Automatically Measuring Individual Consumption Events During Natural Eating Using a Table Embedded Scale

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Electrical Engineering

by
Ryan Stephen Mattfeld
August 2014

Accepted by:
Dr. Adam W. Hoover, Chair
Dr. Eric R. Muth
Dr. Richard E. Groff

Abstract

This thesis is motivated to improve the tools available for tracking energy intake. The goal of this work is to develop a table-embedded scale capable of measuring the weight of individual consumption events during unrestricted eating. The method was tested on a dataset gathered from 276 subjects eating 518 courses consisting of 22,383 marked individual consumption events in a cafeteria environment. Approximately 30% of the consumption events can be detected and weighed. The remaining 70% of the events occur without participants interacting with the scale or when noisy interactions with the scale prevent weight measurement.

The relationship between bite size in grams and BMI was analyzed across all 7,501 measurable bites found using ground truth bite times and across 10,240 automatically detected events. The relationship between bite weight and BMI was found to be 0.28 g/BMI. Without using ground truth bite times, a relationship of 0.21 g/BMI was found for automatically detected events. The trend is diminished but still clearly present even with the presence of false alarms. In addition, when each bin is broken into quartiles, the results indicate that g/bite vs BMI is nearly constant for the smallest 25% of bites, but increases in each quartile. When the largest 25% of bites are analyzed, a relationship of 0.58 g/BMI is found. While these amounts may seem small, the cumulative effect over hundreds or thousands of bites suggests new opportunities for behavior change based on bite size.

Acknowledgments

I would like to thank my adviser, Dr. Hoover, for his countless hours of advice, input, and guidance throughout the entire process of attaining my Master's degree. I would also like to sincerely thank Dr. Muth and Dr. Groff for the time and effort they spent when reviewing this thesis. Finally, I would like to thank my friends and family for their support, especially my parents, Ann and Steve Mattfeld, and my wife, Valerie Mattfeld, who have been instrumental in my pursuit of a Master's degree.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

In the study and treatment of overweight and obese individuals, two main measurements are energy intake (EI) and energy expenditure (EE). EI is a measure of the quantity of energy absorbed through food and drink consumption, and EE measures the energy expended through homeostasis (body maintenance) plus physical activities. EI is determined by food intake, whereas EE is determined by the resting metabolic rate, the thermic effect of food, and the thermic effect of physical activity. Both EI and EE can be measured in the same units, calories or Joules. This following work is concerned with improving methods of recording and measuring EI.

1.2 Energy Intake (EI) and Energy Expenditure (EE)

Measuring EI can be achieved by recording an individual's food and drink intake and gathering energy and nutrient information from the data [1]. To this point in time, the primary means of recording EI are through the use of manual recording methods [2, 3, 4]. The methods available include questionnaires about frequency of food consumption, completing a diary about the quantity of food eaten, and recalling the amount of food eaten over a set period of time. Measuring the frequency of food consumption is accomplished using a food-frequency questionnaire designed to asses patterns in subjects' diets by recording the frequency with which various foods are consumed [5]. Keeping a diary of food consumed is accomplished through the use of self-report food records like the 7-day diet diary [5]. Finally, recording food consumption over time can be accomplished using a 24-hour recall, in which a directed interview is used to quantify the amount of food and beverage consumption that occurred through a single day [6]. More recently, computer based methods have been produced that are capable of scanning bar codes or verbal food descriptions to record EI over the course of a day [7]. Finally, there are additional methods which tie together EI and EE by recording both daily food intake as well as daily exercise to estimate weight changes and the nutritional value of items consumed [8, 9].

The accuracy of the self-report methods discussed are typically determined by comparing EI measures against EE measures for the same time period. Combined with weight gains or losses, the difference between EE and EI is determined. If weight is to remain constant, then EE should match EI. The most accurate method of determining EE is by using a calorimetry chamber, in which subjects live in a fully controlled environment. In the calorimetry chamber, a subject's EI can also be measured by using a technique called bomb calorimetry for all foods provided to the subject. Calorimetry chambers are very expensive and also extremely impractical for widespread use over time [10]. The most accurate technique currently available in a more realistic setting for measuring EE is the use of doubly labeled water [11]. Over a set period of time, frequently one week [12], the average metabolic rate of a subject is measured by using water in which the hydrogen and oxygen are replaced with deuterium and oxygen-18. By replacing these elements in the water consumed, the decay of the isotopes consumed is measured through use of daily saliva, blood, and urine samples. This technique, while applicable in regular daily life, is still very expensive and requires technical expertise to take samples [13]. Even with the drawbacks, this technique is the standard for EI methods to set their benchmarks. There is significant evidence that establishes that current methods of measuring EI that rely on information provided by self-report methods yield inconsistent and unreliable results. Food diaries, recall methods, and food-frequency questionnaires all rely upon self-report [6]. It has been shown that people using these self-report techniques display a tendency to under report their EI by 10% to 50% [14]. The tendency to under report could be caused by a number of underlying factors [15, 16, 17, 18, 19]. When using recall methods, people forget the quantity and type of food consumed. In addition, food-frequency questionnaires require large lists of items, and their effectiveness is determined by how much food lists overlap with people's eating habits. Food records require the person recording the information to be accurate at self-estimation, but people have been shown to exhibit tendencies to overestimate larger portions and underestimate smaller portions [6]. There is a clear need for improving methods of measuring EI for practical applications. Tools that do not rely on self-report would improve EI measurement because automatic tools using technology would allow EI measurement to be objective, hopefully eliminating the under-reporting bias common to self-report methods. In addition, automatic tools could provide real-time feedback during eating to provide methods that may be useful for behavior change. Tools that are more accurate, less expensive, and put less of a burden on the people using the techniques would improve EI measurement [20, 21].

1.3 Obesity

Obesity is a medical condition characterized by the excessive accumulation and storage of fat in the body. Obesity can be defined in a number of ways, but the most widely adopted method of defining obesity is through body mass index (BMI). A person's BMI is measured by dividing weight in kilograms by height in meters, squared. The World Health Organization (WHO) [22] defines the boundaries for overweight, obesity, and morbid obesity by BMIs of 25, 30, and 40, respectively.

According to the World Health Organization, obesity has been rapidly increasing globally, more than doubling since 1980. In addition, obesity is identified to be the leading cause of death in a number of countries whose combined population includes 65% of the total global population [22]. This widespread increase in obesity is damaging the health of the worldwide population and is considered to be a global epidemic [23, 24]. The presence of obesity has been correlated to a variety of diseases including gallbladder disease, high cholesterol levels, type 2 diabetes mellitus, coronary heart disease, osteoarthritis, high blood pressure, cardiovascular disease, many types of cancer, sleep apnea, asthma, infertility, musculoskeletal disorders, and depression [18, 25, 26, 27].

The United States in particular has seen a rapid increase in obesity in children, adolescents, and adults. In 2009 and 2010, the National Health Nutrition Examination Surveys found that 35% of the adult population and 17% of adolescents and children in the United States were obese [28]. This indicates that roughly 78 million American adults and 12.5 million adolescents and children are obese [29]. In response to the rapidly increasing obesity crisis, the U.S. Department of Health and Human Services set goals to reduce the prevalence of obesity among adults and adolescents, deemed the "Healthy People 2012" initiative. The goals of this initiative were to reduce obesity among American adults to 15% and adolescents and children to 5% [30]. The goals of the initiative were not met. Instead, during the "Healthy People 2012" initiative, the prevalence of obesity in children actually increased from 11% to 18%. Based on the obesity trends seen in 2010, the combined medical expenditures caused by obesity over the next 2 decades would reach $549.5 billion dollars. Over that time period, there would be a 33% increase in the prevalence of obesity and a 130% increase in the prevalence of severe obesity. The prevalence of obesity in the year 2030 would reach 42% and severe obesity would reach 11% [31].

Surgery, pharmacology, and behavioral changes are the primary methods available to combat obesity. Surgical procedures include gastric bypass, gastric banding, vertical banded gastroplasty, and laparoscopic cholecystectomy [32, 33]. These procedures have all been shown to produce significant weight loss, but are typically only considered for morbidly obese individuals who have failed to respond to alternative methods for weight loss but are good candidates from a psychological standpoint. Pharmacological methods are those methods that use medication to reduce appetite, decrease nutrient absorption, or increase thermogenesis (EE) [34]. These medications have been shown to have side effects on the human body such as cataracts, neuropathy, hemorrhagic strokes, and valvular heart disease. Behavioral weight loss methods aim to increase EE and decrease EI through changes in a person's behavior [18, 19].

The study of diet and obesity has been an area of increased interest in engineering because of the need for more effective tools to improve EE and EI measurement. Dietetics research is heavily dependent on dietary assessment methodology [20]. The lack of accuracy in this area is a major problem for dietitians because nutrition care plans are dependent on the accuracy of the feedback given. Some improvement has been seen with the introduction of computerized questionnaires on mobile devices and the Internet, but there is still a high level of inaccuracy. The ideal solution would be to create systems that can use sensors to automatically monitor EI and EE.

1.4 EI Measurement Using Scales

The use of scales to measure and record EI has been tested in laboratory settings. A device called the Universal Eating Monitor (UEM) has been developed in an attempt to provide a tool for measuring EI through use of a scale. The UEM was developed to permit continuous weighing of a subject's plate or alternative food reservoirs with a concealed electronic balance [35]. The UEM can record the amount consumed throughout a single-course meal consisting of a relatively homogeneous mixture of foods. The UEM is capable of determining total weight intake over the course of a meal, the duration of the meal, the initial rate of intake, and the deceleration of intake of humans consuming a solid or liquid version of the same food.

The UEM has allowed researchers to detect and measure eating rate and change in eating rate over time. These statistics have been gathered through use of a cumulative food intake curve [36]. The cumulative food intake curve is calculated by treating the first temporal third of the meal as linear, while using a quadratic or linear equation for the other two-thirds of the meal [37, 38, 39, 40]. By fitting curves to the scale data recorded across a meal, the eating rate of the meal can be measured, but the individual events that occur throughout the meal are lost.

An additional study using weight data as the primary resource have focused on the eating rate (ER) of individuals [41]. This study used standardized meals for breakfast and lunch and measured eating rate by using both grams and kilocalories consumed divided by the unit of time in minutes. The study concluded that self-reported eating rates correlated well with measured eating rates. Slower eating rates have been shown to lead to lower EI and greater weight loss [42, 43].

The consistency of laboratory measures of food intake are another issue that has been a topic of study over time. One study, using the UEM in a laboratory setting determined that measures of food intake were stable for men and women even when consuming one of three different types of sandwiches over time [44].

The results of these studies show that EI and ER can be measured through use of a scale or UEM in laboratory settings, including a controlled environment and pre-determined food types. The measures presented find results based on a starting and ending scale weight, and measure eating rate by dividing the weight consumed by time. None of the methods are used in a real-world, uncontrolled scenario, and none of the methods measured individual consumption events.

Using a scale to measure individual consumption events across a meal is a challenging process. The data provided by a scale across a meal varies widely as can be seen in Figure 1.1. The data shown is recorded across a meal over 13 minutes in length. The weight recorded by the scale clearly decreases over time, but individual bites or drinks can be challenging to identify. There is significant noise present as the user presses down on the scale to prepare their food for consumption or when they lift heavy objects, such as a glass of water, off the scale.

Based on the features present in the full meal scale curve, researchers have sought to analyze meals by fitting curves to the scale data. Typically, these curves are fit using linear or quadratic

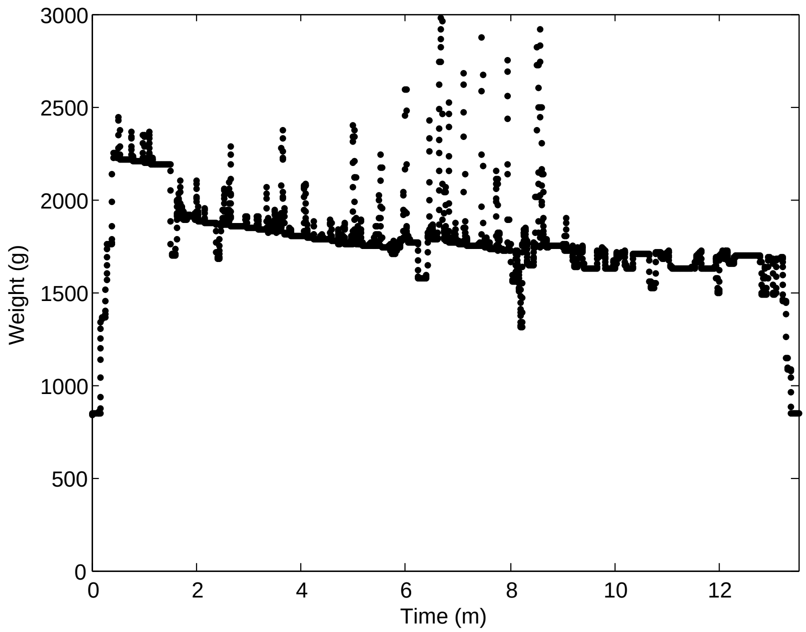


Figure 1.1: Scale data recorded across an entire meal covering over 13 minutes. In general, it can be seen that the weight decreases throughout the meal as food is consumed. But it can also be seen that there are many spikes as drink containers or food masses are removed and later returned to the table, or as the user presses down on the table while retrieving food, or as the user does other activities such as manipulating food or cleaning.

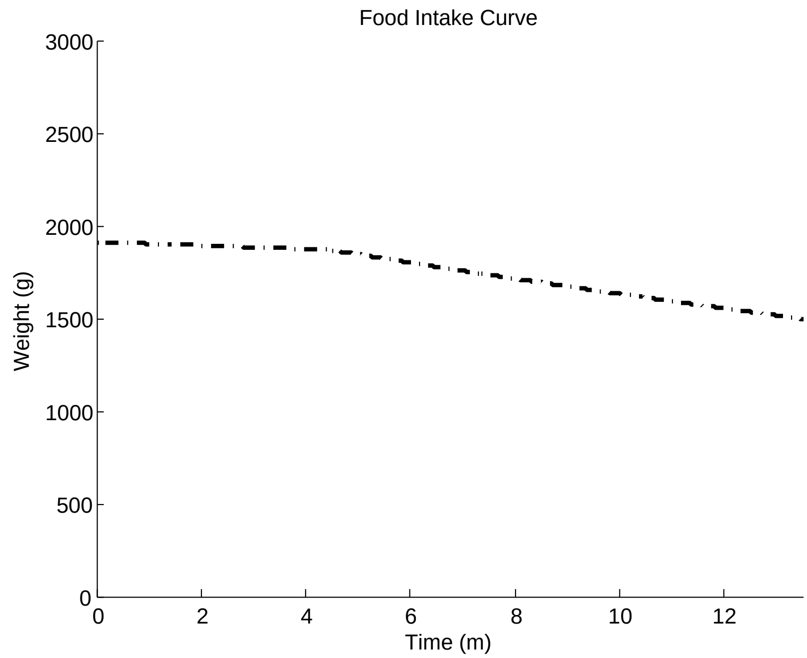


Figure 1.2: Two lines of best fit, one fit to the first third of the meal, and one fit to the final two-thirds, are used to estimate the rate of consumption over the course of a meal.

equations to find a line that best represents the scale data across time. Using lines of best fit to analyze food intake rates for the scale data shown in Figure 1.1 results in a line of best fit as shown in Figure 1.2. This line was found by fitting one trend line to the first temporal third of the scale data and another to the second two temporal thirds. The lines show the rate that the food is consumed over time, but using line fits of this type filters out all of the individual events occurring throughout the meal.

