)

In this approach, estimation of EI was performed by taking the average of sensor model estimated EI and image review estimated EI for each eating episode.

$$\boldsymbol{EI_{AVG}} = \frac{\boldsymbol{EI_{SEN}} + \boldsymbol{EI_{IMG}}}{2}$$

6.2.7 EI Measurement: food diary – ASA24 ($\boldsymbol{EI_{ASA}}$)

During the experiment, the participants used a food diary, Automated Self-Administered 24-Hour (ASA24) Recall developed by the National Cancer Institute, to report dietary intake. The ASA24 uses an extensive food database and provides energy information for each item entered. The ASA24 also included over 17,000 portion size images to aid to facilitate portion size estimation. The ASA24 can be used to collect both food diary as well as 24-hour recalls. In this study, the ASA24 was used as food diary.

To minimize the inherent error of self-report, a short training on portion size estimation was given to all participants. In the training, four meals shown in Fig. 6-5 were formed with standard portion-sized food replicas (NASCO, Fort Atkinson, WI). All participants used ASA24 to enter meals information and were familiarized with standard units of portion measurements (i.e. cups, tablespoon, teaspoon etc.). During the study, at the end of each eating episodes, participants used ASA24 to record all foods items and beverages consumed.

Figure 6-5. Training meals for portion size estimation.

6.2.8 Estimated energy requirements (EER)

To provide a source for comparison in the free-living environment, the estimated energy requirements (EER) was computed. A set of equations and factors that account for an individual's EI, energy expenditure, age, sex, weight, height, and physical activity level

(PAL) are typically used to compute the daily EER. Based on the measurement of DLW studies, the EER consists of predictive equations for calculating the amount of EI that will maintain any individual's body weight [34]. The EER equations for men and women are given below:

Men: $864 - (9.72 \times age[y]) + PA \times (14.2 \times weight[kg] + 503 \times height[m])$

Women: $387 - (7.31 \times age[y]) + PA \times (10.9 \times weight[kg] + 660.7 \times height[m])$

where PA is the physical activity coefficients: PA - 1.00 (sedentary), 1.14 (low active), 1.27 (active), or 1.45 (very active). In this study, $PA = 1.12$ was used assuming that the participants were not involved in intensive physical activities (e.g. marathon, triathlon, playing in sports team).

6.2.9 Statistical Analysis

For laboratory eating episodes, the reporting error in estimated EI in an eating episode was computed for each of the methods (i.e. $\boldsymbol{EI_{SEN}}$, $\boldsymbol{EI_{IMG}}$, $\boldsymbol{EI_{AVG}}$, $\boldsymbol{EI_{ASA}}$) with respect to weighed food records($\boldsymbol{EI_{WFR}}$).

$$Reporting\ error = \left|\frac{\boldsymbol{EI - EI_{WFR}}}{\boldsymbol{EI_{WFR}}} \times 100\%\right| \ ; \boldsymbol{EI} : \ \boldsymbol{EI_{SEN}}, \boldsymbol{EI_{IMG}}, \boldsymbol{EI_{AVG}}, \boldsymbol{EI_{ASA}}$$

The mean and SD of mean absolute percentage errors for each method were calculated. To compare the mean difference between estimated EI from methods and weighed food record EI, Bland-Altman [32] plots, bias, and limit of the agreements (95% confidence intervals) were calculated. For free-living eating episodes, total daily EI from $\boldsymbol{EI_{SEN}}$, $\boldsymbol{EI_{IMG}}$, $\boldsymbol{EI_{AVG}}$ $\boldsymbol{EI_{ASA}}$ and, EER were compared. Then, one way ANOVA was analyzed to determine the significant difference among methods. Tukey–Kramer *post hoc* multiple comparisons analysis was performed to determine which methods differed from each other. All statistical analyses were done using MATLAB 2017 (from Mathworks Inc.).

6.3 Results

The dataset used to develop EI estimation models contained a total of 12 laboratory eating episodes and around 15 free-living eating episodes from four participants. In $\boldsymbol{EI_{SEN}}$ model, the forward feature selection algorithm selected features SD_chewRate_perSeq, chewRate as best features for EI estimation. Reporting errors for all EI estimation methods for laboratory eating episodes are presented in Table 6-2. The results of reporting errors demonstrate that the best accuracy in laboratory eating episodes was obtained in $\boldsymbol{EI_{AVG}}$ with 22.37% ± 15.32% error. The results of the Bland-Altman analysis are presented in Table 6-3. The Bland-Altman analysis indicated that the biases in EI estimation for all three methods and self-report were positive (overestimated). Figures 6-6, 6-7, 6-8 and 6-9 show the Bland-Altman plots each of the methods compared to weighed record.

