Detecting Periods of Eating During Free-Living by Tracking Wrist Motion

Yujie Dong[1], Jenna Scisco[2], Mike Wilson[2], Eric Muth[2] Adam Hoover[1]
[1]Electrical and Computer Engineering Department [2]Psychology Department
Clemson University
Clemson, USA
dongyujie@gmail.com, ahoover@clemson.edu

***Abstract*—This work is motivated by the growing prevalence of obesity, a health problem affecting over 500 million people. Measurements of energy intake are commonly used for the study and treatment of obesity. However, the most widely used tools rely upon self-report and require a considerable manual effort, leading to underreporting of consumption, non-compliance, and discontinued use over the long term. The purpose of this paper is to describe a new method that uses a watch-like configuration of sensors to continuously track wrist motion throughout the day and automatically detect periods of eating. Our method uses the novel idea that meals tend to be preceded and succeeded by periods of vigorous wrist motion. We describe an algorithm that segments and classifies such periods as eating or non-eating activities. We also evaluate our method on a large data set (43 subjects, 449 total hours of data, containing 116 periods of eating) collected during free-living. Our results show an accuracy of 81% for detecting eating at 1 second resolution in comparison to manually marked event logs of periods eating. These results indicate that vigorous wrist motion is a useful indicator for identifying the boundaries of eating activities, and that our method should prove useful in the continued development of body-worn sensor tools for monitoring energy intake.**

***Index Terms*—obesity, energy intake, activity recognition, body motion tracking, accelerometer, gyroscope**

I. Introduction

This work is motivated by the growing prevalence of obesity. The World Health Organization reports that in 2008, 1.4 billion adults (age 20+) were overweight (body mass index > 25) and 500 million adults were obese (BMI > 30) [41]. Reports for 2012 show that one in three adults and one in six children in the United States were obese [13], [25]. Obesity is a major risk factor for diabetes, heart disease, high blood pressure, stroke and cancer [39]. Deaths attributed to obesity continue to increase [12], [23]; 65% of the world's population lives in countries where more people die from complications due to overweight or obesity than from complications due to underweight [41].

Energy expenditure (EE) and energy intake (EI) are commonly used measurements in the study and treatment of obesity [30]. The former measures the energy cost of homeostasis (body maintenance) plus physical activities, the latter measures consumption. Many studies have shown that self-reported estimates of EE suffer from bias [8], [20], [28]; body-worn motion sensors provide more objective measurements with less user burden at less cost [40]. Similarly, numerous studies have shown that people tend to underreport their EI using self-report methods, with estimates of underreporting ranging from 10-30% for normal weight subjects to 20-50% for obese adults and children [6], [14], [16], [19], [24], [32], [37]. The goal of our research is to develop body-worn sensing methods that can objectively measure EI. The need for such tools has been widely advocated within the dietetics community [21], [36] and in funding programs from the US National Science Foundation and National Institutes of Health [11].

The purpose of this paper is to describe a new method that uses a watch-like configuration of sensors to continuously track wrist motion throughout the day and automatically detect periods of eating. The problem of using body-worn sensors to automatically measure EI may be broken into two parts. The first part is identifying periods of consumption amongst all daily activities. The second part is estimating EI during those periods. Previous research has focused on the second part of the problem, such as counting the numbers of chews [31], swallows [27], [29], drinks [4], bites [9], or specific eating gestures [2], [3]. This paper is the first to describe a method that detects entire periods of eating (e.g. meals and snacks) during all-day tracking. It is also significant that we evaluated our method on a data set that was collected during free-living as opposed to in a laboratory environment, so that our method could be tested on unscripted eating behaviors.

II. Methods

Our method assumes a person is wearing a watch-like configuration of accelerometers and gyroscopes, as depicted in Figure 1. The sensors track the linear and rotational motion of the wrist. We have discovered that prior to an eating activity (e.g. a meal/snack), there tends to be a period of larger wrist motion energy, caused by things like bringing food to a table, adjusting the position of utensils, opening food containers, and unwrapping food. During an eating activity, the total wrist motion energy tends to be reduced. At the end of an eating activity, there tends to be another period of larger wrist motion energy, caused by things like putting remaining food away, washing hands, standing up, and putting dishes away. We have designed an algorithm that uses this idea to detect periods of eating. It calculates a continuous estimate of wrist motion energy and uses a hysteresis-based peak detector to segment periods of time in-between vigorous motions. For each segmented period, features are calculated and used to

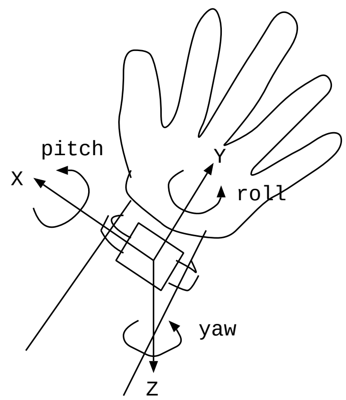

Fig. 1. A watch-like configuration of accelerometers and gyroscopes tracks wrist motion continuously throughout the day (shown here on the right hand).

classify the period as an eating or non-eating activity. We first describe the details of the algorithm. We then describe the data collected and the evaluation metrics used to determine the efficacy of this approach.