1.5 Challenges in Detecting Individual Consumption Events

The goal of using a scale to automatically detect, measure, and record energy intake (EI) events in a real-world, restriction-free environment, requires that the scale data is broken down in a way that allows analysis of events that result in EI. In order to break down the data in this way, energy intake events should be defined. In this thesis, the term "bite" refers to ingestion resulting from the consumption of any form of sustenance, whether food or drink. A "food bite" is ingestion through the consumption of food, and a "drink bite" is ingestion through the consumption of a liquid beverage. These EI activities can be measured by determining the weight change, in grams, detected through use of a scale. In this thesis, a "scale event" is a set of scale activities characterized by a region of stability, followed by a region of instability and an additional region of stability. A scale event as described above allows for the measurement of the weight change that occurs when the scale is interacted with, and is used as the basic unit for identifying and measuring the weight of individual bites throughout a meal.

Over the course of a meal, individual bites can appear in a variety of forms. A food bite consisting of a single piece of food that is never returned to the scale could cause scale events such as those shown in Figure 1.3. The figure shows three consecutive food bites, where the moment of ingestion is marked by a vertical line and occurs immediately after a period of scale instability. The period of instability that occurs during the scale event is the time frame during which the food to be consumed is picked up and can be used to determine the weight of food consumed. The first bite in Figure 1.3 shows a spike in the weight recorded by the scale when the participant picks up a bite of food. This spike is followed by a small bump in the scale reading as the motion to pick up the bite of food is completed. The difference in scale weight can be found by subtracting the scale weight from the period of scale stability following the bite from the scale weight from the region of stability preceding the bite.For the first bite shown, the weight of food consumed is 8.91 grams. The second bite in the figure again shows a spike in scale weight as the food is picked up, but no additional noise follows the spike. The third and final bite in the figure shows the distinctive spike in scale data as food is picked up again, and is followed by more activity as the food is moved on the plate before being picked up and consumed.

In addition to the food bites described above, there are also some food bites that occur when the participant picks up a food mass, such as a sandwich or piece of pizza, takes a bite, then returns the remainder of the food mass to the tray. This type of bite presents a number of unique challenges. The first challenge is determining the scale events that indicate when the food mass leaves the tray and when it is returned. In addition, using the scale data alone, it is impossible to know how many bites were taken out of the food mass during the time between when the food mass was picked up and when it was returned to the tray. Since the food mass is held above the tray for an extended period of time, no activity is detected on the scale. In order for this type of food bite to be measurable, the participant must pick up the food mass, take a single bite, and return the food

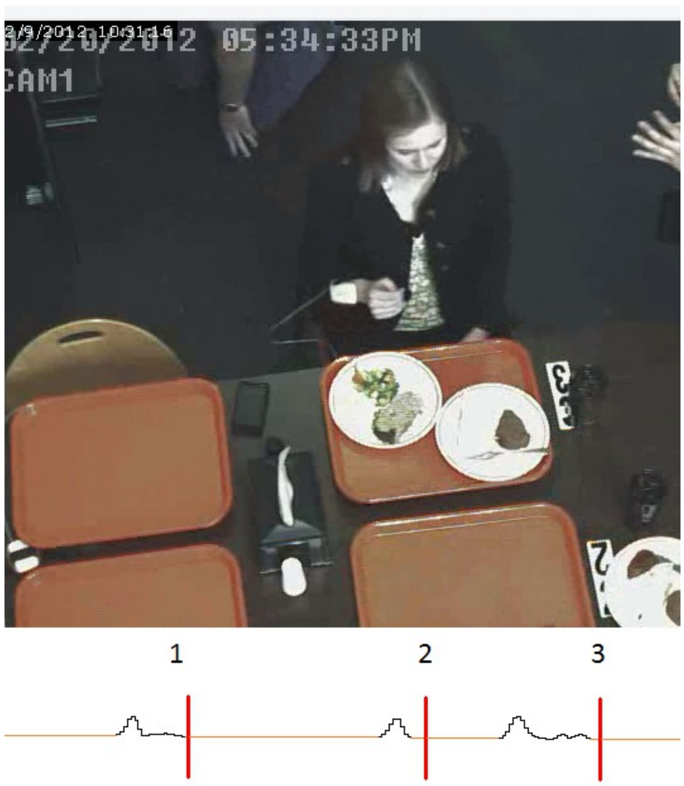


Figure 1.3: Example of scale data recorded as three single bites of food are picked up from the tray and consumed.

mass to the tray. Some sample scale data corresponding to this activity appears as shown in Figure 1.4. As can be seen in the figure, there is a scale event showing a clear decrease in weight as the food mass is picked up from the tray, followed by an extended period of time in which the scale is stable and the bite is taken, and finally a second scale event showing a clear increase in weight when the food mass is returned to the tray. The moment of ingestion occurs at the point in time indicated by the vertical line. This pattern shows two scale events that can be used to detect and measure the weight of a bite in which both the removal of the food mass from the tray and its return are clearly defined scale events. The weight of the bite can be calculated by finding the weight before the food mass is picked up and subtracting the weight recorded after the remainder of the food mass is returned to the tray. In addition, it is important that only one bite is taken because if multiple bites were taken, as is a common occurrence in this type of bite, the weight of each individual bite cannot be determined.

A third type of bite is one in which a drink is taken and can be detected through use of two scale events: the first when the drink is picked up from the tray and the second when the drink is returned to the tray after some of the beverage is consumed. This set of two scale events is used to detect and measure a drink bite, and is characterized by a large decrease in scale weight as the drink is picked up, a period of stability as the drink is taken, and a large increase in scale weight as the drink is returned to the tray. An example of a drink bite is shown in Figure 1.5. This example is ideal because the weight decrease as the drink is picked up is clear and distinctive, there is a clear stable period as the drink is taken (the moment of ingestion is marked by a vertical line), and a clear and distinctive weight increase as the drink is returned to the tray. Drink bites proved to provide the most consistently distinctive and identifiable pattern, though participants often took multiple drinks within one time period, making it impossible to individually measure each drink consumed.

There are many examples of ambiguous scale events that are difficult to interpret. There are several instances in which the scale events indicate actions that appear to be food bites even though no food is consumed. There are a number of examples of challenging single bites shown in Figure 1.6. For all of the bites in this example, there is no clear spike in the scale when food is picked up as there is in the single bites shown previously in Figure 1.3. In addition, there are scale events that appear to indicate that food bites are being taken but no bite occurs. Figure 1.6 is labeled in order to more clearly indicate points of interest. In the figure, labels 1, 3, 4, and 6 show the moments of ingestion of four bites, indicated by vertical lines, that occur after segments of scale data in which

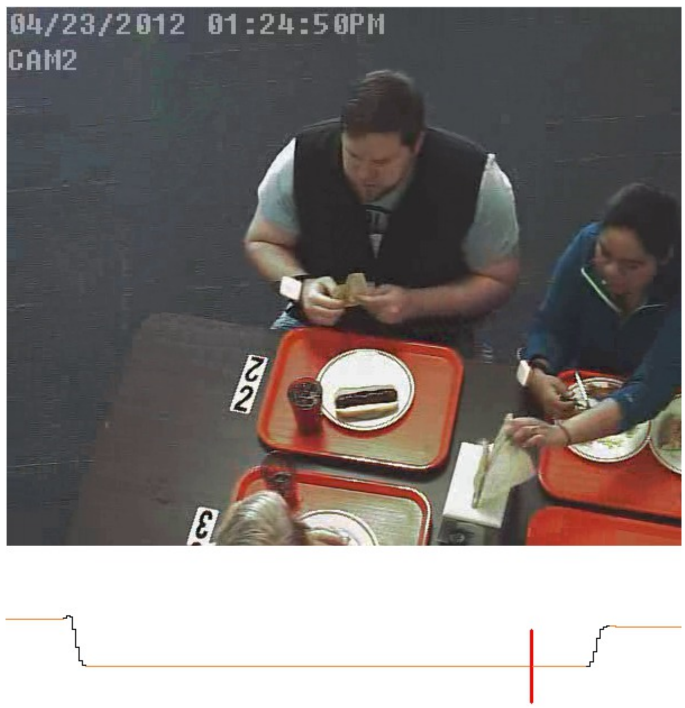


Figure 1.4: Example of scale data recorded as a food mass (hot dog in this case) is picked up, a bite is consumed, and the remainder of the food mass is returned to the scale. The vertical line indicates the moment of food ingestion.

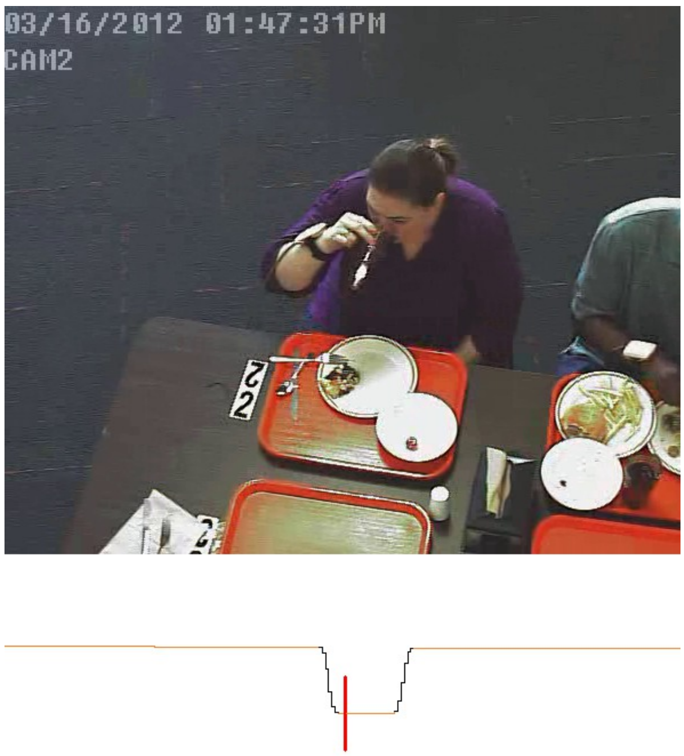


Figure 1.5: Example of scale data recorded as a drink is picked up, some of the drink is consumed, then the cup is returned to the scale.

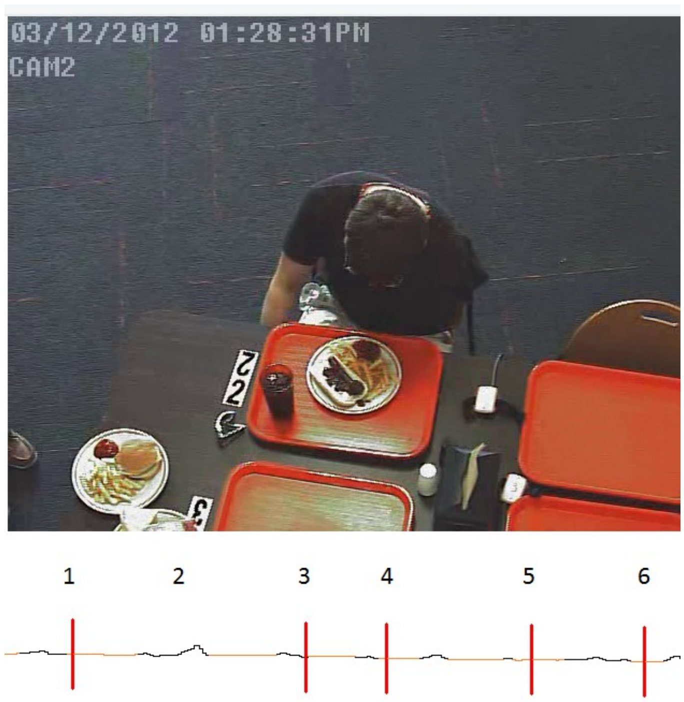


Figure 1.6: This shows a segment of scale data in which there are a number of single bites that are less distinctive. In addition, there is an event (2) that appears to be a single bite but during which no bite is taken.

there is no distinctive spike in the scale reading as the food is picked up. Label 2 shows a segment of data similar to data that typically precedes a bite, but after which, no bite is taken. This is an example of a scale event that may be mistaken for a bite. At label 5, a bite is taken, but the weight is not able to be determined because there are two scale events before the bite. In this case, it is impossible to determine, from the scale data alone, whether the weight of the bite can be determined from the scale event immediately preceding the bite, from the sum of the two preceding scale events, or from the second preceding scale event alone.

One challenge that appears when participants are allowed to eat in a real-word, restriction-free environment is that participants will often pick up food but become distracted before consuming the food. After being distracted, perhaps due to communicating with a friend, the participant may return the food to the tray without consuming it or they may pick up more food before taking a bite. In Figure 1.7, the first scale event present in the data consists of a period of stability before food is picked up, a period of instability as the participant picks up food, and a second period of stability as the participant talks to a friend. This scale event cannot be used to measure the weight of food consumed because the participant proceeds to pick up more food before taking a bite. The problem is further exacerbated because another bite of food is picked up and consumed during the unstable period from the second scale event. The end result is that neither the first nor second bite taken in the figure can have their individual weights accurately recorded.

Another common cause of immeasurable bites occurs when the participant eats too quickly for the scale to provide a stable weight before and after the bite. When participants eat very quickly, the resulting scale data can appear as in Figure 1.8. Throughout all of the scale data shown, 8 bites are taken at the moments of ingestion indicated by vertical lines, but no individual bites can be measured because each bite cannot be individually tied to a scale event. All of the noise is present because the participant continues to pick up or manipulate their food frequently, and the scale is not allowed to stabilize.

A similar problem occurs when participants rest their arm on the tray, leading to a large number of bites having immeasurable bite weights. In Figure 1.9, during the first scale event, the participant is resting their arm on the tray, causing too much noise on the scale for the weights of the first 5 bites to be read. The figure also shows a second scale event which appears to indicate a single bite being taken, but a bite only occurs before the unstable region of data, making that bite weight immeasurable as well. Finally, one accurate reading of a single bite occurs during the third scale event shown. This figure shows both the challenges presented by participants resting their arm on the scale as well as the difficulty of identifying single bites from just scale data.