Table 6-2 Reporting errors (in %) for EI estimation for laboratory meals with respect to the weighed records EI

Methods	EI estimation method	Reporting error (%) Mean	Reporting error (%) SD
1	$\boldsymbol{EI_{SEN}}$	25.54	24.15
2	$\boldsymbol{EI_{IMG}}$	30.03	26.02
3	$\boldsymbol{EI_{AVG}}$	22.37	15.32
4	$\boldsymbol{EI_{ASA}}$	48.28	48.51

Table 6-3 Bland- Altman analysis for the EI estimation for laboratory meals

Methods	Energy Intake estimation	Bias	SD of individual difference	Lower LOA	Higher LOA
1	$\boldsymbol{EI_{SEN}}$	-2.56	252.57	-497.61	492.48
2	$\boldsymbol{EI_{IMG}}$	-82.46	306.11	-682.43	517.51
3	$\boldsymbol{EI_{AVG}}$	-13.98	228.39	-461.62	433.66
4	$\boldsymbol{EI_{ASA}}$	-194.84	300.42	-783.66	393.99

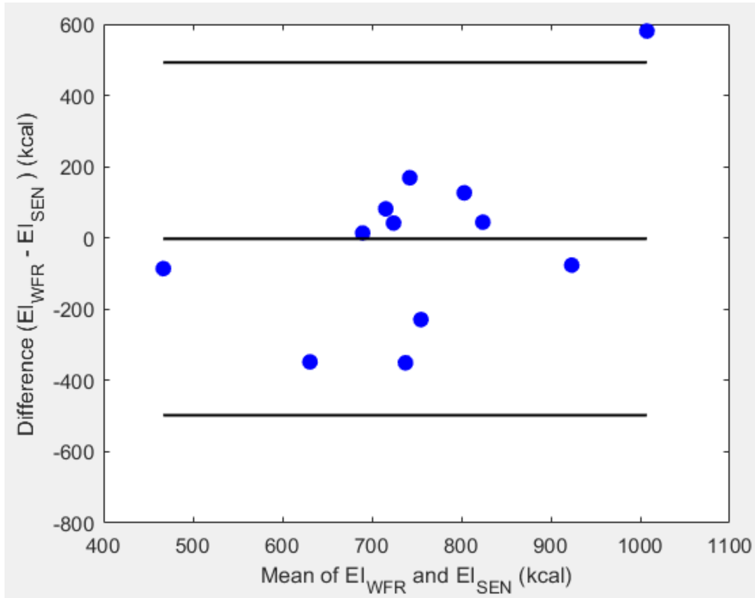

Figure 6-6. Bland-Altman plots for $\boldsymbol{EI_{WFR}}$ vs $\boldsymbol{EI_{SEN}}$. The blue dots represent estimated EI for each eating episodes.

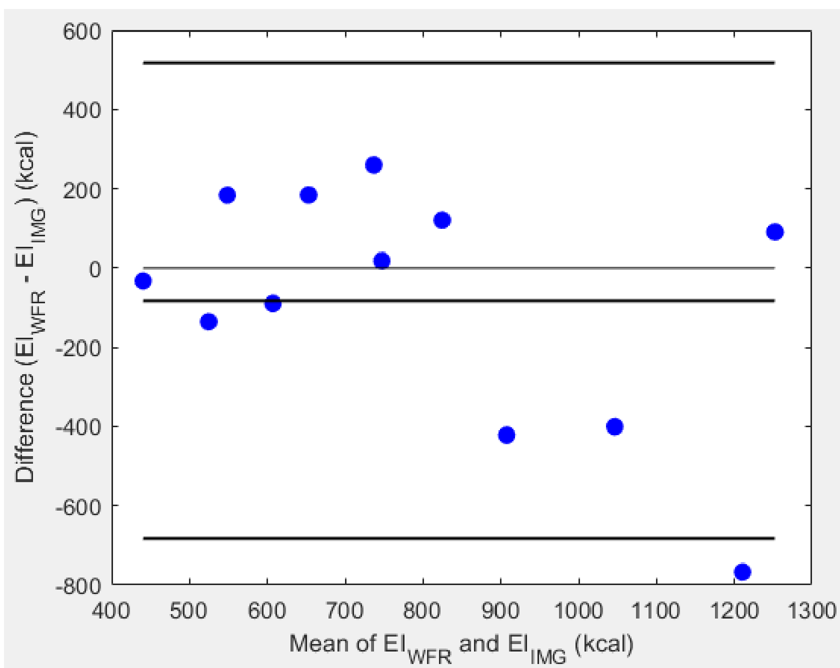

Figure 6-7. Bland-Altman plots for $\boldsymbol{EI_{WFR}}$ vs $\boldsymbol{EI_{IMG}}$. The blue dots represent estimated EI for each eating episodes.