A. Algorithm

1) Preprocessing: Data from the sensors are first smoothed to reduce the effects of noise:

$$S_t = \sum_{i=-N}^{0} R_{t+i} \frac{\exp(\frac{-t^2}{2R^2})}{\sum_{x=0}^{N} \exp(\frac{-(x-N)^2}{2R^2})} \tag{1}$$

where R_t is the raw datum and S_t is the smoothed datum at time t. Equation 1 implements a Gaussian-weighted window centered on the current measurement, so that only half of a Gaussian distribution is used for smoothing. The variable N is a window size and R is the sigma of the Gaussian. The particular values used for these variables will depend upon the quality of the sensors used; we provide values for our testing device later. Equation 1 is applied independently to the data from each accelerometer and gyroscope axis.

2) Segmentation: Wrist motion energy can be characterized by the total amount of motion. We tested both the sum of accelerometer readings and the sum of gyroscope readings, finding similar results [10]. Because accelerometers use approximately one-tenth the power of gyroscopes [34], [35], it is preferable to use accelerometers for continuous all-day monitoring. We therefore calculate wrist motion energy as

$$E_t = \frac{1}{W+1} \sum_{i=t-\frac{W}{2}}^{t+\frac{W}{2}} |S_{x,t}| + |S_{y,t}| + |S_{z,t}| \tag{2}$$

where $S_{x,t}$, $S_{y,t}$ and $S_{z,t}$ are the smoothed acceleration readings at time t. The parameter W is a window size; we have found a sliding 1 minute window to be sufficient for smoothing over brief vigorous motions while still capturing longer vigorous motions indicative of the boundaries of eating activities [10].

The following example demonstrates the presence of vigorous wrist motions before and after eating, and is helpful for explaining the algorithm. Figure 2 shows the wrist motion

```
loop t (data index)
  reset (T1,T2)
  while (E[t] < T2)
    t=t+1
    if (E[t] < T1)
      reset (T1,T2)
  end while
  while (E[t] > T1)
    t=t+1
  end while
end loop
```

Fig. 3. Pseudocode for peak detector used to segment data.

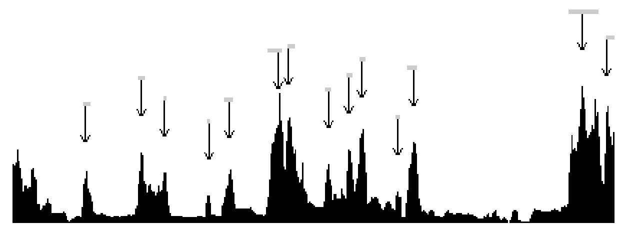

Fig. 4. Detected peaks on the first 2 hours of data from figure 2. Arrows indicate the points used for segmentation, lines above arrows indicate the spans of the detected peaks.

energy of a person over a 12 hour period (the Y-axis is clipped to save space). The start and stop times of the meals/snacks shown in the figure were manually logged by the person being recorded. It can be seen that all 4 eating activities show a pronounced peak before and after eating. Of course, other peaks occur throughout the day, so this feature alone cannot be used for classification. But it does provide a reasonable mechanism for segmenting the data.

To automatically identify peaks we developed a custom peak detector using the concept of a hysteresis threshold [33]. Our detector identifies peaks at local maxima that are sufficiently pronounced while suppressing marginal local maxima. Pseudocode for the algorithm is given in Figure 3. The algorithm loops through the data from beginning to end, with each pass through the two while loops identifying a single peak. The two thresholds $T1$ and $T2$ are set equal to the value of the signal (wrist motion energy) at the current index, and two times that value. The first while loop iterates until the signal exceeds the second (larger) threshold, in essence requiring the signal to go 2x above its previously observed minimum. During this search, if a signal value is found that is lower than $T1$, then the required thresholds are recalculated. Once the signal has exceeded the second threshold, the second while loop iterates until the signal falls below the first (lower) threshold. The index of the peak is taken as the location with the maximum signal value found during the two while loops. Figure 4 shows the result of the peak detector on the first 2 hours of data from Figure 2.