1.6 Novelty of Work

This thesis proposes the development of an algorithm which can be applied to create a tool capable of measuring the weight of individual bites taken by subjects during unrestricted eating. Previous research in this area typically restricts scale measurements to a laboratory setting with

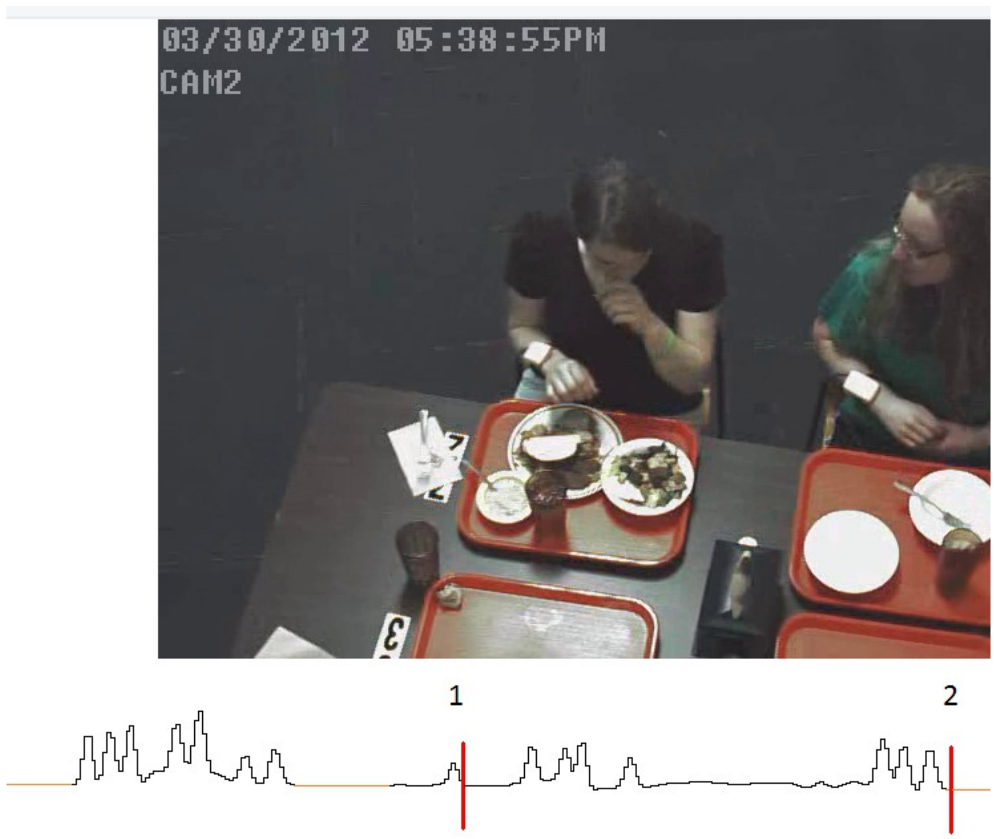


Figure 1.7: An example of a challenging segment of data in which two bites occur, but neither of the bites' weights can be determined accurately.

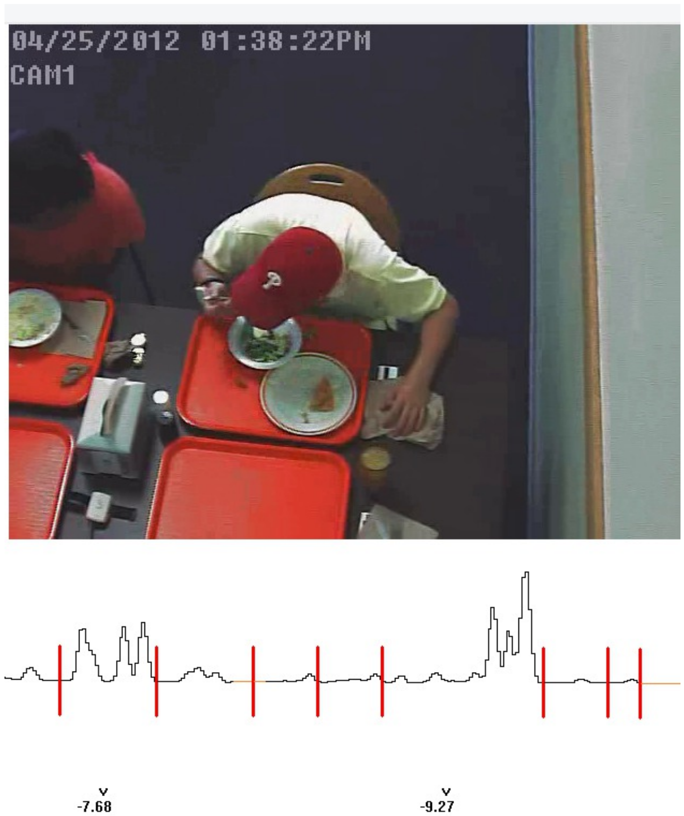


Figure 1.8: An example of a challenging segment of data in which the participant is taking quick bites and manipulating their food frequently, resulting in 8 bites whose weights cannot be determined due to the scale never being allowed to stabilize.

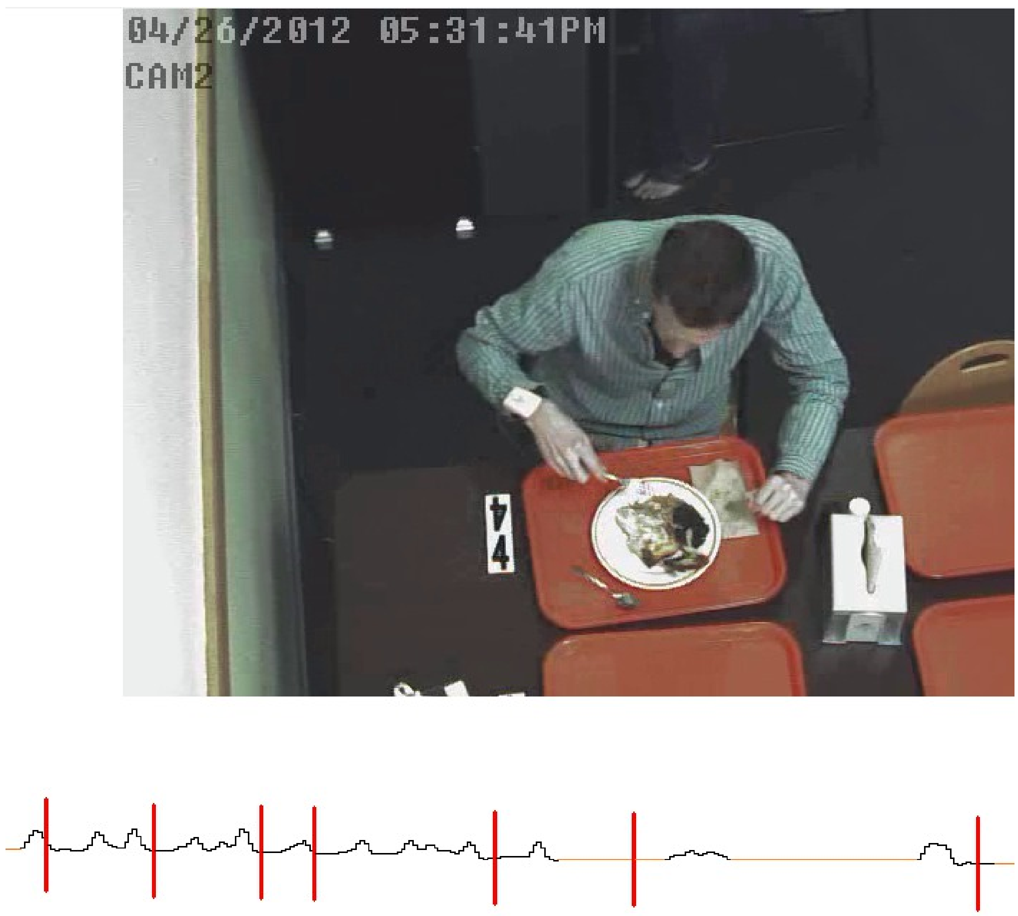


Figure 1.9: An example of a challenging segment of data in which the participant rests their arm on the tray, making 5 bite weights immeasurable. In addition, there is an example of a scale event that does not correlate to a bite and a scale event that does correlate to a bite.

a limited variety of foods available for consumption. Prior research has estimated the weight of individual bites by using an external observer to manually record the number of bites taken, either through video feed or a two-way mirror [45, 46, 35]. Using the manually observed bite count, scale weights from before and after the meal, and the duration of the meal, the average weight of each bite and the eating rate across the entire meal could be estimated. The methods used in this thesis are capable of detecting and measuring the weight of individual bite events and could potentially be used to provide real-time feedback to users. In addition, the algorithm is tested across a large dataset of 276 participants eating 518 courses. Results are reported evaluating the correlation of BMI to individually detected consumption events.

Chapter 2

Methods

2.1 Introduction

In this chapter, methods for measuring the mass of food consumption at a bite level throughout a meal are presented. There are three main goals to be accomplished: 1) identify the mass of bites consumed during a meal using data provided by a scale, 2) evaluate relationships between the mass per bite and the BMI of the individual, and 3) build an automated bite detection algorithm to be applied to data collected through use of a scale.

2.2 Scale Hardware

For this study, a scale was mounted on an instrumented table used to collect weight data. The scale used in this data collection was an OHAUS Scout Pro SP4001. The SP4001 is readable to 0.1 grams with a reported standard deviation of 0.1 grams. It supports a maximum capacity of 4 kg and has a reported stabilization time of 3 seconds. The scale is digitally calibrated through use of a keypad and has a net weight of 0.8 kg.

In order to verify the statistics, specifically the standard deviation of the scale, we performed scale noise tests. In order to experimentally determine the scale noise, we recorded scale readings over a time period of two minutes on three separate occasions in an environment simulating the conditions present in a cafeteria. In order to simulate a cafeteria environment, we loaded the scale with a tray containing a typical meal's worth of mass and simulated the noisy environment

Figure 2.1: An image of the scale recessed under the table.

typical of the cafeteria by walking around the table and occasionally bumping the table. From these simulations, we determined that the scale produced noise with one standard deviation of 0.29 grams.

The scale was mounted on a platform recessed into the table top. This setup resulted in the scale pan resting roughly one centimeter above the surrounding tabletop as seen in Figure 2.1. Each scale pan was attached to a food tray through use of Velcro strips. This system created some difficulties when measuring scale data due to the small clearance between the scale and the table, combined with the pliable plastic food tray. These two parts of the system allowed the food tray to bend enough to contact the table top if heavy items were placed near the edges of the tray. This phenomenon occurred when participants moved their drinks near the tray's edge during a meal and when participants shifted plates around near the edges of the tray. Consequently, when the tray would make contact with the table top, the resulting scale measurement would be inaccurate, due to some of the weight present being borne by the table rather than the scale.

2.3 Data Collected

The data used for this study was recorded in a large data collection effort supported by the National Institutes of Health via grant 1R41DK091141-A1.

The data was used to test the accuracy of a wrist-worn device with a variety of food choices, utensils, containers, and subjects inside a cafeteria in a previous study. The Harcombe Dining Hall of Clemson University was the facilitator for data collection. The facility seats up to 800 guests and provides a wide range of foods and beverages, allowing people to build their own meal. For example, some of the choices include omelets, sandwiches, pizza, pasta, fruits and vegetables, meat cuts, deserts, juices, milk, sodas, teas and coffee. The foods are served in a wide variety of containers, including plates, bowls, wraps, pouches, trays, cartons, cups, and glasses, and they are consumed using a variety of utensils including forks, knives, spoons, and fingers.

The full data set is composed of 276 subjects, including 131 males and 145 females subjects ranging age from 18 to 75. Body mass index (BMI) ranged from 17.4 kg/m^2 to 46.2 kg/m^2. Ethnicity is predominantly Caucasian (192), but also includes African-American (27), American Indian or Alaska Native (2), Asian or Pacific Islander (29), Hispanic (11), and Other (15).

In the dining hall, an instrumented table was prepared to record data simultaneously from four participants. Four digital cameras in the ceiling (approximately 5 meters height) were used to record the participant's mouth, torso, and tray while they consumed their meal. Also, a custom wrist-worn device was used to record the motion in the participant's wrist during his or her meal consumption. Each participant had his device wired to a laptop where data was stored. Anywhere from two to four participants were seated at the table at a time for each meal. A scale was located under the subject's tray to monitor food weight while eating. The scale data and wrist motion data were sampled at a rate of 15 hz. Figure 2.2 shows a picture of the instrumented table.

2.4 Ground Truth Bites

There were no restrictions on foods or eating style during data collection. Participants were allowed to eat as normally as possible, including natural movements unrelated to eating (e.g. conversations with people, using a phone, gesturing, using a napkin, etc.). Subjects were free to build their own meal, as well as drink any type of beverage using the container of their choice. Figure 2.3 shows an image of four subjects eating at the instrumented table.

Figure 2.2: An image of the instrumented table.

Figure 2.3: An image of the instrumented table with participants ready to eat.

After the initial data was collected, the entire data set was manually reviewed and moments in time in which bites were taken were marked. The manually marked bite times included additional data about the bite being taken including the container used, hand used, type of food consumed, and utensil used. The data used to create these ground truth marks were created by having reviewers manually watch the synchronized video of the participants eating and mark the times at which the bites were taken. The bites were marked at each moment in time in which the participant ingested either food or drink. The only difference between food and drink bites can be found in the container used and food consumed. Drink bites occurred when mugs or cups were used, while food was consumed in either plates or bowls. In addition to marking the time at which bites were taken, the container used, food or drink consumed, utensil used, and which hand was used were recorded.

The data was evaluated and bites recorded across 276 different subjects during 518 courses containing a total of 22,383 marked ground truth bites. Throughout the dataset, 380 different types of food were consumed using 4 different utensils from 4 different containers [47].

2.5 Ground Truth Scale Data

While ground truth bite times had been recorded across the full dataset previously, these time indicators do not provide the weight of food consumed for each bite. In order to examine the scale data with the goal of detecting and measuring bite weights, a custom tool was developed. The tool was coded using Microsoft Visual C in Visual Studio 2010. Navigation throughout the meal was controlled using the keyboard to manually step forwards and backwards through time as well as to play or pause. Hand labels could be created by using specific keys. Video and scale information were displayed to the user as shown in Figure 2.4. The scale data is shown across the bottom portion of the screen. The vertical bars indicate the ground truth bite times previously marked in the dataset. Both food and drink bites are shown with identical marks.

The labels in Figure 2.4 as follows:

1. Index currently seen in the video.
2. The current cumulative time in the meal.
3. The weight being registered by the scale at the current index.
4. Allows jumping to any index.