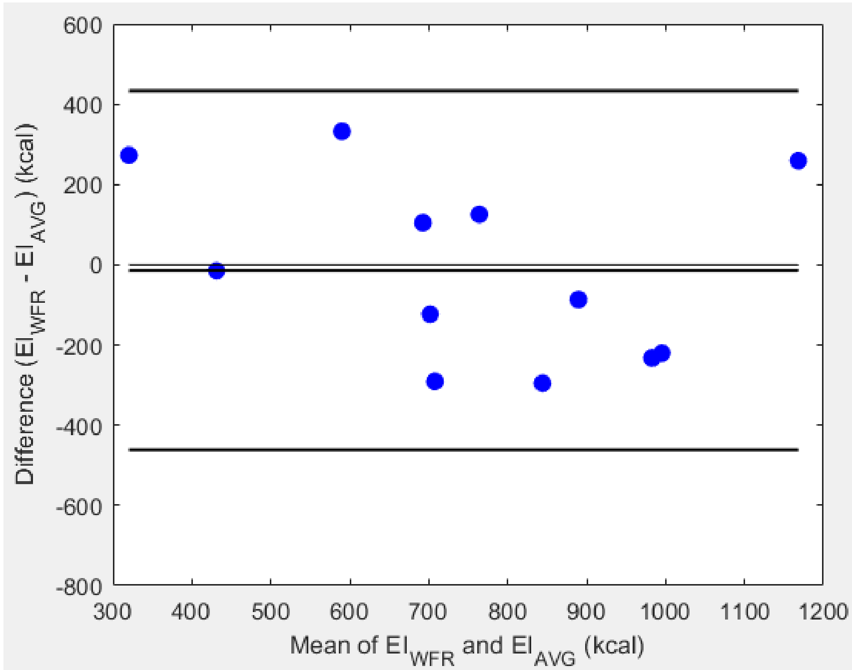

Figure 6-8. Bland-Altman plots for $\boldsymbol{EI_{WFR}}$ vs $\boldsymbol{EI_{AVG}}$. The blue dots represent estimated EI for each eating episodes.

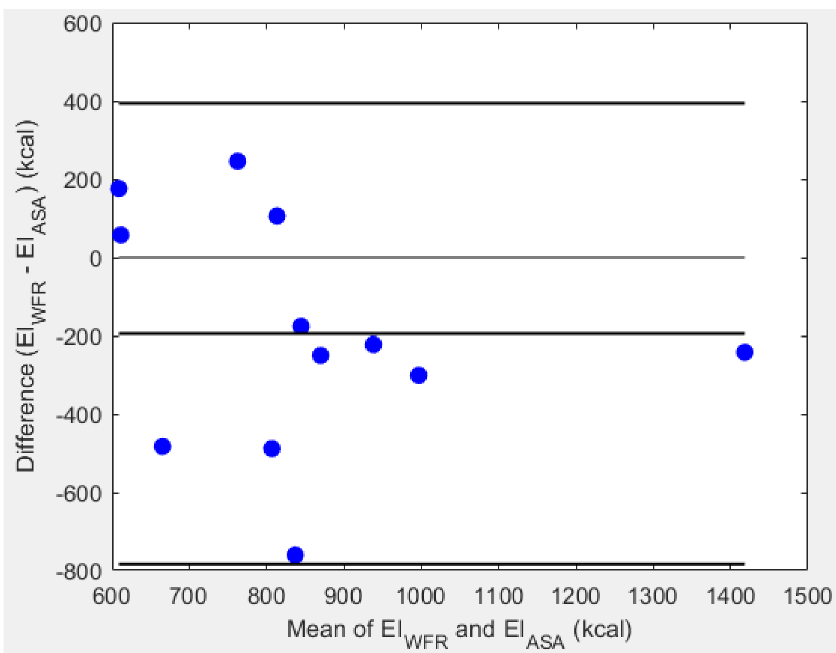

Figure 6-9. Bland-Altman plots for $\boldsymbol{EI_{WFR}}$ vs $\boldsymbol{EI_{ASA}}$. The blue dots represent estimated EI for each eating episodes.

For free-living eating episodes, estimated daily EI from EI_{SEN}, EI_{IMG}, EI_{AVG} , EI_{ASA} and EER are presented in Table 6-4. The one-way ANOVA test showed that the means of daily EI computed from all measurements (EI_{SEN}, EI_{IMG}, EI_{AVG}, EI_{ASA} and EER) were not significantly different (p-value > 0.05). Boxplots of estimated daily EI from each of the methods are shown in Fig. 6-10.