The indices of the detected peaks are used to segment the data. We have noticed that sometimes, a meal/snack can have a peak inside the period of eating. An example of this can be seen in the period labeled “dinner” in Figure 2. This is

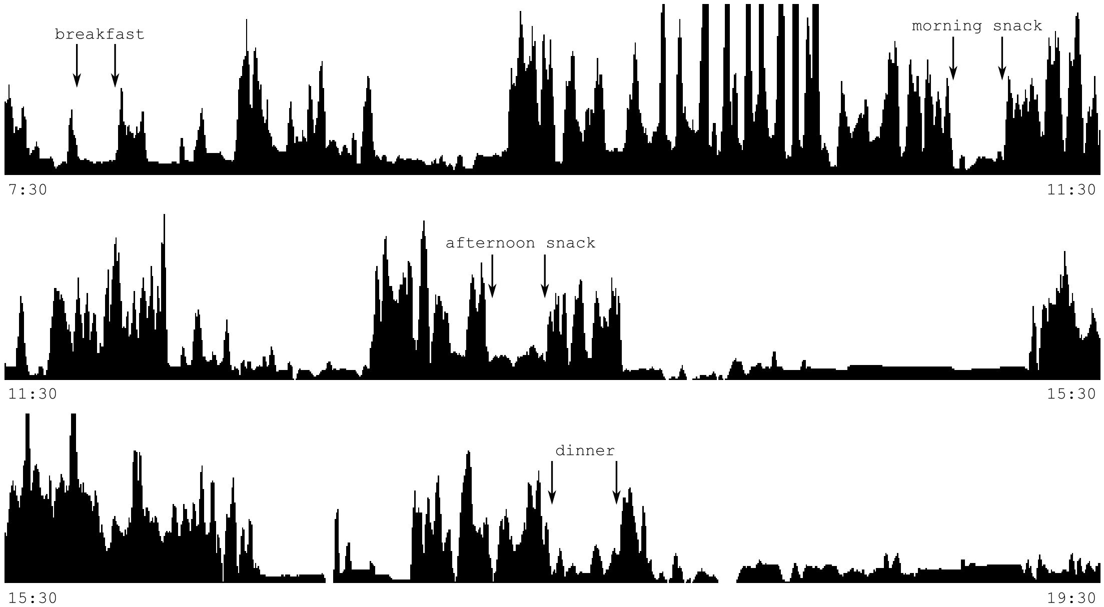

Fig. 2. An example of accelerometer-based wrist motion energy of a person over a 12 hour period. Manually logged meal times are marked.

likely caused by the person conducting activities like extended application of condiments, or preparing a second course. Since the wrist motion energy is calculated over a sliding 1 minute window, brief periods of intense motion such as single gestures will not trigger our segmentation algorithm. In the case of a longer vigorous activity, the result is that the period of eating is oversegmented. If desired, this could be overcome by merging consecutive segmented periods after classification.

3) Features: We investigated numerous features for classification [10]; this paper reports on the 4 features found to be most useful. Each feature is calculated over each inter-peak segmented period. We refer to the first feature as manipulation. It is calculated as:

$$f_{1,w} = \frac{1}{W} \sum^{W} \frac{|S_{\phi,t}| + |S_{\theta,t}| + |S_{\psi,t}|}{|S_{x,t}| + |S_{y,t}| + |S_{z,t}|} \tag{3}$$

where W is the span of the segmented period, t is the index that iterates across that span, and S is the smoothed datum (ϕ, θ, ψ = yaw, pitch, roll). This feature measures the ratio of rotational motion to linear motion. The second feature is linear acceleration, and is calculated as:

$$f_{2,w} = \frac{1}{W} \sum^{W} |S_{x,t}| + |S_{y,t}| + |S_{z,t}| \tag{4}$$

The third feature is the amount of wrist roll motion, and is calculated as:

$$f_{3,w} = \frac{1}{W} \sum^{W} |S_{\psi,t} - \frac{1}{W} \sum^{W} S_{\psi,t}| \tag{5}$$

The fourth feature is the regularity of wrist roll motion, and is calculated as:

$$f_{4,w} = \frac{1}{W} \int_{W} 1 \; \forall \; t \in [|S_{\psi,t}| > 10^{\circ} \; ... \; t + 8\text{sec}] \tag{6}$$

This feature takes on a value between 0 and 1, representing the percentage of time that the wrist is in roll motion. The calculation includes the time the wrist roll is at least 10 deg/sec, plus a period of 8 sec after each occurrence of wrist roll motion falling below 10 deg/sec. The latter two features are inspired by our previous work [9] in which wrist roll was used to detect bites during eating. The values 10 deg/sec and 8 sec were found to be optimal for characterizing a typical bite motion and interval.

4) Classification: For classification we used a naive Bayes classifier [22]. The Bayesian approach to classification is to assign the most probable class $c_i \in C$, given feature values $f_1, f_2, ..., f_N$. Using the naive assumption of independence of features, the classification problem can be written as:

$$c_i = \underset{C}{argmax} \, P(c_i) \prod_{j} P(f_j | c_i) \tag{7}$$

For our problem there are only two classes, eating (c_0) and non-eating (c_1). We set each $P(c_i) = 0.5$. We modeled the probabilities of each feature given each class using a normal distribution:

$$P(f_j|c_i) = \frac{1}{\sqrt{2\pi\sigma_{i,j}^2}} \exp\left(-\frac{(f_j - \mu_{i,j})^2}{2\sigma_{i,j}^2}\right) \tag{8}$$

where $\mu_{i,j}$ and $\sigma_{i,j}^2$ are the mean and variance of feature j for class i.