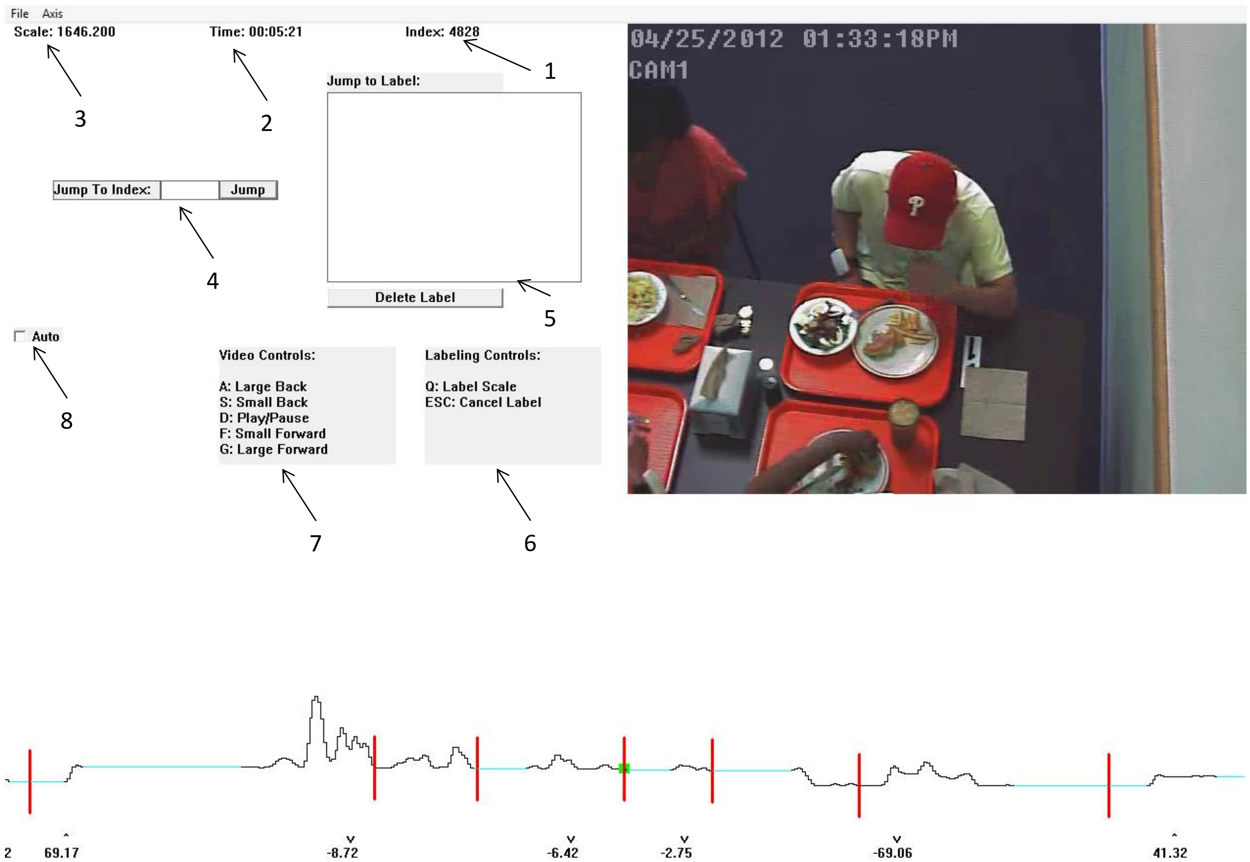


Figure 2.4: Tool used to observe files, label them, and determine algorithm effectiveness.

5. Box populated with any labels previously created.

6. The keys used to label the scale data.

7. Information about how to control the video.

8. Switches between manual labels and labels created by algorithms.

In order to determine the weight of bites, two scale weights must be measured, one before the food or drink is removed from the table, and another after the bite is taken. The weight differential between the two scale readings should indicate the weight of food consumed. Two examples of manually marked scale weights can be seen in Figure 2.5. Both of the bites shown have measurable weights. The first bite is a food bite in which the food that is picked up is immediately consumed. The weight of the bite can be determined by measuring the stable scale weight before the food is picked up and subtracting the stable scale weight that occurs while the bite is being taken. The extended vertical lines show the times as which scale weights were recorded, and the connecting bar directly below the video shows the weight differential between the two scale readings indicated as well as the number of bites taken between the two indices. The result is a 2.97 gram bite weight as the participant picked up and consumed a french fry and is the result of a single food bite. The second bite shown is a food mass bite that can also be measured. In this case, the participant picks up a chicken sandwich, takes a bite, and returns the remainder of the sandwich to the tray. Because of this, two scale events are needed in order to measure the weight of the bite taken. The stable scale weight before the sandwich is picked up and the stable scale weight after the sandwich has been returned to the tray are used to measure the weight of the bite taken. As can be seen in the figure, the bite taken in this case was 14.88 grams.

Using the process described above, data was manually analyzed across 20 randomly selected meals in order to determine the percentage of bites that could be individually measured and those whose weights were impossible to determine. By manually moving through the synchronized video recording and scale data, each scale event could be analyzed using video in order to determine the activities responsible for causing each scale event. Based on the information gathered using this process, it is possible to determine whether each bite is measurable and what events are responsible for making bite weights immeasurable in an ideal scenario, using video evidence to support the scale data.

Across the 20 meals selected, 870 bites were recorded and classified into four categories based on the complexity of the scale data signal and the accuracy of the scale reading. The first category of bites are those bites with a clear measurement procedure and good scale accuracy. For these bites, scale data shows a clear pattern allowing for easy recognition of food or drink consumption, and the scale differential indicating the weight of the bite seems consistent with the video record of the food or drink consumed. The second category of bites are those with a complex measurement procedure and accurate scale data. This category includes bites in which the scale events surrounding the bite do not show a clear pattern to detect an individual food or drink bite using video, but a weight could be determined that seems consistent with the video record of the food or drink consumed (perhaps by averaging multiple bites together or using video to confirm the bite weight). The third category

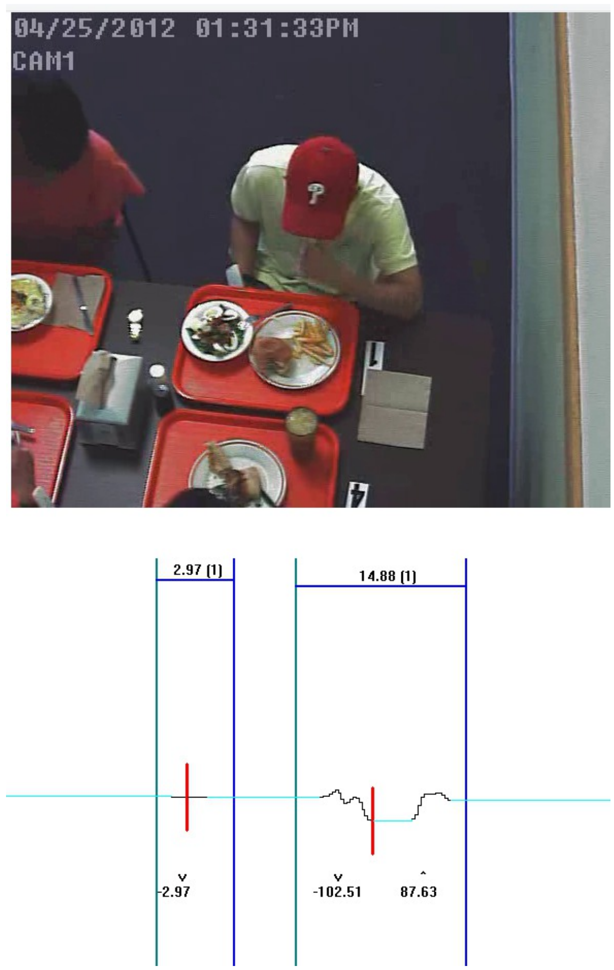


Figure 2.5: Tool used to observe files, label them, and determine algorithm effectiveness. Two separate bites are shown as well as the times used to find scale recordings indicating the weight of each bite.

Table 2.1: Classification of 870 manually measured bites from 20 randomly selected meals

	Accurate Scale Data	Inaccurate Scale Data
Clear Procedure	157	19
Complex Procedure	504	190

of bites includes those with a clear measurement procedure but poor scale accuracy. This category includes bites for which surrounding scale events clearly indicate food or drink consumption, but the bite weight identified by the scale does not match the video record of the food or drink consumed (perhaps due to the tray making contact with the table and thereby invalidating the scale reading). Finally, the fourth category of bites are those in which the scale events surrounding the bite do not show a clear pattern for identification of food or drink bites, and the weight for the bite does not seem to be consistent with the video record (possibly due to the participant resting their arm on the tray and the scale makes contact with the table). All 870 bites found in the 20 meals selected were categorized into these four classifications as shown in Table 2.1.

From Table 2.1, it can be seen that 157 bites, or 18 percent of the bites found in the 20 randomly selected meals, have clear scale events with accurate scale readings, meaning that 18 percent of the bites can be easily identified and have accurate weights recorded. An additional 504 bites, or 58 percent of the total bites, appear to have accurate scale readings, but are complex to measure. This complexity could result from challenging patterns of scale data or from the recognition that multiple bites were taken during one scale event. In the case that the scale data presents challenging patterns, an algorithm could be developed to capture as many of these patterns as possible, but in the case that there are multiple bites taken during one scale event, the weights of the individual bites can not be determined from the scale data. For any bites with inaccurate scale data, the remaining 209 bites, the bite weights cannot be accurately determined. The challenge when developing a method to automatically find bite weights, therefore, is to ensure that all bites with accurate scale data and a clear procedure are identified as bites while those with inaccurate scale data are not. In addition, this method should attempt to identify as many bites as possible that have a complex procedure and accurate scale data, as long as individual bite weights can be determined.

2.6 Weighing Bites at Known Times

In order to identify bite weights, regions of scale stability must be identified. Stable regions in scale data represent times at which there is no activity affecting the scale's weight reading. These regions can be used to find accurate scale readings that can be used to measure bite weights.

Using previously marked ground truth bite times to determine when a bite occurred, an algorithm to measure bite weights was developed. For each ground truth bite time, the algorithm determines if the bite's weight is measurable, and, if so, the program proceeds to determine and record the bite weight. The algorithm used examines the regions of stability and instability located around the bite time and uses these regions to detect if each bite can be measured. Each bite is classified as one of three things; a drink bite, a single food bite, or a food mass bite.

- Drink bites include any instance in which the participant takes a drink from a cup or mug.
- Single food bites are instances in which the participant picks up a bite of food and consumes the food in a single bite, returning nothing to the scale.
- Food mass bites include events in which the participant picks up a mass of food, take a bite, then returns the remainder of the food mass to the scale.

2.6.1 Scale Stability

The first step in analyzing the scale data and determining when food was removed from the tray and ingested begins with identifying the regions of scale stability. In order to ensure that all scale noise was accounted for in the regions of stability, we set the threshold for stability to 5 standard deviations of scale noise. As shown in section 2.2, we found the scale noise to have 1 standard deviation of 0.29 grams. In addition, stable regions were required to have at least 2 seconds of stability in order to qualify as a stable region. This 2 second minimum was determined to be an effective limit based on observing the impact of using various time restrictions on 20 randomly selected meals.

Stability was determined by processing each point of scale data throughout each meal. Each point was tested as the center of a potential stable region. From 1 second before each data point until 1 second after each data point, the average of the scale values was taken. If none of the data points within the 2 second time window differed from the average weight by more than the stability threshold, then the entire 2 second time window was labeled as a stable region. By using this method, all regions of stability were determined for each course recorded. With stable regions (and consequently, unstable regions) labeled, processing could begin in order to identify activities taking place in unstable regions.

2.6.2 Action Classifications

When classifying actions, the primary goal is to determine which actions result in food or beverage consumption and which actions do not. An algorithm was developed to identify if each previously marked bite occurrence corresponded to any of the given the categories of food mass bites, single bites, or drink bites. The algorithm is outlined in Figure 2.6. The algorithm takes advantage of the information already stored when the ground truth bites were recorded to identify if the bite is a food bite. If the bite is a food bite, the algorithm proceeds to determine if the bite has a measurable weight.

The algorithm first determines if the bite is a single food bite. A single food bite is characterized by two scale events, one preceding the bite and one following the bite. During the preceding scale event, the food is picked up. This scale event on its own is not enough to verify that a single food bite occurs, but, once a single food bite is verified, it is enough to determine the weight of food consumed. In order to verify that a single food bite has occurred, the scale event following the bite is required to show a decrease in scale weight. This restriction helps eliminate non-bite scale events from being identified as bites. If weight is returned to the scale during the scale event, it may indicate that a utensil was picked up and later returned to the tray or that a food mass was picked up from the table, only to be returned later. Either of these activities would cause an inaccurate bite weight to be measured. Formally, this pattern can be seen in Figure 2.7 and is defined by Equations 2.1-2.4. A bite marked at time t must meet 5 total criteria to be classified as a single bite and to have its weight recorded. The first criteria is shown in Equation 2.1 and requires that the stable period occurring as the bite is taken must be at least two seconds long. This criteria ensures that the scale is able to reach stability long enough for an accurate weight to be recorded. The second criteria, shown in Equation 2.2, sets limits for where the bite can occur relative to the unstable scale regions surrounding it. A bite must either occur in a stable region of scale data or within 2 seconds of the end of an unstable region of scale data. This is necessary in order to ensure that the food ingested at time t was picked up from the tray in the region of instability from t_a to

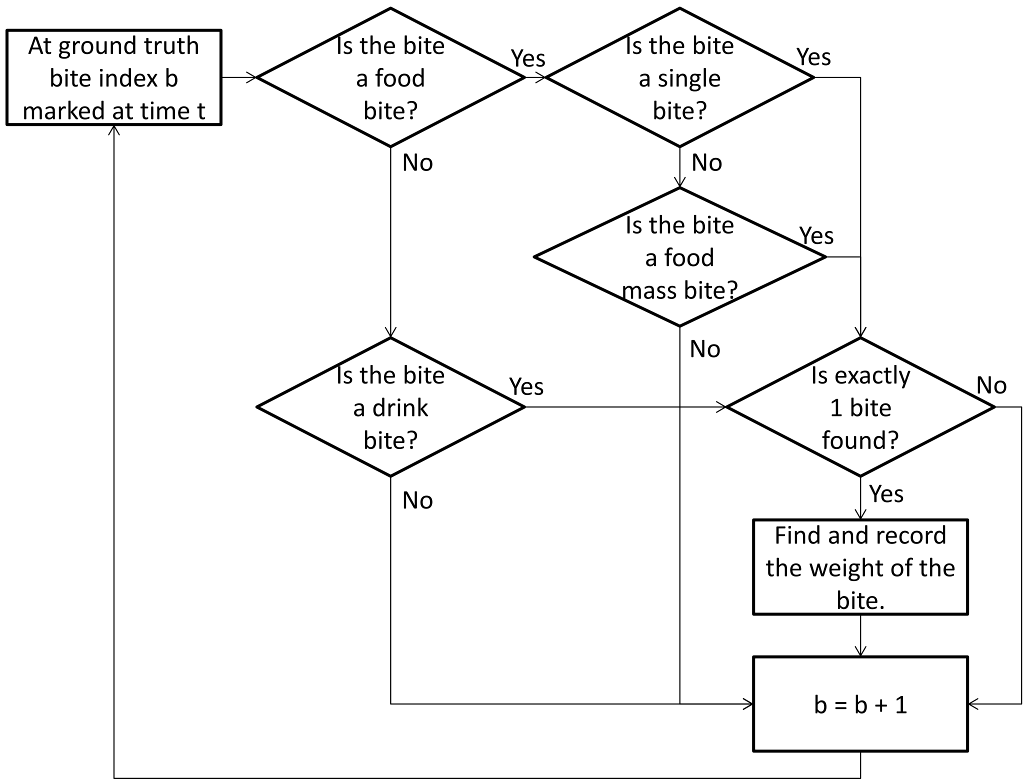


Figure 2.6: Flow chart describing the algorithm used to find bite weights using previously marked ground truth bites.