Table 6-4 Daily EI for free-living eating episodes

Participant	EI_{SEN}	EI_{IMG}	EI_{AVG}	EI_{ASA}	EER
1	1464.37	1279.81	1372.09	1679.87	2759.96
2	2726.54	1977.94	2352.24	1996.21	2754.14
3	2585.40	3283.42	2934.41	4076.24	3094.25
4	1840.32	2019.58	1929.95	2974.73	2948.02
Average	2154.16	2140.19	2147.17	2681.76	2889.09

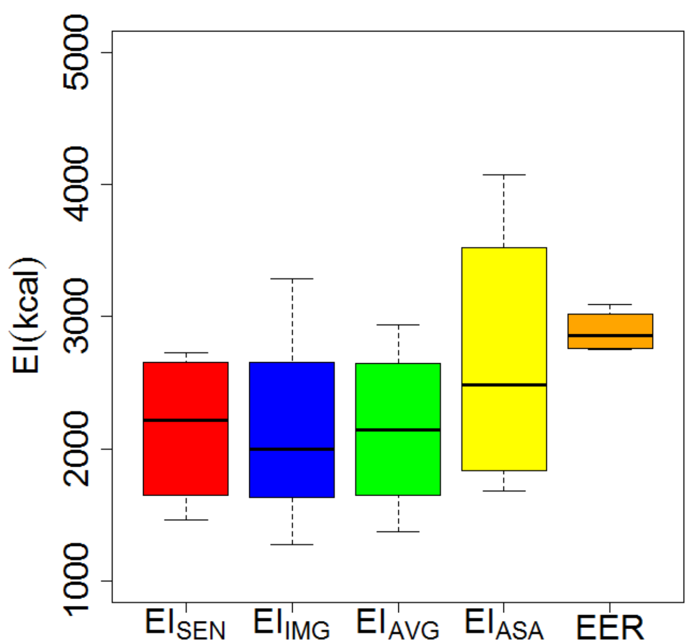

Figure 6-10. Boxplots of daily EI for free-living eating episodes.

6.4 Discussion

Several studies have attempted to estimate daily EI using image-based methods and wearable on-body sensors. However, while some of the methods show great promise under certain circumstances, objective, accurate and cost-efficient methods for estimation of EI are yet to be developed. The main objective of the present work was to explore the ability of wearable device consisting of sensors and camera that can provide a semi-automatic approach to estimate EI accurately both in the laboratory and free-living environments. The most important finding of this work was that the accuracy of EI estimation can be improved utilizing both image and sensor information.

Estimation of EI was performed in four ways – sensor model based on sensor-derived chew counts, sensor-assisted image model from manual review of the food images, the average of the sensor and image models and self-reported food diaries. Results suggest that the $\boldsymbol{EI_{AVG}}$ exhibited the best EI estimation reporting error of 22.4% ± 15.3% for laboratory eating episodes which outperforms our previous reported accuracy in [23]. As per the hypothesis, the performance of EI estimation improved when the sensor model and image review were averaged. One potential possibility of this improvement could be the individual contribution of $\boldsymbol{EI_{SEN}}$ and $\boldsymbol{EI_{IMG}}$. Since our sensor primarily sensitive to chewing, therefore the sensor model performed better in estimating EI related to solid food intake. On the other hand, the image review led to a better performance in detecting liquid food items compared to the sensor model. However, there exists a possibility of inaccuracy in portion size estimation of solid food items during image review. Therefore, the averaging of these two methods may potentially balance out the effect of error and provide improved accuracy in EI estimation.

Apart from accuracy, the current work proposed a group EI model rather than an individually calibrated model, which may potentially offer a more robust model to account for inter-person variations. Additionally, the group model may potentially perform better in free-living conditions as the model can be generalized to larger populations. It was also observed that the reporting error of $\boldsymbol{EI_{SEN}}$ method was better than $\boldsymbol{EI_{IMG}}$ by at least 5%. For large dataset, the accuracy may increase and therefore one can expect better performance utilizing sensor-estimated mass in the calculation of EI rather than using image reviewed portion size. One potential reason could be the portion size estimation in image review. Literature implies that three major elements affect portion-size reports perception, conceptualization, and memory. Due to these human limitations, even well-trained nutritionists may make errors in portion size estimation.

The lower reporting bias of -2.56 was observed in $\boldsymbol{EI_{SEN}}$ model compared to other models. However, the narrowest limits of agreement was found in $\boldsymbol{EI_{AVG}}$ method. This results along with the reporting error indicate that $\boldsymbol{EI_{AVG}}$ method may perform better in free-living environment where variety of food items may impact on EI estimates. Another observation is that the SD of individual differences of $\boldsymbol{EI_{IMG}}$ and $\boldsymbol{EI_{ASA}}$ are high and quite comparable. One potential reason could be that both methods employ the human expertise of self-report. The highest reporting error of ~48% was found in $\boldsymbol{EI_{ASA}}$ method, which could be due to error of portion size estimation.