B. Data collection

An iPhone 4 (Apple Inc., 1 Infinite Loop, Cupertino, CA 95014, http://www.apple.com/iPhone/) was used to collect data to develop and evaluate our algorithm. This device was chosen because it is programmable, equipped with the appropriate sensors, and has a sufficiently large memory (16GB) and battery (1420 mAh) to record continuous data for an entire day. Commercial activity monitors exist in the form of wrist watches, but they only contain accelerometers (no gyroscopes),

Fig. 5. Data collection using an iPhone 4 on the wrist.

and so could not be used for this work. Although the iPhone is larger than a watch, it is important to note that a much smaller device could be constructed; this is discussed further in section IV.

The iPhone was placed inside a pouch which could be wrapped snugly around the forearm (see Figure 5). The top of the device was aligned with the wrist joint but positioned so that it would not inhibit movement of the wrist. A custom program was written to run on the iPhone, recording the raw data for later transfer to a computer through a USB port. Our segmentation and classification algorithms were implemented in the C programming language using a Win32 graphical user interface to visualize the data and results. For smoothing sensor data from the iPhone, we found that a window size N of 1 second with a Gaussian sigma R of 10 produced good results.

The Clemson University Institutional Review Board approved the data collection and each subject provided informed consent. Subjects were given the device in a brief laboratory visit prior to the day of their recording, and were instructed in its use. They were asked to put the device on and start the custom program soon after waking in the morning, and to conduct all activities throughout the day as naturally as possible while the device continuously recorded their wrist motion. Subjects were asked to remove the device only when engaging in activities that could damage it, such as taking a shower. On the day following recording, each subject returned the iPhone to the experimenter for data download and review.

Data was collected in two separate batches. In the first batch, 30 subjects (12 male, 18 female, ages 18-32) were instructed to manually write down the start and stop times of their actual meals and snacks in a provided log book, using the time displayed on the device for reference. The iPhone program recorded data at 60 Hz and drained the battery after approximately 8.5 hours. Subsequent to this batch, we learned that 15 Hz data was sufficient for our method, and were able to extend recording time to approximately 12 hours. We also learned that subjects had trouble using the provided written log to record the times of eating. We therefore discontinued using the manually written log and instead added an event marker button to the iPhone program that subjects were instructed to press when they started and ended meals or snacks. We also removed the function from the program that allowed participants to halt/resume recording, to avoid confusion. In the second batch of data collection, 25 subjects (8 male, 17 female, ages 20-50) used the updated version of the iPhone program.

During post-review, the experimenter interviewed each subject to identify possible errors for exclusion. Out of the first batch of 30 recordings, 10 had to be discarded due to poor compliance with keeping records. Two subjects forgot to write down start or stop times for 1 or more meals/snacks. Three subjects stated that they filled the log out at the end of the day based upon memory, instead of writing down the start and stop times as they occurred. Five subjects misinterpreted our instructions and started/stopped the iPhone recording program for meals only. These problems motivated the reprogramming of the iPhone recording program to remove the halt/resume button, and to include an event marker on the iPhone screen in place of using a written time log. For the second batch of 25 recordings, button press logs were reviewed with subjects the day after recording to eliminate inadvertent markers. Out of 294 total marks, 172 recorded 86 discrete, verified eating activities with event marks at the actual start and stop boundaries as verified by the participants. Most of the remaining 122 marks were identifiable as inadvertent due to being single marks (as opposed to marks that could be paired into start/stop sets). Given the sensitivity of the iPhone touchscreen, the size of the event marker button (5×2 cm), and the fact that each subject wore the device for a whole day, this number of inadvertent presses was not surprising. Nine marks were reported as intentional by the subjects to test that the device was still recording, but were not associated with meals. Six marks were identified as double presses of the button due to being less than 10 seconds apart. Two subjects reported forgetting to press the button at the end of one or more meals; these recordings were discarded.

In total, our data collection yielded 449 hours of data from 43 subjects, including a cumulative 22.4 hours of eating over 116 total meals/snacks. It is important to note that the goal of the data collection was to capture a sample of eating activities covering a variety of individuals, meals, environments and times of day. The purpose of the data set was to enable algorithm development for automatically detecting such periods. It was not a goal of the data collection to capture total daily intake. We asked that participants try to capture all their eating activities, but the goal of this work was not contingent on meeting this criteria.