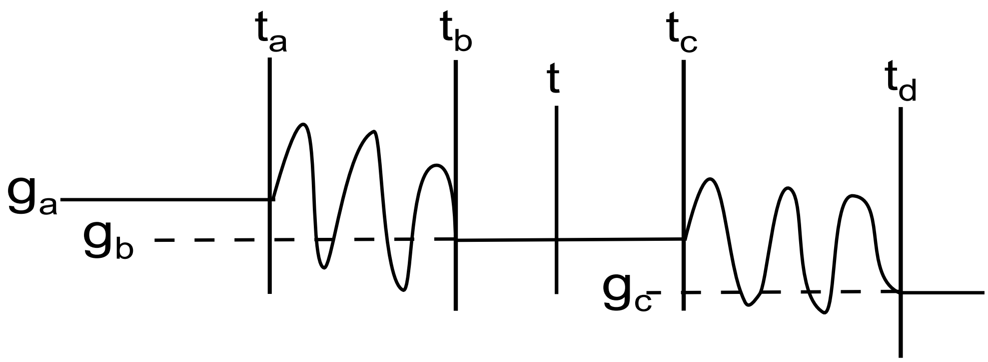


Figure 2.7: A hand drawn depiction of a typical single bite, where t is the time at which a bite is taken.

$\mathrm{t_b}$. In addition to the two time restrictions discussed above, there are two weight restrictions for each potential single bite. First, the weight change before the bite occurs must be restricted. Based on histograms found for bite weights across 20 meals, we found that all measurable bites for those meals had weights between 0 and 30 grams, resulting in the restriction shown in Equation 2.3. To ensure that the food is not returned to the table after the bite is taken, thereby indicating that either a bite was not taken or only some of the food was consumed, Equation 2.4 was also set as a requirement for a single bite. The final requirement for a single bite is that only one ground truth marked bite can occur between $\mathrm{t_a}$ and $\mathrm{t_c}$. This is to ensure that an individual bite weight can be determined by the scale, rather than an average of multiple bites.

$$t_c - t_b \geq 2\text{sec} \tag{2.1}$$

$$t_b \leq t \leq t_c \; or \; t \geq t_b - 2 \geq t_a \tag{2.2}$$

$$0\text{g} \leq g_a - g_b \leq 30\text{g} \tag{2.3}$$

$$g_b - g_c \geq 0\text{g} \tag{2.4}$$

If the requirements for a single bite are not met, then the requirements for a food mass bite are tested. A food mass bite is also characterized by two consecutive scale events. In the case of a food mass bite, the first scale event should result in a weight decrease as the food mass is picked up. The second scale event occurs after a bite is taken from the food mass and the remainder of the food mass is returned to the scale. Formally, this pattern can be seen in Figure 2.8 and defined by Equations 2.1-2.2 and 2.5-2.6. The first two restrictions for a food mass bite are time restrictions identical to those required of a single bite. These two restrictions are shown in Equation 2.1 and Equation 2.2. In addition to these two time restrictions, there are two additional weight restrictions for a food mass bite. The first restriction is a restriction on the weight change that occurs before the bite is taken as shown in Equation 2.5. Based on food mass bites recorded in 20 randomly selected meals, a range of 100g to 300g was determined to be a characteristic of all recordable food mass bites. In addition to the pre-bite weight restriction, a post bite weight restriction is used to restrict which bites can be classified as food mass bites. The restriction is shown in Equation 2.6, and the result of the restriction is that the food mass bite must have its net weight change restricted to the range of 0g to 30g. This range is identical to the range specified for single bites. The final restriction for a food mass bite is that only one bite can be marked between $\mathrm{t_a}$ and $\mathrm{t_d}$ in order to ensure that only a single bite weight, rather than an average of multiple bite weights, is measured. This requirement is similar to the single bite restriction for single bites, except $\mathrm{t_d}$ is the final time index checked rather than $\mathrm{t_c}$. This change is necessary because the weight calculation for a food mass bite finds the difference between $\mathrm{g_a}$ and $\mathrm{g_c}$, whereas the weight of a single food bite is calculated by finding the difference between $\mathrm{g_a}$ and $\mathrm{g_b}$. This change in weight calculation means that, for a single bite, the weight calculation includes any bites taken between $\mathrm{t_a}$ and $\mathrm{t_c}$, but for a food mass bite, the weight calculation includes any bites taken between $\mathrm{t_a}$ and $\mathrm{t_d}$.

$$100\text{g} \leq g_a - g_b \leq 300\text{g} \tag{2.5}$$

$$g_a - g_b - 30\text{g} \leq g_c - g_b \leq g_a - g_b \tag{2.6}$$

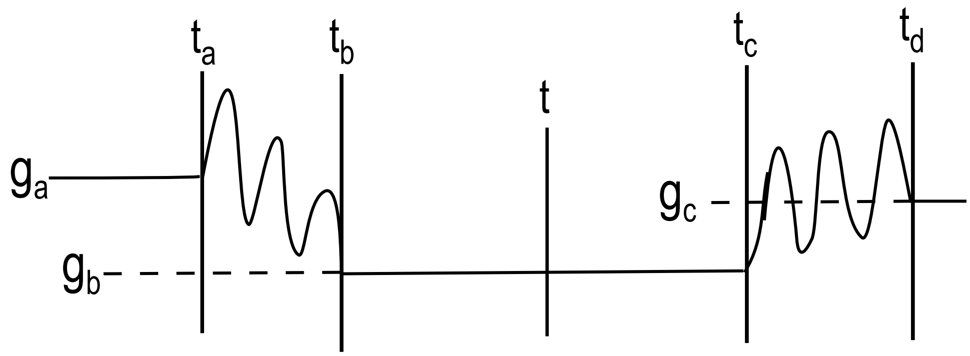


Figure 2.8: A hand drawn depiction of a typical food mass bite, where t is the time at which a bite is taken.

Assuming the requirements of a single bite are not met and the requirements of a food mass bite are not met, only a drink bite could be measured at the bite of interest. If the bite is determined to be a drink based on the information recorded when the bite was manually marked, then the restrictions for a drink to be measurable are checked. A drink bite is characterized by two scale events. During the first scale event, the scale weight should decrease significantly as the drink is picked up. Between the two scale events, a drink should occur. During the second scale event, the scale weight should increase significantly as the remainder of the drink is returned to the tray. Formally, this pattern can be seen in Figure 2.9 and is defined by Equations 2.1-2.2 and 2.7-2.8. The first restrictions for a drink bite are identical to the first two for both single bites and food mass bites. These restrictions are identified in Equation 2.1 and Equation 2.2. In addition to the two specified time restrictions, there are a set of two unique weight restrictions for a measurable drink bite to meet. Based on analysis of 20 randomly selected meals, it was determined that most drinks, when picked up, meet a weight threshold of 80g to 550g as identified in Equation 2.7. In addition, it was discovered that the weight of liquid consumed typically falls between 0g and 80g, prompting the restriction shown in Equation 2.8. Finally, for a drink bite to be individually measured, only one bite can be marked between t_a and t_d. This restriction is the same as the restriction for food mass bites because the weight of a food mass bite and a drink bite is calculated in an identical manner. If multiple bites were marked in the time from t_a to t_d, then the weight of each individual bite could

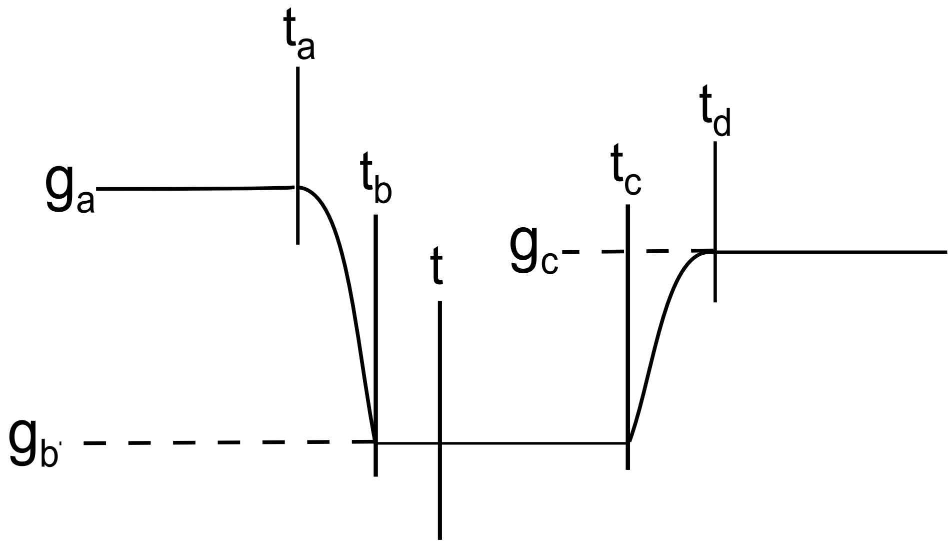


Figure 2.9: A hand drawn depiction of a typical drink bite, where t is the time at which a bite is taken.

not be measured based on the available data.

$$80\text{g} \leq g_a - g_b \leq 550\text{g} \tag{2.7}$$

$$g_a - g_b - 80\text{g} \leq g_c - g_b \leq g_a - g_b \tag{2.8}$$

2.7 Automatic Event Detection and Classification

The method for automatically detecting EI events is based on the information gathered when determining the weight of ground truth bites. The method developed uses the same basic templates for single bites, food mass bites, and drink bites, but every data point in the scale data is treated as a potential bite. When using this method, additional restrictions must be placed on the potential bites found at any point in time due to the fact that each point in time is not a previously identified bite. In addition, the algorithm was changed due to the lack of available information about whether any individual bite is a drink bite or a food bite. The algorithm used for the automatic detector is shown in Figure 2.10.

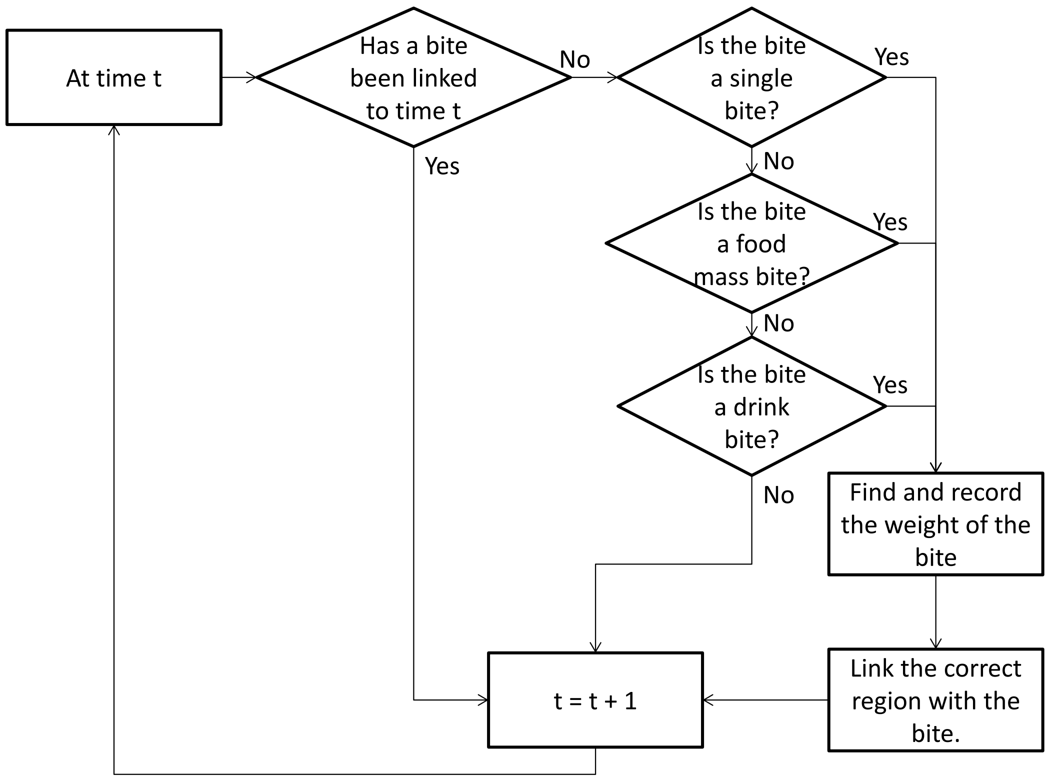


Figure 2.10: Flow chart describing the algorithm used to find bite weights automatically.

The first step in understanding the automatic algorithm is to identify what it means for a bite to be linked to a time t. When treating each instant in time as a potential bite, steps must be taken to ensure that no bite weight is measured multiple times as multiple bites. When a bite is identified, a region of time is marked as linked to the bite. Based on the type of bite, we can clearly define this region. Assuming a single bite is identified using the ideal example shown in Figure 2.7, then the full region of time from t_a and t_c is then linked with the bite found at time t. Therefore, if a single bite is found, no more bites of any type can be found between t_a and t_c. Food mass bite and drink bites must be linked to different regions. If a food mass bite is found following the template shown in Figure 2.8, then the full time region from t_a to t_d should be linked to the bite, and no more bites should be recorded within that time frame. Similarly, for a drink bite following the template shown in Figure 2.9, the time from t_a to t_d should be linked to the drink bite, and no more bites should be allowed within those limits.

After bypassing any time slots with which a bite has already been linked, the automatic algorithm checks each of the three possible classifications of bite, though the weight restrictions for each bite type are slightly altered. The ground truth algorithm operates with knowledge that each point being tested has been visually verified as a bite, so a minimum bite weight of 0g is appropriate. The automatic algorithm has no such guarantee for each time index being tested, so the minimum bite weight restriction must be altered to compensate for scale noise. In order to filter out events caused by noise in the scale, each bite in the automatic algorithm had a minimum weight threshold of 3 times the scale noise, or 0.88 grams. The result of this change can be represented mathematically for the equations defining each of the three different bite classifications.

Each of the three bite classifications used in the automatic algorithm share many similarities with the versions used in the ground truth algorithm. For each classification of bite, the time restrictions identified for the ground truth algorithm earlier in Equation 2.1 and Equation 2.2 still hold for each classification of bite in the automatic algorithm. The single bite algorithm still follows the template shown in Figure 2.7, though the weight restrictions are altered slightly. The equations defining the weight requirements are redefined to eliminate scale noise from being registered as a bite by replacing Equation 2.3 with Equation 2.9, and Equation 2.4 is replaced by Equation 2.10. In addition, food mass bites still follow the template shown in Figure 2.8 with different weight restrictions. For food mass bites, Equation 2.6 is replaced by Equation 2.11 in order to eliminate scale noise from showing up as a bite. Finally, drink bites still follow the template shown in Figure

2.9, but Equation 2.8 is replaced by Equation 2.12 in order to eliminate scale noise from being recognized as a bite.

$$3\sigma_{noise} \leq g_a - g_b \leq 30\text{g} \tag{2.9}$$

$$g_b - g_c \geq 3\sigma_{noise} \tag{2.10}$$

$$g_a - g_b - 30\text{g} \leq g_c - g_b \leq g_a - g_b - 3\sigma_{noise} \tag{2.11}$$

$$g_a - g_b - 80\text{g} \leq g_c - g_b \leq g_a - g_b - 3\sigma_{noise} \tag{2.12}$$

Chapter 3

Results

In order to evaluate the results of the automatic detection algorithm, the results provided through use of ground truth marked bites were used as a baseline. Any bite marked as a ground truth bite without having a measurable weight is a bite for which no weight can be detected using scale data alone. These examples of bites whose weights cannot be detected are a result of limitations built in to using a scale to collect data in a free eating environment in addition to the limitations of the setup we developed to observe and record data on the meals used in the dataset.