For free-living eating episodes, methods were compared in estimating daily EI. The validity of estimated and self-reported EI was further assessed using EER equations developed from DLW studies. The ANOVA results suggest that there was no significant difference between the methods. Further studies will need to involve more participants to explore differences in estimating EI.

The most important features for $\boldsymbol{EI_{SEN}}$ model were found to be related to chew rates. Previous research [33] has demonstrated the possibility of change in the quantity of intake with chewing rates. The study showed that obese persons chewed less and ingested the meal more quickly than did lean persons. Although the current study only included healthy participants, the chewing rate related might influence EI. Future studies can be carried out for justification.

Another important contribution of this work is that the custom-designed annotation software. To the best of our knowledge, there is no such software that can review images, analyze food intake events detected by sensors and simultaneously perform the dietary assessment.

The major strengths of this work were the performance comparison between three EI estimation methods. A limitation of this study was that the sensor used can only detect chewing, therefore the liquid intake detection from the sensor system was not readily available. As the only the images of food intake events were reviewed, therefore liquid events may have missed and contributed to reporting error. Another limitation is that the study was conducted in a small population. Further improvement can be done in the EI models by incorporating more subjects in a free-living environment. The current work proposes a semi-automatic approach of estimating EI, in which the food intake detection was done automatically but the review of image was performed manually. In the future, instead of manual image review, deep learning based computer vision techniques could potentially be used for automatic food recognition from the images.

6.5 Conclusion

In this paper, we propose a novel wearable sensor system that can estimate EI in individual from sensors and images without the need of any fiducial markers in the images. In leave-one-subject-out cross-validation experiments, the average of EI model using chewing sensor-derive metrics and image review estimates had the best reporting error 22.4% ± 15.3%. Further study is needed in broader populations and validation in a free-living environment.

6.6 References

[1] K. D. Hall, S. B. Heymsfield, J. W. Kemnitz, S. Klein, D. A. Schoeller, and J. R. Speakman, “Energy balance and its components: implications for body weight regulation123,” Am. J. Clin. Nutr., vol. 95, no. 4, pp. 989–994, Apr. 2012.

[2] J. O. Hill, H. R. Wyatt, and J. C. Peters, “Energy Balance and Obesity,” Circulation, vol. 126, no. 1, pp. 126–132, Jul. 2012.

[3] W. A. Coward, “The doubly-labelled-water method: principles and practice,” Proc. Nutr. Soc., vol. 47, no. 3, pp. 209–218, Sep. 1988.

[4] A. H. C. Goris and K. R. Westerterp, “Underreporting of Habitual Food Intake Is Explained by Undereating in Highly Motivated Lean Women,” J. Nutr., vol. 129, no. 4, pp. 878–882, Apr. 1999.

[5] A. H. Goris, M. S. Westerterp-Plantenga, and K. R. Westerterp, “Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake,” Am. J. Clin. Nutr., vol. 71, no. 1, pp. 130–134, 2000.

[6] G. Johansson, A. Wikman, A. M. Ahrén, G. Hallmans, and I. Johansson, “Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living,” Public Health Nutr., vol. 4, no. 4, pp. 919–927, Aug. 2001.

[7] “Free Calorie Counter, Diet & Exercise Journal | MyFitnessPal.com.” [Online]. Available: https://www.myfitnesspal.com/. [Accessed: 25-Jun-2018].

[8] “Lose It! - Weight loss that fits.” [Online]. Available: https://www.loseit.com/. [Accessed: 21-Nov-2017].

[9] “Online Food Diary and Calorie Counter, with free iPhone, Android, and iPad mobile apps.” [Online]. Available: https://www.mynetdiary.com/. [Accessed: 25-Jun-2018].

[10] F. Cordeiro et al., “Barriers and Negative Nudges: Exploring Challenges in Food Journaling,” Proc. SIGCHI Conf. Hum. Factors Comput. Syst. CHI Conf., vol. 2015, pp. 1159–1162, Apr. 2015.

[11] C. K. Martin, H. Han, S. M. Coulon, H. R. Allen, C. M. Champagne, and 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., vol. 101, no. 03, pp. 446–456, 2009.

[12] F. Zhu et al., “The Use of Mobile Devices in Aiding Dietary Assessment and Evaluation,” IEEE J. Sel. Top. Signal Process., vol. 4, no. 4, pp. 756–766, Aug. 2010.

[13] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, “Measuring Calorie and Nutrition From Food Image,” IEEE Trans. Instrum. Meas., vol. 63, no. 8, pp. 1947–1956, Aug. 2014.

[14] A. Myers et al., “Im2Calories: Towards an Automated Mobile Vision Food Diary,” in 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1233–1241.