C. Evaluation

We used two sets of evaluation metrics. The first metrics evaluate the classifier by the total amount of time correctly classified, the second metrics evaluate the classifier by the total amount of eating activities (segments) correctly classified. The boundaries of manually logged periods of eating were recorded at 1 second resolution. The boundaries of automatically classified periods of time were rounded to the nearest second.

For the first metrics, true positives (TP) were counted as the number of seconds of time that were labeled as eating in the manual logs and classified as eating. False positives (FP) were counted as the number of seconds of time that were labeled as non-eating in the manual logs and classified

notation	feature	eating		non-eating	
		mean	var	mean	var
f_1	manipulation ((deg/sec)/G)	791	45785	395	57284
f_2	acceleration (G)	0.039	0.0002	0.054	0.0043
f_3	roll motion (deg/sec)	9.1	18.2	6.8	39.2
f_4	roll regularity (%time)	0.58	0.02	0.37	0.07

TABLE I
AVERAGE FEATURE VALUES FOUND DURING TRAINING.

as eating. True negatives (TN) and false negatives (FN) were counted similarly by comparing the manual log labels to the data classified as non-eating. Sensitivity and specificity were calculated as TP/(TP+FN) and TN/(TN+FP). Accuracy was calculated as:

$$\text{accuracy} = \frac{\text{TP} \times 20 + \text{TN}}{(\text{TP + FN}) \times 20 + (\text{TN + FP})} \tag{9}$$

The factor of 20 in equation 9 weights true positives to true negatives at a ratio of 20:1. This is used because eating occurs much less frequently than non-eating in general free-living. The importance of using 20:1 weighting during evaluation is demonstrated in our results and further discussed in section IV.

For the second metrics, consecutive segments that were labeled as eating by the classifier were merged into single whole segments (see end of section II-A2). True detections were counted as the number of manually logged entries that overlapped segments that were labeled eating by the classifier. Undetected eating activities were counted as the remainder of the manually logged entries. False detections were counted as the remainder of the segments that were labeled eating by the classifier.

III. RESULTS

The classifier was trained using leave-one-out cross-validation. Thus, for testing each of the 43 recordings (1 per person), the classifier was trained using the other 42 recordings to calculate values for the classifier probabilities ($\mu_{i,j}$ and $\sigma_{i,j}$). Table I lists the average means and variances for the features for each class. We use the notation f_1 to refer generically to the feature manipulation (see equation 3), and $f_{1,w}$ to refer to that feature calcualted over a specific window W. The probabilities for the eating class were calculated as the average feature values for all segments labeled as eating by the subjects. For the non-eating class, all the remaining data from the recordings was broken into 5 minute windows and the probabilities were calculated as the average feature values. As can be seen in Table I, during eating there tends to be higher values for manipulation, roll motion and roll regularity, and lower values for linear acceleration. The variances for all features for the non-eating class are higher than for eating, due to the variety of activities grouped together in this class. Two-tailed independent t-tests comparing all paired distributions showed the differences are statistically significant (all p's $<$ 0.001).

Table II shows the results of testing the classifier using leave-one-out cross validation combined across the two batches of data collected. The accuracy achieved was 79%. Using just the first two features (manipulation and linear acceleration) produced the same accuracy as using all 4 features. However, since our data was recorded in two batches, we also analyzed the results using separate leave-one-out cross validation for each batch. Specifically, for the 20 recordings in the first batch, each was tested using the other 19 for training the classifier; for the 23 recordings in the second batch, each was tested using the other 22 for training the classifier. Table III shows these results. In this case an accuracy of 81% was achieved, and the use of all 4 features improved sensitivity, specificity, and accuracy. We hypothesize that this is due to the different amounts of data recorded in each batch. The first batch averaged 8.5 hours per recording, spanning 10AM to 6:30PM on average. The second batch averaged 12 hours per recording, spanning 10AM to 10PM on average. Since the second set included more evening and night activities, we suppose that the different training produced more accurate feature values for the classifier.

features	sensitivity	specificity	accuracy
f_1, f_2	80%	79%	79%
f_1, f_2, f_3, f_4	76%	82%	79%

TABLE II
RESULTS USING LEAVE-ONE-OUT CROSS VALIDATION ON ALL DATA.

features	sensitivity	specificity	accuracy
f_1, f_2	78%	79%	79%
f_1, f_2, f_3, f_4	81%	82%	81%

TABLE III
RESULTS USING LEAVE-ONE-OUT CROSS VALIDATION SEPARATELY ON EACH OF THE TWO BATCHES OF DATA.