Any scale events in which multiple bites occurred were deemed immeasurable. Even though these bite weights could be estimated by dividing the weight differential by the number of bites taken, this estimation does not find the true weight of each individual bite. The variations in the weight of each bite could lead to incorrect analysis, so the algorithm is limited to just those times when one bite corresponds to one scale event. Table 3.1 shows the number of bites available throughout the entire dataset as well as the results from both the ground truth and automatic algorithms. The table also summarizes the number of confirmed measurable bites found by both algorithms. From this table, it can be seen that both the ground truth and automatic algorithms were able to capture roughly 25% of the food bites available in the full dataset. In addition, both algorithms were able to capture over 50% of the drink bites. These results indicate that drink bites were easier to distinguish in the scale data than food bites across both algorithms. The automatic algorithm was able to identify over 80% of the measurable bites found by the ground truth algorithm, though nearly 40% of the events identified by the automatic algorithm were false alarms.

Table 3.1: Bites measured by both algorithms.

	Full dataset	Bite times known	Automatic detection
All bites	22380	7501	6155
Food bites	18872	5023	4390
Drink bites	3508	2478	1765
Errors (false alarms)	N/A	N/A	4085

3.1 Relating BMI and Bite Weight

Since the BMI of each participant was recorded, a BMI can be tied to a bite weight for each occurrence of a verified bite weight. Our dataset consists of thousands of bites corresponding to hundreds of participants with BMIs ranging from 17.4 to 46.2 and bite weights ranging from 0 to 79.8 grams. Showing 5% of all data points collected for clarity, the plot of BMI to bite weight can be seen in Figure 3.1. Due to the volume of data points as well as the dense clustering between 20 to 25 BMI and 0 to 10 grams, treating each bite individually would skew the relationship between BMI and bite weight because the data density increases near the center of mass. Binning the data based on participant would lead to the same effect due to the high density of participants in the 20 to 25 BMI range. We determined that binning bites based on participant BMI and averaging the bite weights in each bin would allow the relationship between BMI and bite weight to be observed more clearly. We sorted the bite weights into 10 bins based on participant BMI and averaged all of the bites in each bin using Equation 3.1. In the equation, A is the average bite weight for the bin, N is the number of bites in the bin, and W is the weight of a bite in the bin. The relationship between bite weights and BMI was taken with the exception of the highest 3 BMI bins. The highest three bins were removed from all results reported from all of our sources due to there being a much smaller sample size of bites (less than 100 bites per bin).

$$A = \frac{\sum_{i=1}^{N} W_i}{N} \tag{3.1}$$

After binning an averaging, we found a line of best fit in order to determine the change in bite weight based on BMI. To accomplish this task, we used the polyfit function in Matlab to find the slope and intercept of a first order line of best fit. An example of the results we found is shown in Figure 3.2. The slope of the line of best fit is 0.28 g/BMI and corresponds to the rate of

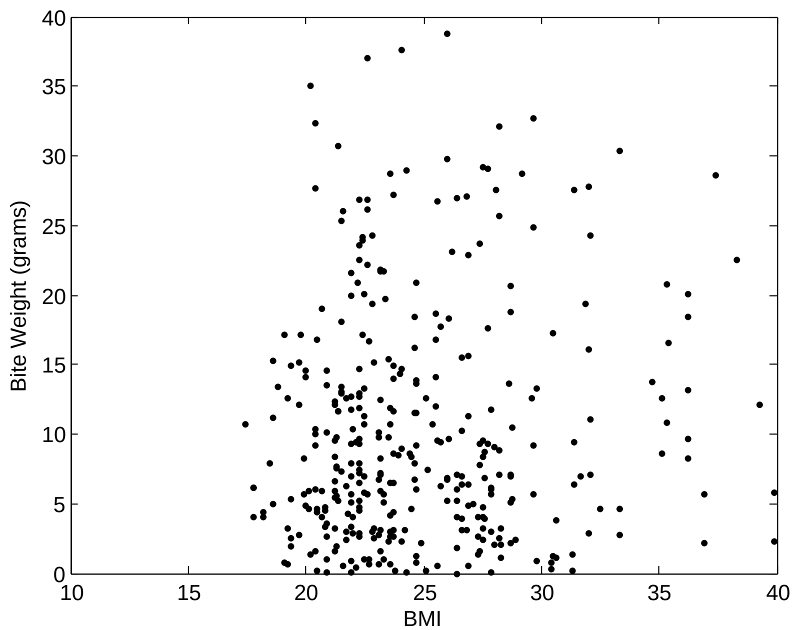


Figure 3.1: A comparison of BMI to bite weight for the raw data collected from the ground truth driven algorithm. Only 5% of the total bites are shown for clarity.

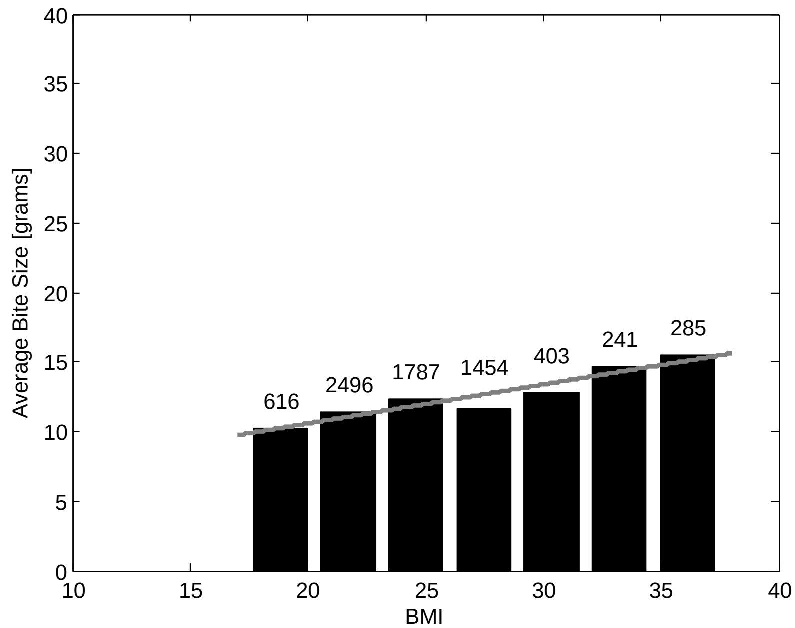


Figure 3.2: Average bite weight grouped by BMI for the ground truth driven algorithm. The line of best fit and the number of bites in each bin are shown.

increase in average bite size as BMI increases. The number of bites in each bin is displayed above each corresponding BMI bin.

3.2 Weighing Known Events

The results from the ground truth algorithm showed a clear pattern of increasing average bite weight across BMI. While 0.28 g/BMI slope may appear to be a small difference, if summed up over bites in each meal every day, it becomes clear that this small weight difference has a profound cumulative effect on energy intake.

With the average of each bin having been found using Equation 3.1, the variance of each bin should be analyzed. When the bite weight distribution, as shown in Figure 3.3, is studied, it becomes clear that standard deviation is not the best statistic for measuring the variability present in the data because the data resembles a half Gaussian. In order to observe the variance in the data, we instead broke each BMI bin into quartiles and observed the change in g/BMI for each quartile.

Table 3.2: GT algorithm grams per BMI slopes broken into quartiles

	1st quartile	2nd quartile	3rd quartile	4th quartile
g/BMI slope	0.05	0.17	0.30	0.58

The ability to break this data into quartiles is only possible due to the measurement of individual consumption events, rather than fitting a trend line to a full meal. The slope for each quartile can be seen in Table 3.2. The slope for each quartile was found by finding the line of best fit for BMI to bite weight plots for each quartile of bites in each bin. The average bite weight for each quartile of bites was found using Equation 3.2, where A is the average bite weight within the bin for each quartile, q=1...4, corresponding to each of the four possible quartiles, N is the number of bites in the bin, and W is the list of bite weights in the bin sorted from smallest to largest. The results show that as the size of the bite increases, the relationship between g/BMI becomes steeper. This indicates that people of all BMIs take similar size bites 25% of the time, for the smallest quartile of bites. For each successive quartile, however, people with larger BMIs take larger bites than people with smaller BMIs. At the largest quartile, people are taking 0.58g more per bite per BMI. This suggests that behavior treatment could be targeted towards resisting the urge to take larger bites.

$$A_q = \frac{\sum_{i=(q-1)_N+1}^{q_N+1} W_i}{N/4} \tag{3.2}$$

The next step in breaking down the data gathered by the GT BMI is separating the food and drink bites. Observing food and drink bites can provide more insight into the source of the bite weight differential present in people with different BMIs. It was found that for food bites in particular, food bite weights remain more consistent across all BMIs. Using the same method of binning the results by BMI and averaging the food bite weights in each bin, the plot shown in Figure 3.4a was produced. The line of best fit for the figure indicates that food bites have a weight increase of .12 g/BMI. The same relationship is shown to be .30 g/BMI for drinks as shown in Figure 3.4b. The results shown indicate that when it comes to drinks, there is a larger increase in the average bite weight as BMI increases.

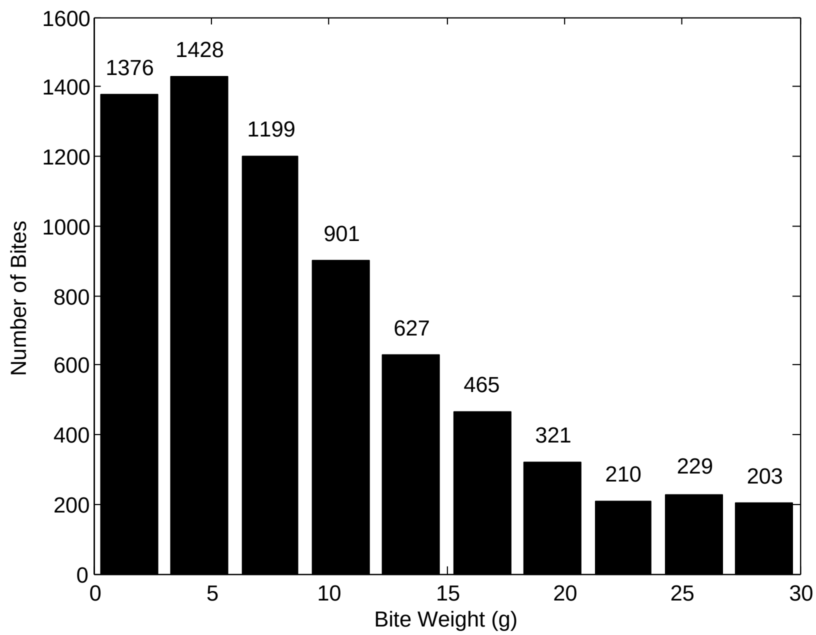


Figure 3.3: Histogram for bite weights detected by the ground truth algorithm.

Table 3.3: Automatic Algorithm Results.

Bite Type	Bite Count	Slope (g/BMI)
All automatically detected events	10240	0.21

3.2.1 Zero Calorie Beverages Compared to High Calorie Beverages

From individual drink analysis, the relationship between zero calorie drinks and high calorie drinks was analyzed. If participants tended to take larger drinks with high calorie beverages, for example, it could provide more insight into BMI differences related to beverage choice. The distribution of grams per bite for both zero calorie and high calorie beverages was examined. As can be seen in Figures 3.5a and 3.5b, histograms of zero calorie drink bites and high calorie drink bites respectively, there is little to no difference between the two beverage types. In addition, when observed based on BMI, the distributions are also identical. The conclusion from this analysis is that the relationship between bite size and BMI is not affected by the type of beverage, whether zero calorie or high calorie.

3.3 Automatically Detected Scale Events

For automatically detected scale events, 10,240 events through the 518 courses in the dataset were found. In order to evaluate the accuracy of the automatically detected events, the 10,240 events found were compared to all the measurable bites identified through use of manually marked ground truth bite times. There were 6,155 automatically detected events matched to measurable ground truth bites. Evaluated against the 22,380 total actual bites, this indicates that the algorithm has a sensitivity of 0.34 and a positive predictive value of 0.71 for detecting all bites.

The relationship between weight and BMI for automatically detected events showed a .21 g/BMI increase as shown in Figure 3.6 and summarized in Table 3.3. The results still show a strong positive increase in g/BMI at an event level, even without full knowledge of bite times. Compared to the ground truth driven algorithm, the results show a 0.07 g/BMI smaller increase due to the noise events detected in the automatic algorithm.

In order to evaluate bite weight to BMI relationship for automatically detected events that were confirmed to be measurable bites, the automatically detected events found were matched to

Table 3.4: Slope in g/BMI for all bites types.

	Ground truth algorithm	Automatic algorithm
All events	0.28	0.27
Food bites	0.12	0.10
Drink bites	0.30	0.17
Errors (false alarms)	N/A	0.0023

measurable bites identified using ground truth bites times and the results were evaluated. When dealing with this subset of events, we discovered that the rate of g/BMI increase has a slope of 0.27 from the plot shown in Figure 3.7. This indicates that the automatically detected events that were confirmed as measurable bites show the same pattern as all measurable bites as determined by manually marked ground truth bites. We also broke the matched bites into food and drink bites to observe whether the same patterns persisted for this subset of bites. The results showed that the confirmed food bites had a slope of 0.10 g/BMI as seen in Figure 3.8a and the confirmed drink bites had a slope of .17 g/BMI as seen in Figure 3.8b. A full comparison of the relationship between bite weight and BMI for automatically detected events that are confirmed to be measurable bites is summarized in Table 3.4. The slopes are similar between the two algorithms with the exception of drink bites. This slope changed significantly because the automatically detected events that were verified as drink bites include a number of events in which a drink was picked up and multiple sips or gulps were taken before the drink was returned to the scale. This caused the weights calculated for automatically detected drink events to be significantly larger across all BMIs. This increase in weight across all BMIs lessened the slope.