[15] W. Jia et al., “Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera,” Public Health Nutr., vol. 17, no. 8, pp. 1671–1681, Aug. 2014.

[16] A. Doulah et al., “Meal Microstructure Characterization from Sensor-Based Food Intake Detection,” Front. Nutr., vol. 4, 2017.

[17] O. Amft, M. Kusserow, and G. Troster, “Bite Weight Prediction From Acoustic Recognition of Chewing,” IEEE Trans. Biomed. Eng., vol. 56, no. 6, pp. 1663–1672, Jun. 2009.

[18] J. N. Salley, A. W. Hoover, M. L. Wilson, and E. R. Muth, “Comparison between Human and Bite-Based Methods of Estimating Caloric Intake,” J. Acad. Nutr. Diet., vol. 116, no. 10, pp. 1568–1577, Oct. 2016.

[19] Y. Dong, A. Hoover, J. Scisco, and E. Muth, “A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking,” Appl. Psychophysiol. Biofeedback, vol. 37, no. 3, pp. 205–215, Sep. 2012.

[20] M. Mirtchouk, C. Merck, and S. Kleinberg, “Automated Estimation of Food Type and Amount Consumed from Body-worn Audio and Motion Sensors,” in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 2016, pp. 451–462.

[21] N. Hezarjaribi, S. Mazrouee, and H. Ghasemzadeh, “Speech2Health: A Mobile Framework for Monitoring Dietary Composition from Spoken Data,” IEEE J. Biomed. Health Inform., vol. PP, no. 99, pp. 1–1, 2017.

[22] E. Sazonov et al., “Toward Objective Monitoring of Ingestive Behavior in Free-living Population,” Obesity, vol. 17, no. 10, pp. 1971–1975, May 2009.

[23] J. M. Fontana et al., “Energy intake estimation from counts of chews and swallows,” Appetite, vol. 85, pp. 14–21, Feb. 2015.

[24] J. M. Fontana, M. Farooq, and E. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” IEEE Trans. Biomed. Eng., vol. 61, no. 6, pp. 1772–1779, Jun. 2014.

[25] “Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool.” [Online]. Available: https://epi.grants.cancer.gov/asa24/. [Accessed: 21-Nov-2017].

[26] “aTimeLogger 2 - mobile time tracking app | aLoggers.” [Online]. Available: http://www.atimelogger.com/. [Accessed: 18-Jun-2018].

[27] “ASA24.” [Online]. Available: https://asa24.nci.nih.gov/. [Accessed: 29-Jun-2018].

[28] “Food Processor Nutrition Analysis Software,” ESHA Research. .

[29] “Nutritics | Nutrition Analysis Software For Professionals.” [Online]. Available: https://www.nutritics.com/p/home. [Accessed: 27-Jun-2018].

[30] "Nutrition Data System for Research – Nutritional Analysis Software - ndsr87072 - University of Minnesota Office for Technology Commercialization." [Online]. Available: http://license.umn.edu/technologies/ndsr87072_nutrition-data-system-for-research-nutritional-analysis-software. [Accessed: 27-Jun-2018].

[31] "USDA National Nutrient Database for Standard Reference." 2013.

[32] J. M. Bland and D. G. Altman, "Measuring agreement in method comparison studies," Stat. Methods Med. Res., vol. 8, no. 2, pp. 135–160, Apr. 1999.

[33] J. Li et al., "Improvement in chewing activity reduces energy intake in one meal and modulates plasma gut hormone concentrations in obese and lean young Chinese men," Am. J. Clin. Nutr., vol. 94, no. 3, pp. 709–716, Sep. 2011.

[34] I. of Medicine, Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. 2002.

CHAPTER 7
CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

This dissertation contributes to the work on development and validation of a wearable sensor system for dietary assessment and EI estimation. The main contribution of this work is the development of related signal processing and pattern recognition algorithms for automatic, accurate and objective monitoring of eating episodes and estimation of EI.

Firstly, the dissertation explores the existing works in EI estimation with a systematic review. Different methods such as traditional self-report, mathematical models, image-based method and wearable sensors were investigated. The review demonstrates that some of the methods do show great potentials under certain circumstances, however, there is a critical need for objective, accurate and cost-efficient methods for estimating of daily EI. Also the methods must provide good accuracy in free-living condition, accurate identification of food items and minimal burden to the user.

Then, the dissertation examines the desired time resolution of sensor-based food intake detection to obtain the meal microstructure parameters of food intake events that may help in EI estimation. Based on the analysis, the findings suggest that a sensor time-resolution of < 5 seconds may potential be sufficient to evaluate the key meal microstructure parameters. Next, the dissertation explores a wearable camera to capture images of eating episodes. The wearable camera captures both food and non-food images. The chapter-4 of dissertation introduces a methodology for clustering of food images into food and non-food groups based on histogram matching, without recognizing food items on the image.