The values $P(c0), P(c1)$ in our classifier (see section II-A4) determine the likelihood of a segment being classified as an eating or non-eating activity. We tested across the range of values $P(c0), P(c1) = \{0, 1\}, \{.05, .95\}, \{.1, .9\}, ...\{1, 0\}$ to find the maximum accuracy. Table IV shows the importance of weighting accuracy at 20:1 for evaluating total time correctly classified. Weighted at 1:1, our classifier achieves a maximum accuracy of 95% but this occurs at a sensitivity of 0%, in other words when all data are labeled as non-eating. Weighted at 20:1, our classifier achieves a maximum accuracy of 81% which is less than 95%, but the sensitivity and specificity are balanced.

The accuracy per person ranged from 35-97%, with a median of 82%. The accuracy was above 70% for 38 out of

weighting	sensitivity	specificity	accuracy
1:1	0%	100%	95%
20:1	81%	82%	81%

TABLE IV
EVALUATION OF CLASSIFIER AT MAXIMUM ACCURACY USING 1:1 VERSUS 20:1 WEIGHTING OF TIME CORRECTLY CLASSIFIED AS EATING VERSUS NON-EATING.

43 people (88%), with the five remaining having much lower accuracies. This suggests that for most people, our method may be suitable for detecting eating activities, but that for some people our method may not work. It could be that some people are less likely to engage in vigorous wrist motions before and after eating activities. However, each subject was only recorded for a single day, so this result may be more a function of the particular meals/snacks eaten on the day of recording than individual habits.

We also evaluated our classifier at the segment level using the second set of metrics described in section II-C. The classifier correctly detected 100 actual eating activities, missed 16, and had 379 false detections. The average time between the start of manual log entries and correctly detected eating activities was -0.6 minutes. The average time between the end of manual log entries and correctly detected eating activities was +1.5 minutes. This suggests that the peaks detected by our method occur slightly before and slightly after actual eating begins and ends, respectively.

IV. DISCUSSION

Table V summarizes the present study in comparison to related works. The first contribution of this paper is that it is the first to describe a method to detect entire periods of eating (i.e. meals, snacks) as opposed to counting individual swallows, chews, bites or specific gestures during eating. For example, Amft and colleagues studied the recognition of four different gestures related to eating: using a fork and knife to eat from a plate, using a spoon to eat from a bowl, using hands to eat, and drinking from a glass [3], [17]. In a related work they recognized gestures specific to 11 food categories [2], and in a recent study they detected drinking gestures (sip and fetch motions) [4]. Sazonov and colleagues studied the recognition of chews [31] and swallows [29]. Päßler and colleagues developed a method to recognize swallows associated with different types of foods [27]. Our group developed a method to recognize and count bites of food and sips of liquid taken during a meal [9]. All these methods make progress towards the goal of using body-worn sensors to automatically measure EI. However, when operated all day, all these methods face the challenge of false positives occuring during non-eating activities. The method described in this paper could be used as an automated on/off switch, activating any of these methods only during a detected meal/snack. This has the potential to greatly reduce the incidence of false positives in all-day automated EI measures.

A second contribution of this paper is that we tested our method on data collected during free-living, as compared to scripted eating activities in the lab. For example, most previous works used between 4 and 11 specific foods and directed subjects through a scripted sequence of eating and rest activities [2], [3], [17], [27], [29], [31]. In one study designed to detect drinking gestures, each subject was recorded in an approximately one hour session that included scripted activities of office work, gaming, and leisure [4]. In our previous work that studied an automated bite counting method, no restrictions were placed on food or beverage types, but all data was still collected in a lab [9]. In contrast, for this work each of our subjects was recorded for a day (10.4 hours on average) during normal daily free-living, and we placed no restrictions on eating behavior. To our knowledge this is the first study using body-worn sensors to detect eating activities in free-living conditions. Although lab studies offer a controlled environment in which eating period detection could be objectively confirmed by video recording or direct observation, body-worn sensors that detect eating are designed with the ultimate intention of being used in free-living. Free-living studies face the challenge of recording the actual activities of subjects in order to evaluate the automated methods, as seen in this paper. However, event markers are commonly used in mobile physiological monitors, such as Holter EKG monitors, for a wearer to note a significant event, e.g. a panic attack. Hence, event markers are an accepted approach to identifying significant behavioral events from among other free living activities. The advantage of studying eating in free-living is that eating behaviors, schedules and activities are as natural as possible.