3.3.1 Analysis of Failed and Extra Event Detections

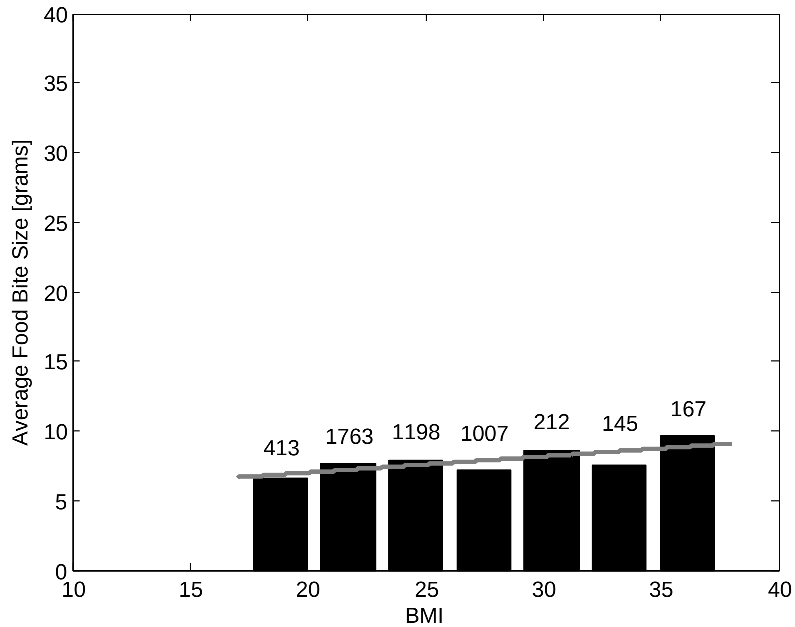
In order to examine the nature of false alarms more closely, 20 meals were randomly selected to provide further analysis. Any automatically detected events that were not confirmed as measurable bites were examined in detail. By watching the video during these segments, it was found that a significant number of errors fall into one of three categories. The first category of errors are those during which participants took multiple bites of food during a single scale event. The second category is one in which scale artifacts produce events that look like a bite, but during which, no bite is taken. The third category is one in which bites are taken at the start of an unstable period in the scale data. This is a problem for two reasons; 1) identifying which scale event can be used to find the bite's weight becomes ambiguous and 2) when the scale is unstable while a bite is being taken, it is an indicator that the participant is resting their arm on the tray or shifting items on the tray. In order to evaluate whether the errors could be filtered out easily, we compared each of the distributions to the distribution of bites that were confirmed to have measurable weights. The distribution of measurable bite weights can be seen in Figure 3.9a.

When multiple bites are taken during one scale event, preventing individual bite weights from being taken, the scale event in which the food is picked up is generally larger than a typical measurable bite. The distribution for this type of false alarm can bee seen in Figure 3.9b. While the mean for this type of false alarm is higher than that for measurable bites, the difference is not enough to allow for this type of error to be filtered easily.

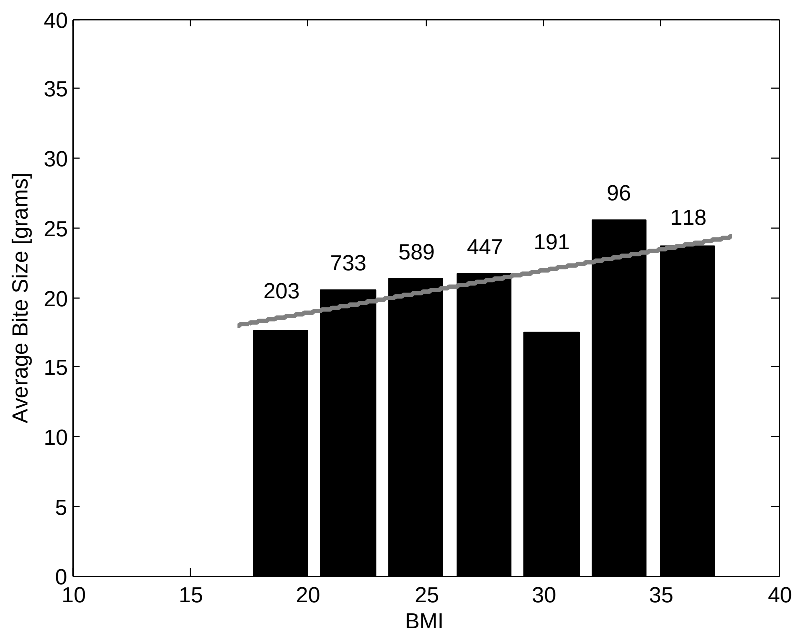
The second error results from noise created by scale artifacts. Scale artifacts produce a mean that is significantly below the mean for actual bites as can be seen in Figure 3.9c. This mean shift makes sense, since most scale artifacts are caused by items being shifted on the scale, causing the tray to partially rest on the table. One flaw in the experimental setup was that the clearance between the tray and the table is small enough that moving a glass of water or plate of food near the edges of the tray can cause it to flex and contact the table. This type of event as well as similar events caused by participants resting their arms on the tray create these scale artifacts. It is clear that while the distribution of scale artifacts differs from the distribution of measurable bites, the overlap is strong enough that any removal of scale artifacts also removes a significant portion of actual bites.

The third and final class of error is an error in which a bite is taken near the start of a period of instability. Taking a bite early on in a period of instability indicates that the participant was doing more than simply picking up a bite and consuming it. If no other activity occurs while the bite is taken, the scale would be stable as the food is consumed. Based on this observation, false alarms resulting from this eating pattern typically occur when the participant is either resting their an arm on the tray (which leads to inaccurate scale readings) or shifting items around on the tray while taking a bite (which also leads to inaccurate scale readings). These activities frequently invalidate scale readings, and the bites characterized by these activities have a distribution as shown in Figure 3.9d. The distribution of these false alarms is not differentiable from the distribution of measurable bites, so this type of false alarm also cannot be easily filtered out using raw scale data.

All three major types of false alarms found exhibit distributions too similar to measurable bites for the events to be easily filtered out.

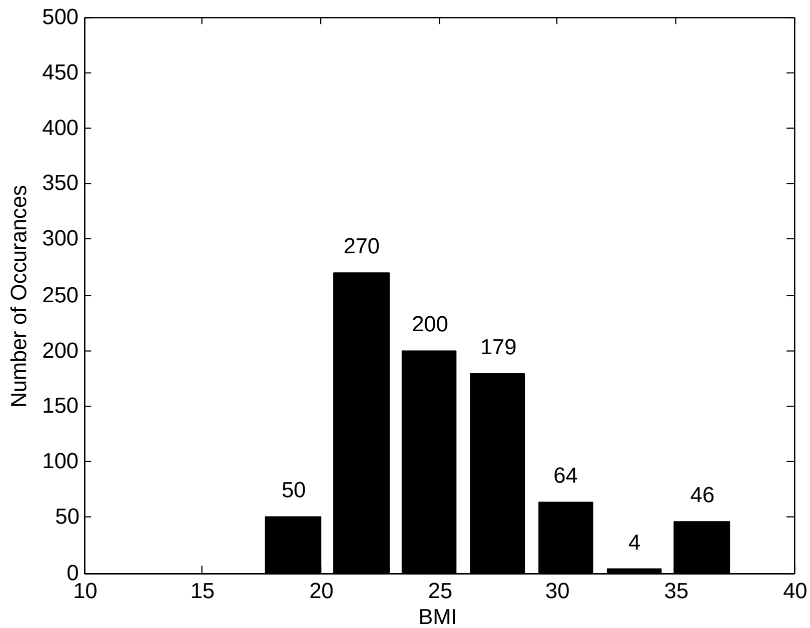


(a) Average food bite weight grouped by BMI.

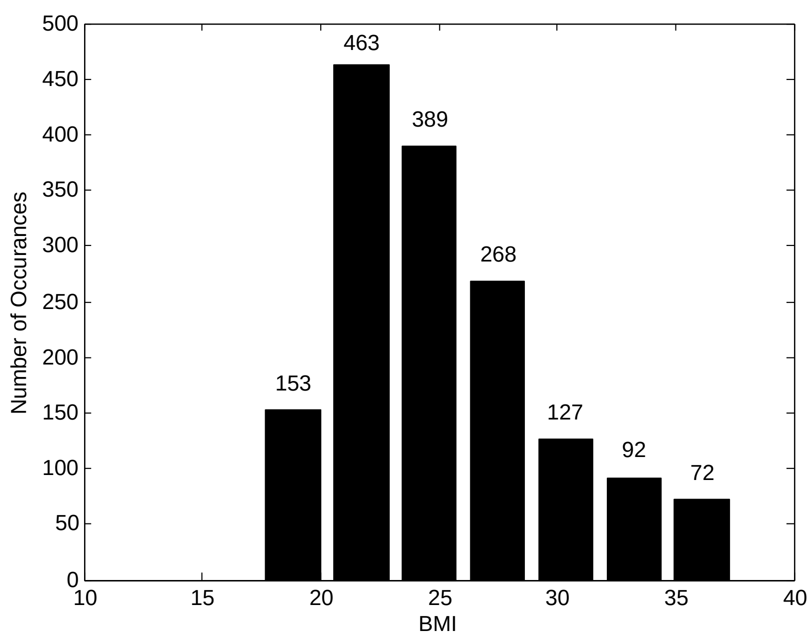


(b) Average drink weight grouped by BMI.

Figure 3.4: Average food and drink weights grouped by BMI.



(a) Histogram for the weight of zero calorie beverage drinks.



(b) Histogram for the weight of high calorie beverage drinks.

Figure 3.5: Histograms comparing the distribution of high calorie and zero calorie drink bites. It can be seen that the distributions appear similar.

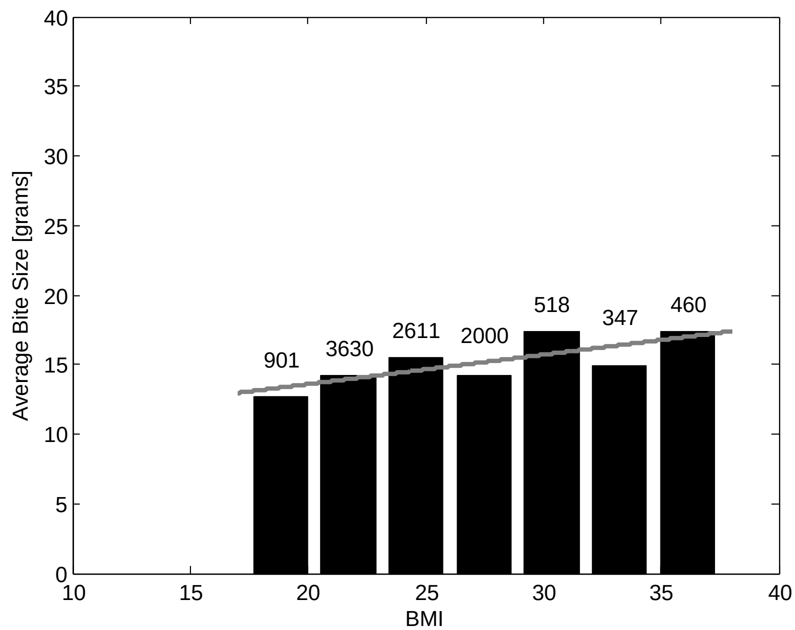


Figure 3.6: Average weight for all events detected by the automatic algorithm broken into bins by BMI.

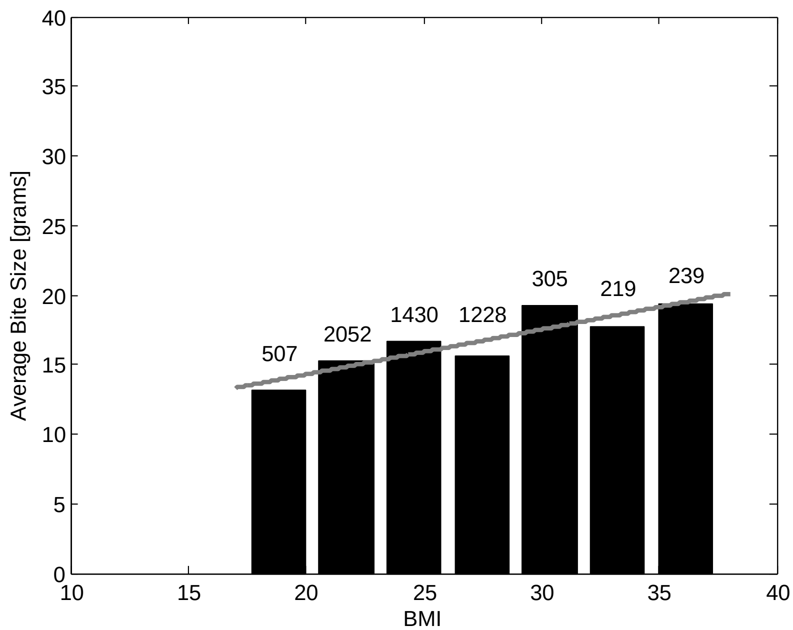
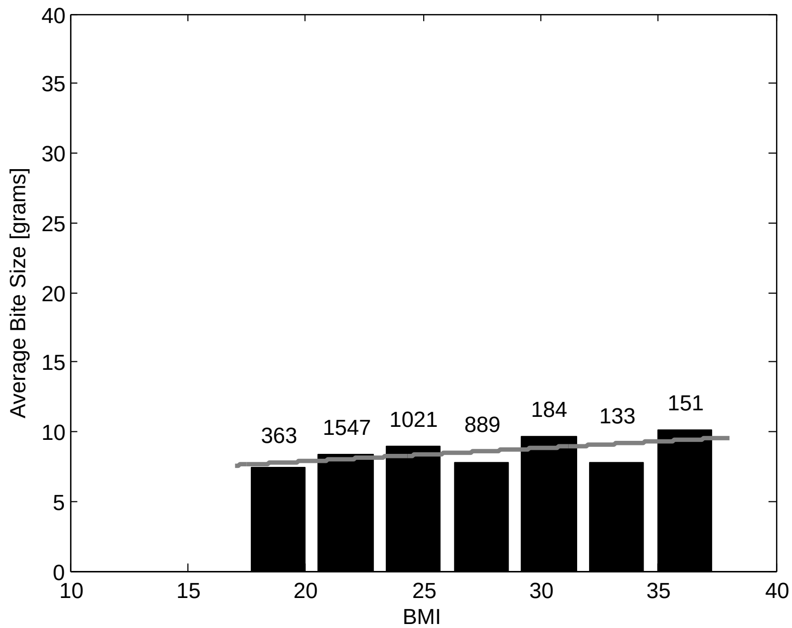
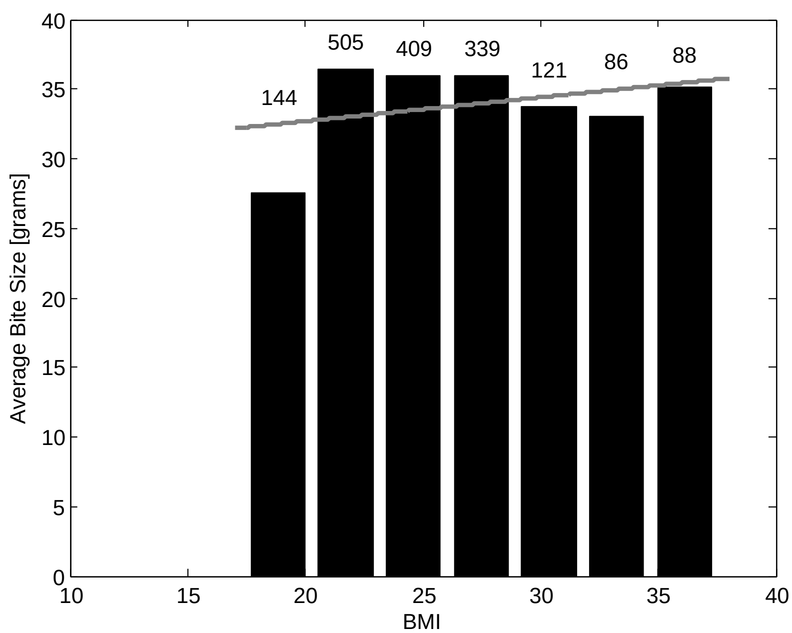


Figure 3.7: Average bite weight for ground-truth verified bites detected by the automatic algorithm broken into bins by BMI.