Incorporating sensor with wearable camera, the dissertation then proposes a novel wearable sensor system (AIM-2) that can detect eating instances, provide accurate chew counts during food intake events and automatically capture passive images. To obtain chew counts from eating events, a novel chewing sensor is introduced. The work further demonstrates the number of images can be reduced significantly by analyzing the images during food intake event only. The method also addresses potential possibility of alleviating some of the privacy concerns of wearable cameras. The results suggest that accurate monitoring of food intake can be achieved in both controlled laboratory conditions as well as unrestricted, daily free-living conditions. Moreover, the sensor system can be potentially used for EI estimation in free-living environment.

By utilizing the AIM-2 device, the dissertation finally presents the preliminary results of automatic EI estimation for humans in free-living environment. A human subject study was conducted on healthy participants both in the laboratory and free-living conditions to test the accuracy of AIM-2. The major contribution of the work compared to existing EI estimation techniques is to utilize both sensors and images to estimate EI. To assist reaching to the goal, a software is also designed for reviewing images, sensor-detected eating events and estimate ingested mass and EI. The preliminary results suggested that AIM-2 can be potentially used for EI estimation in free-living environment.

7.2 Limitations and Future works

The primary limitation of the work presented in the dissertation is the small sample size of 4 participants. However, the presented results are promising. In future, long-term studies with large population are needed to fully develop the feasibility of proposed approach. Further works can also be done on the validation of sensor system and the proposed methods in a weeklong free living conditions against the doubly labeled water method.

Another limitation is that the sensor system primarily uses chewing information to detect food intake events. Therefore, if the user performs an eating activity which includes significant amount of chewing but no swallowing (e.g. chewing gum), then the proposed system will output false positives. In addition, the drinking episode which does not involve chewing, may not get detected by the system. Future works are needed to be done on liquid intake detection. The jaw movements during consumption of liquids can be potentially explored to detect liquid intake.

Currently, the device is capturing continuous image in 15 secs interval. Future works can be done on the device where the camera will be triggered based on the sensor signals and capture images only during food intake events. The food images are now being reviewed manually by nutritionist. In future, deep learning based computer vision techniques could potentially be used for recognition of the type of food consumed.

The EI estimation models can be further improved by utilizing large number of eating episode information from more participants. Further studies are also needed to explore the long-term use of the device and identify issues related to user comfort and compliance.

Going forward, a method for real-time food intake detection and immediate feedback to the users can be explored. Currently, the signal processing of the sensor data coming from AIM-2 was done off-line. However, for real-time food intake detection, all the processing needs to be moved to the AIM-2 framework. Future works can be done to develop these new methodologies, for making the AIM-2 more robust and reliable wearable system for dietary assessment.

REFERENCES

[1] C. M. Champagne *et al.*, “Energy Intake and Energy Expenditure: A Controlled Study Comparing Dietitians and Non-dietitians,” *J Am Diet Assoc*, vol. 102, no. 10, pp. 1428–1432, 2002.

[2] N. Day, N. McKeown, M. Wong, A. Welch, and S. Bingham, “Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium,” *International Journal of Epidemiology*, vol. 30, no. 2, pp. 309–317, Apr. 2001.

[3] L. S. Muhlheim, D. B. Allison, S. Heshka, and S. B. Heymsfield, “Do unsuccessful dieters intentionally underreport food intake?,” *Int J Eat Disord*, vol. 24, no. 3, pp. 259–266, Nov. 1998.

[4] F. E. Thompson and A. F. Subar, “Dietary assessment methodology,” in *Nutrition in the Prevention and Treatment of Disease*, 2nd ed., Academic Press, San Diego, CA, 2008.

[5] A. H. C. Goris and K. R. Westerterp, “Underreporting of Habitual Food Intake Is Explained by Undereating in Highly Motivated Lean Women,” *J. Nutr.*, vol. 129, no. 4, pp. 878–882, Apr. 1999.

[6] D.-H. Wang, M. Kogashiwa, and S. Kira, “Development of a new instrument for evaluating individuals’ dietary intakes,” *J Am Diet Assoc*, vol. 106, no. 10, pp. 1588–1593, Oct. 2006.

[7] D.-H. Wang, M. Kogashiwa, S. Ohta, and S. Kira, “Validity and reliability of a dietary assessment method: the application of a digital camera with a mobile phone card attachment,” *J. Nutr. Sci. Vitaminol.*, vol. 48, no. 6, pp. 498–504, Dec. 2002.