A third contribution of this paper is that we tested our method on the largest data set (449 total hours) that has been reported in the related literature. This is partly due to the fact that eating occurs much less frequently than non-eating in free-living. In our data set we observed 22.4 hours of eating out of 449 total hours, a ratio of 1 to 20. In contrast, previous works that used data collected in the lab were based on an unrealistic equal ratio of eating and non-eating data [3], [27], [29]. This confounds comparisons of accuracies between these methods and the results reported in this paper. As shown in our results, using equal weighting for eating and non-eating data achieves 95% accuracy for a classifier that blindly labels all data as non-eating; this is obviously not desirable. We achieved 81% accuracy classifying 1 second epochs as eating or non-eating at a more realistic 20:1 weighting that more closely conforms to actual behavior. Contrasting our results against previous works, Sazonov and colleagues achieved 97% accuracy classifying 1.5 second epochs as containing swallows or not, 85% accuracy detecting individual swallows, and 81% accuracy classifying 30 second epochs as containing chews or not [29], [31]. Päßler and colleagues achieved 83% accuracy detecting swallows and 79% accuracy recognizing the food type swallowed [27]. Amft and colleagues achieved 80% accuracy recognizing 11 different foods being eaten by a single subject [2], and 94% precision with 84% recall recognizing drinking motions of 6 subjects [4]. Our previous work achieved 86% sensitivity at 81% positive predictive value at detecting bites during meals [9]. However, we reiterate that all the related works weighted the detection of eating activities vs non-eating activities at a 1:1 ratio, whereas we weighted 20:1.

study	activities detected	test environment	total subjects	total hours	sensor location
[3], [17]	gestures	lab, 4 foods	4	5	arm, wrist, back
[2]	gestures	lab, 11 foods	1	-	neck, ear, wrist, arm
[4]	drinks	lab	6	6	wrist
[29]	swallows	lab, 5 foods	20	65	neck, throat
[31]	chews	lab, 5 foods	20	65	jaw
[27]	swallows	lab, 8 foods	51	21	jaw
[9]	bites	lab	47	14	wrist
present	meals/snacks	free-living	43	449	wrist

TABLE V
SUMMARY OF DISCUSSION COMPARING THIS WORK TO RELATED WORKS.

It also bears repeating that the related works limited eating conditions by collecting data in the lab.

The envisioned embodiment of our device is a small watch worn on the wrist. Compared to some related works, this configuration is simpler, and potentially less embarrassing (compared for example to head mounted sensors), which has implications for long-term use in free-living. We used an iPhone to collect data for algorithm development and evaluation so that we could record a full day of raw data. However, in practice our method would only need to store raw data until the current data segment was classified, greatly reducing the needed memory. It would also only need to power gyroscopes during data segments suspected of containing eating activity. This is important because a MEMS accelerometer uses approximately 10% of the power of a MEMS gyroscope [34], [35]; a typical coin-sized battery can power a single MEMS gyroscope for part of one day, while it can power an accelerometer for over one week.

Another application for the method described in this paper is the automated measurement of a daily pattern of eating activities. Daily patterns are known to be associated with variations in EI [18]. For example, night eating syndrome (NES) is characterized by evening hyperphagia and morning anorexia [5]. Different studies have found varying links between NES and obesity and other disorders [7]. One factor inhibiting study is the difficulty of objectively quantifying and measuring diagnostic criteria involving eating patterns [1]. Binge eating disorder [15] and eating disorders linked to night shift working [38] are other problems associated with temporal eating patterns. The method described in this paper has the potential to provide an objective eating activity calendar for studying these types of problems.

A limitation of our evaluation is that the classifier was trained at the group level. We hypothesize that individual-level training could improve these results. However, we only recorded 1 day of data per person, with an average of $116/43 = 2.7$ periods of eating per person. In future work we would like to collect free-living data from individuals over a longer duration (e.g., a week or more). This would allow the classifier to be trained to the individual, perhaps improving performance. This would also allow us to further explore the idea that our method may not be suitable for some people who do not exhibit vigorous wrist motion at the boundaries of eating activities. Another limitation of our method is that it groups all non-eating activities into a single class. It may be that a multi-class approach with a more sophisticated classifier would achieve a higher recognition accuracy. These ideas are all subjects for future work, for which the current work provides the foundation that the detection of eating boundaries is feasible based on a relatively simple wrist sensor.

REFERENCES

[1] K. Allison, et. al., “Proposed Diagnostic Criteria for Night Eating Syndrome”, in *Int’l J. of Eating Disorders*, vol. 43 no. 3, 2010, pp. 241-247.

[2] O. Amft, M. Kusserow and G. Troster, “Probabilistic parsing of dietary activity events”, in *Proc. of Int’l Workshop on Wearable and Implantable Body Sensor Networks*, 2007, pp. 242-247.

[3] O. Amft, and G. Troster, “Recognition of Dietary Activity Events Using On-body Sensors”, in *Artificial Intelligence in Medicine*, vol. 42, 2008, pp. 121-136.

[4] O. Amft, D. Bannach, G. Pirkl, M. Kreil and P. Lukowicz, “Towards Wearable Sensing-Based Assessment of Fluid Intake”, in *Proc of IEEE 8th Int’l Conf on Pervasive Computing and Communications Workshops*, 2010, pp. 298-303.

[5] G. Birketvedt, et. al., “Behavioral and Neuroendocrine Characteristics of the Night-Eating Syndrome”, in *J. of the American Medical Assoc.*, vol. 282 no. 7, 1999, pp. 657-663.