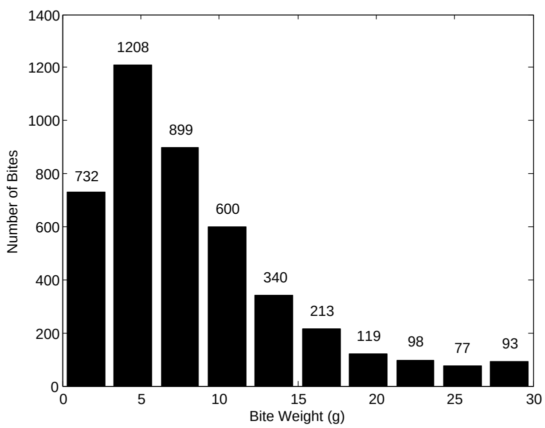


(a) Average bite weight for ground-truth verified food bites detected by the automatic algorithm broken into bins by BMI.

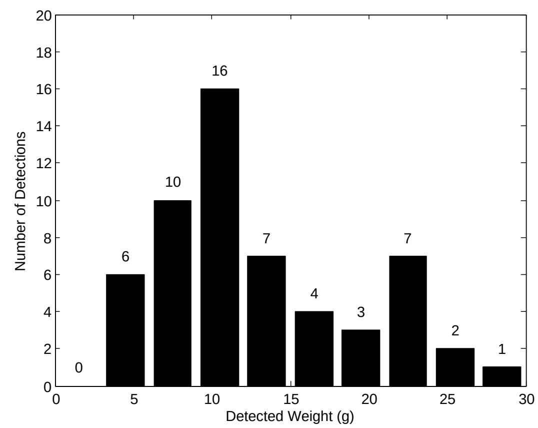


(b) Average bite weight for ground-truth verified drink bites detected by the automatic algorithm and broken into bins by BMI.

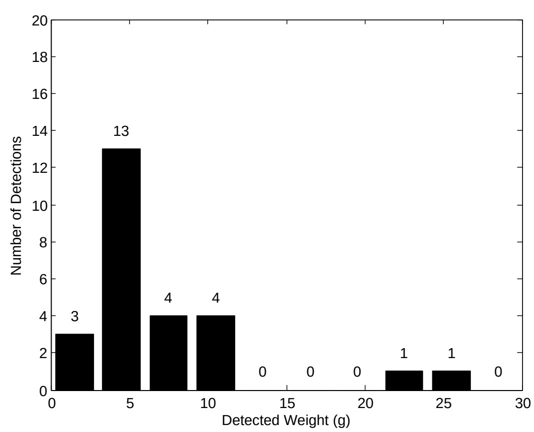
Figure 3.8: Average weights ground truth verified bites detected automatically.



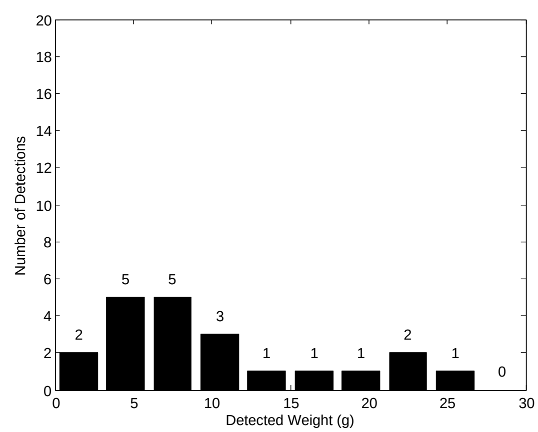
(a) Histogram of weight of actual bites.



(b) Distribution of erroneous bite detections in which bites were taken at the start of an unstable period.



(c) Distribution of erroneous bite detections in which scale artifacts produce unstable periods resembling bites.



(d) Distribution of erroneous bite detections in which the bite is taken near the start of an unstable period, indicating a bite that is likely to be mis-measured.

Figure 3.9: Comparison of the distributions of automatic events verified as actual bites and false alarms.

Chapter 4

Conclusion

This work is motivated to improve the tools available for tracking energy intake. A tool capable of measuring the weight of individual consumption events would be a powerful tool for people seeking to take a more active role in managing their health. The data gathered in this thesis showed that in a real-world, free-eating environment, about 30% of the consumption events can be detected and weighed. The remaining 70% of the events either take place without interacting with the scale (as when people take multiple bites from one food mass before returning it to the scale) or when there is noisy interaction with the scale that prevents measurement (as when people rest their arm on the scale).

The relationship between bite size in grams and BMI was analyzed across all 7,501 measurable bites found using ground truth marked bite times and across 10,240 automatically detected events. Though the data showed a heavy density near the center of the bite weight and BMI distributions, by binning the results based on BMI and finding the average bite weight, a relationship between bite size and BMI was found. When analyzing the measurable bites found using ground truth marked times, the relationship between bite weight and BMI is shown to be 0.28 g/BMI. This relationship is on a per-bite basis, so while the relationship is a relatively small weight change between BMIs, the cumulative effect is quite significant. Without using previously marked bite times, a relationship of 0.21 g/BMI was found. While the trend is diminished, it is still clearly present even with the presence of false alarms detected using the automatic algorithm.

If each bin is broken into quartiles, the relationship between g/BMI can be further analyzed. When this analysis was performed, the results showed that grams/bite vs BMI is roughly constant for the smallest 25% of bites but increases in each quartile. When the largest quartile of bites are analyzed, there is a relationship of 0.58 g/BMI. This suggests that behavior treatment could be targeted towards resisting the urge to take larger bites.

Individual consumption events were correctly detected 28% of the time in the real-world, free-eating environment. False alarms occurred roughly twice for every three true consumption events detected. This study is the first to our knowledge to detect and measure individual consumption events during natural eating with unrestricted foods.

Bibliography

[1] K. Acheson, I. Campbell, O. Edholm, D. Miller, and M. Stock, "The measurement of food and energy intake in man-an evaluation of some techniques." *The American Journal of Clinical Nutrition*, vol. 33, no. 5, pp. 1147–54, 1980.

[2] F. Thompson and A. Subar, *Dietary Assessment Methodology*, 2nd ed. Academic Press/Elsevier, 2008.

[3] L. Muhlheim, D. Allison, S. Heshka, and S. Heymsfield, "Do unsuccessful dieters intentionally underreport food intake?" *International Journal of Eating Disorders*, vol. 24, no. 3, pp. 259–266, 1998.

[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, 2001.

[5] A. Coulston and C. Boushey, *Nutrition in the Prevention and Treatment of Disease.* Academic Press, 2008.

[6] J. Witschi, "Short-term dietary recall and recording methods," *Nutritional epidemiology*, vol. 4, pp. 52–68, 1990.

[7] K. Siek, K. Connelly, Y. Rogers, P. Rohwer, D. Lambert, and J. Welch, "When do we eat? an evaluation of food items input into an electronic food monitoring application," in *Pervasive Health Conference and Workshops, 2006*, December 2006, pp. 1–10.

[8] E. Wohlers, J. Sirard, C. Barden, and J. Moon, "Smart phones are useful for food intake and physical activity surveys," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, September 2009, pp. 5183–5186.

[9] I. Lopes, B. Silva, J. Rodrigues, J. Lloret, and M. Proenca, "A mobile health monitoring solution for weight control," in *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, November 2011, pp. 1–5.

[10] A. Smeets and M. Westerterp-Plantenga, "Acute effects on metabolism and appetite profile of one meal difference in the lower range of meal frequency," *British Journal of Nutrition*, vol. 99, pp. 1316–1321, 5 2008.

[11] J. Speakman, *Doubly Labelled Water: Theory and Practice.* Springer, 1997.

[12] G. Plasqui, A. Joosen, A. Kester, A. Goris, and K. Westerterp, "Measuring free-living energy expenditure and physical activity with triaxial accelerometry," *Obesity Research*, vol. 13, no. 8, pp. 1363–1369, 2005.

[13] D. Schoeller *et al.*, "Measurement of energy expenditure in free-living humans by using doubly labeled water," *The Journal of Nutrition*, vol. 118, no. 11, pp. 1278–1289, 1988.

[14] C. Champagne, G. Bray, J. Delany, A. Kurtz, J. Monteiro, E. Tucker, and J. Volaufova, "Energy intake and energy expenditure: a controlled study comparing dietitians and non-dietitians," *Journal of the American Dietetic Association*, vol. 120, no. 10, pp. 1428–1432, 2002.

[15] D. Schoeller, "Limitations in the assessment of dietary energy intake by self-report," *Metabolism*, vol. 44, Supplement 2, no. 0, pp. 18–22, 1995.

[16] M. Livingstone, A. Prentice, J. Strain, W. Coward, A. Black, M. Barker, P. Mckenna, and R. Whitehead, "Accuracy of weighed dietary records in studies of diet and health." *BMJ: British Medical Journal*, vol. 300, no. 6726, p. 708, 1990.

[17] S. Jonnalagadda, D. Mitchell, H. Smiciklas-Wright, K. Meaker, N. Heel, W. Karmally, A. Ershow, and P. Kris-Etherton, "Accuracy of energy intake data estimated by a multiplepass, 24-hour dietary recall technique," *Journal of the American Dietetic Association*, vol. 100, no. 3, pp. 303–311, 2000.

[18] G. Ulen, M. Huizinga, B. Beech, and T. Elasy, "Weight regain prevention," *Clinical Diabetes*, vol. 26, no. 3, pp. 100–113, 2008.

[19] L. Burke, J. Wang, and M. Sevick, "Self-monitoring in weight loss: A systematic review of the literature," *Journal of the American Dietetic Association*, vol. 111, no. 1, pp. 92–102, 2011.

[20] B. McCabe-Sellers, "Advancing the art and science of dietary assessment through technology," *Journal of the American Dietetic Association*, vol. 110, no. 1, pp. 52–54, 2010.

[21] F. Thompson, A. Subar, C. Loria, J. Reedy, and T. Baranowski, "Need for technological innovation in dietary assessment," *Journal of the American Dietetic Association*, vol. 110, no. 1, pp. 48–51, 2010.

[22] World Health Organization. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs311/en/http://www.test.org/doe/

[23] Harvard School of Public Health. [Online]. Available: http://www.hsph.harvard.edu/obesity-prevention-source/obesity-trends/

[24] B. Caballero, "The global epidemic of obesity: An overview," *Epidemiologic Reviews*, vol. 29, no. 1, pp. 1–5, 2007.

[25] A. Must, J. Spadano, E. Coakley, A. Field, G. Colditz, and W. Dietz, "The disease burden associated with overweight and obesity," *JAMA: The Journal of the American Medical Association*, vol. 282, no. 16, pp. 1523–1529, 1999.

[26] E. Finkelstein, I. Fiebelkorn, and G. Wang, "National medical spending attributable to overweight and obesity: how much, and who's paying?" *Health Affairs Web Exclusives*, pp. 219–26, 2003.

[27] J. Dixon, "The effect of obesity on health outcomes," *Molecular and Cellular Endocrinology*, vol. 316, no. 2, pp. 104–108, 2010.

[28] National Center for Health Statistics. National Health and Nutrition Examination Survey. Questionnaires, datasets, and related documentation. [Online]. Available: http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm

[29] C. Ogden, M. Carroll, B. Kit, and K. Flegal, *Prevalence of obesity in the United States, 2009-2010*. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, 2012.

[30] U.S. Department of Health and Human Services. Final review, Healthy People 2010: Nutrition and overweight. [Online]. Available: http://www.cdc.gov/nchs/data/hpdata2010/hp2010_final_review_focus_area_19.pdf

[31] E. Finkelstein, O. Khavjou, H. Thompson, J. Trogdon, L. Pan, B. Sherry, and W. Dietz, "Obesity and severe obesity forecasts through 2030," *American Journal of Preventive Medicine*, vol. 42, no. 6, pp. 563–570, 2012.

[32] T. O'Connell, "An overview of obesity and weight loss surgery," *Clinical Diabetes*, vol. 22, no. 3, pp. 115–120, 2004.

[33] J. Colquitt, J. Picot, E. Loveman, and A. Clegg, *Surgery for obesity*, 2nd ed. John Wiley & Sons, Ltd, 2009.

[34] M. Huizinga, "Weight-loss pharmacotherapy: A brief review," *Clinical Diabetes*, vol. 25, no. 4, pp. 135–140, 2007.

[35] H. Kissileff, G. Klingsberg, and T. Van Itallie, "Universal eating monitor for continuous recording of solid or liquid consumption in man," *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, vol. 238, no. R14-R22, January 1980.

[36] M. Westerterp-Plantenga, "Eating behavior in humans, characterized by cumulative food intake curvesa review." *Neuroscience and Biobehavioral Reviews*, vol. 24, pp. 239–248, 2000.

[37] M. Westerterp-Plantenga, K. Westerterp, N. Nicolson, A. Mordant, P. Schoffelen, and F. ten Hoor, "The shape of the cumulative food intake curve in humans, during basic and manipulated meals." *Physiol Behav*, vol. 47, pp. 569–576, 1990.

[38] M. Westerterp-Plantenga, L. Wouters, and F. ten Hoor, "Deceleration in cumulative food intake curves changes in body temperature and diet-induced thermogenesis." *Physiol Behav*, vol. 48, pp. 831–836, 1990.

[39] M. Westerterp-Plantenga, L. Wouters, and F. ten Hoor, "Restrained eating obesity, and cumulative food intake curves during four course meals." *Appetite*, vol. 16, pp. 149–158, 1991.

[40] M. Westerterp-Plantenga, E. van den Heuvel, L. Wouters, and F. ten Hoor, "Diet induced thermogenesis and cumulative food intake curves, as a function of familiarity with food and dietary restraint in humans." *Pysiol Behav*, vol. 51, pp. 457–465, 1992.

[41] A. Petty, K. Melanson, and G. Greene, "Self-reported eating rate aligns with laboratory measured eating rate but not with free-living meals," *Appetite*, vol. 63, pp. 36–41, 2012.

[42] T. Spiegel, T. Wadden, and G. Foster, "Objective measure of eating rate during behavioral treatment of obesity." *Behavior Therapy*, vol. 22, pp. 61–67, 1991.

[43] A. Andrade, G. Greene, and K. Melanson, "Eating slowly led to decreases in energy intake within meals in healthy women." *Journal of the American Dietetic Association*, vol. 108, pp. 1186–1191, 2008.

[44] C. Martin, D. Williamson, P. Geiselmana, H. Walden, M. Smeets, S. Morales, and S. Redmann Jr., "Consistency of food intake over four eating sessions in the laboratory," *Eating Behaviors*, vol. 6, pp. 365–372, 2005.

[45] J. Fisher, B. Rolls, and L. Birch, "Children's bite size and intake of an entree are greater with large portions than with age-appropriate or self-selected portions," *The American Journal of Clinical Nutrition*, vol. 77, no. 5, pp. 1164–1170, 2003.

[46] K. Burger, J. Fisher, and S. Johnson, "Mechanisms behind the portion size effect: Visibility and bite size," *Obesity*, vol. 19, no. 3, pp. 546–551, 2010.

[47] Z. Huang, "An Assessment of the Accuracy of an Automated Bite Counting Method in a Cafeteria Setting," Master Thesis, College of Engineering and Science, Clemson University, 2013.