[8] C. K. Martin, T. Nicklas, B. Gunturk, J. B. Correa, H. R. Allen, and C. Champagne, “Measuring food intake with digital photography,” *J Hum Nutr Diet*, vol. 27 Suppl 1, pp. 72–81, Jan. 2014.

[9] A. Doulah *et al.*, “Meal Microstructure Characterization from Sensor-Based Food Intake Detection,” *Front. Nutr.*, vol. 4, 2017.

[10] 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, Jun. 2014.

[11] J. M. Fontana *et al.*, “Energy intake estimation from counts of chews and swallows,” *Appetite*, vol. 85, pp. 14–21, Feb. 2015.

APPENDIX A IRB APPROVAL

THE UNIVERSITY OF ALABAMA | *Office of the Vice President for* Research & Economic Development
Office for Research Compliance

August 22, 2017

Edward Sazonov, Ph.D.
Professor
Electrical & Computer Engineering
College of Engineering
The University of Alabama
Box 870286

Re: IRB Protocol # 17-005-ME
"Evaluation of a Wearable Camera for Food Intake and Ingestive Behavior Monitoring"

Dr. Sazonov:

The University of Alabama IRB has received the revisions requested by the full board on 8/16/17. The board has reviewed the revisions and your protocol is now approved for a one-year period. Please be advised that your protocol will expire one year from the date of approval, 8/10/17.

If your research will continue beyond this date, complete the Renewal Application Form. If you need to modify the study, please submit the Modification of An Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, please complete the Request for Study Closure Form.

Should you need to submit any further correspondence regarding this proposal, please include the assigned IRB application number. Please use reproductions of the IRB approved stamped consent form to obtain consent from your participants.

Good luck with your research.

Sincerely,

[redacted]

J. Grier Stewart, MD, FACP
Medical IRB Chair

358 Rose Administration Building | Box 870127 | Tuscaloosa, AL 35487-0127
205-348-8461 | Fax 205-348-7189 | Toll Free 1-877-820-3066

APPENDIX B PUBLICATIONS

Published:

Journals and Conferences:

A. Doulah, Farooq M, Yang X, Parton J, McCrory MA, Higgins JA and Sazonov E (2017) **Meal Microstructure Characterization from Sensor-Based Food Intake Detection.** Front. Nutr. 4:31. doi: 10.3389/fnut.2017.00031

A. Doulah, E. Sazonov (2017) **Clustering of Food Intake Images into Food and Non-food Categories.** Springer International Publishing AG; Bioinformatics and Biomedical Engineering, IWBBIO 2017, Part I, LNBI 10208, pp. 454 - 463, 2017. DOI: 10.1007/978-3-319-56148-6_40.

A. Doulah, X. Yang, J. Parton, J. Higgins, M. McCrory, and E. Sazonov, "**The importance of field experiments in testing of sensors for dietary assessment and eating behavior monitoring,"** presented at the The 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'18), Honolulu, HI, 2018.

A. Doulah; X. Shen; E. Sazonov, **A method for early detection of the initiation of sit-to-stand posture transitions.** Physiol. Meas. 2016, 37, 515, doi:10.1088/0967-3334/37/4/515.

A. Doulah; X. Shen; E. Sazonov, **Early detection of the initiation of sit-to-stand posture transitions using orthosis-mounted sensors.** Sensors 2017, 17(12), 2712; doi:10.3390/s17122712

Abstract:

M Farooq, **A. Doulah**, Yang X, Roopwattie Jeanot, Parton J, McCrory MA, Higgins JA and Sazonov E (2016) **Inter-and Intra-rater Reliability of Video Based Annotation of Eating Episodes,** The obesity society 2016.

Submitted:

X. Yang, **A. Doulah**, M Farooq, Parton J, McCrory MA, Higgins JA and Sazonov E , "**Statistical models for meal level estimation of mass and energy intake using features derived from video observation and a chewing sensor**" in preparation for Scientific Reports, Nature

M Farooq, **A. Doulah**, McCrory MA, and Sazonov E , "**A multi-day study on the use of video observations for the validation of the AIM in unconstrained environment**" in preparation.

In preparation:

A. Doulah, J. Higgins, M. McCrory, and E. Sazonov, "**A Systematic Review of Technology-driven Methodologies for Estimation of Energy Intake**" to be submitted to IEEE Access.

A. Doulah, and E. Sazonov, **"Automatic Ingestion Monitor Version 2" – a Novel Wearable Device for Automatic Food Intake Detection, Chew Count Estimation and Passive Capture of Food Images"** to be submitted to IEEE JBHI.

A. Doulah, J. Parton, J. Higgins, M. McCrory, and E. Sazonov, **"a Novel Method of Estimating Energy Intake in Laboratory and Free-Living Conditions Using a Wearable Sensor and Food Imagery"** to be submitted to IEEE JBHI.