[6] C. Champagne, G. Bray, A. Kurtz, J. Monteiro, E. Tucker, J. Volaufova and J. Delany, ”Energy Intake and Energy Expenditure: A Controlled Study Comparing Dieticians and Non-Dieticians”, in *Journal of the American Dietetic Association*, vol. 102, no. 10, 2002, pp. 1428-1432.

[7] J. Cleator, J. Abbott, P. Judd, C. Sutton and J. Wilding, “Night eating syndrome: implications for severe obesity”, in *Nutrition and Diabetes*, vol. 2 no. 9, 2012.

[8] J. Conway, M. Irwin and B. Ainsworth, “Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires - A doubly labeled water validation”, in *J. of Clinical Epidemiology*, vol. 55 no. 4, 2002, pp. 293-299.

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

[10] Y. Dong, “Tracking Wrist Motion to Detect and Measure the Eating Consumption of Free-Living Humans”, PhD dissertation, Electrical & Computer Engineering Dept, Clemson University, 2012.

[11] A. Ershow, A. Ortega, J. Baldwin, and J. Hill, “Engineering approaches to energy balance and obesity: Opportunities for novel collaborations and research: Report of a joint National Science Foundation and National Institutes of Health workshop”, in *Journal of Diabetes Science and Technology*, vol. 1 no. 1, 2007, pp. 96-105.

[12] K. Flegal, B. Graubard, D. Williamson, and M. Gail, “Excess deaths associated with underweight, overweight, and obesity”, in *Journal of the American Medical Association*, vol. 293 no. 15, 2005, pp. 1861-1867.

[13] K. Flegal, M. Carroll, B. Kit and C. Ogden, “Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010”, in *Journal of the American Medical Association*, vol. 307 no. 5, 2012, pp. 491-497.

[14] K. Glanz, J. Brug and P. van Assema, “Are Awareness of Dietary Food Intake and Actual Fat Consumption Associated? A Dutch-American Comparison”, in *European Journal of Clinical Nutrition*, vol. 51, 1997, pp. 542-547.

[15] A. Harb et. al., “Night eating patterns and chronotypes: A correlation with binge eating behaviors”, in *Psychiatry Research*, vol. 200 no. 2, 2012, pp. 489-493.

[16] S. Jonnalagadda, D. Mitchell, H. Smiciklas-Wright, K. Meaker, N. Heel, W. Karmally, A. Ershow and P. Mkris-Etherton, “Accuracy of Energy Intake Data Estimated by a Multiple-pass, 24-hour Dietary Recall Technique”, in *Journal of the American Dietetic Association*, vol. 100, no. 3, 2000, pp. 303-311.

[17] H. Junker, O. Amft, P. Lukowicz, and G. Troster, “Gesture Spotting with Body-worn Inertial Sensors to Detect User Activities”, in *Pattern Recognition*, vol. 41, no. 6, 2008, pp. 2010-2024.

[18] J. Kerver, E. Yang, S. Obayashi, L. Bianchi and W. Song, “Meal and Snack Patterns Are Associated with Dietary Intake of Energy and Nutrients in US Adults”, in *J. of the Amer. Dietetic Assoc.*, vol. 106 no. 1, 2006, pp. 46-53.

[19] S. Lichtman, K. Pisarska, E. Berman, M. Pestone, H. Dowling, E. Offenbacher, H. Weisel, S. Heshka, D. Matthews, and S. Heymsfield, “Discrepancy between Self-reported and Actual Caloric Intake and Exercise in Obese Subjects”, in *the New England Journal of Medicine*, vol. 327, no. 27, 1992, pp. 1894-1898.

[20] S. Mahabir, et. al., “Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women”, in *Am. J. of Clinical Nutrition*, vol. 84 no. 1, 2006, pp. 230-236.

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

[22] T. Mitchell, ”Machine Learning”, McGraw Hill, 1997.

[23] A. Mokdad, J. Marks, D. Stroup, and J. Gerberding, “Actual causes of death in the United States, 2000”, in *Journal of the American Medical Association*, vol. 291 no. 10, 2004, pp. 1238-1245.

[24] L. Muhlheim, D. Allison, S. Heshka and S. Heymsfield, “Do Unsuccessful Dieters Intentionally Underreport Food Intake?”, in *International Journal of Eating Disorders*, vol. 24, 1998, pp. 259-266.

[25] C. Ogden, M. Carroll, B. Kit and K. Flegal, “Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010”, in *Journal of the American Medical Association*, vol. 307 no. 5, 2012, pp. 483-490.

[26] J. O’Reardon, et. al., “Circadian Eating and Sleeping Patterns in the Night Eating Syndrome”, in *Obesity Research*, vol. 12 no. 11, 2004, pp. 1789-1796.

[27] S. Päßler, M. Wolff and W. Fischer, “Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food”, in *Physiological Measurement*, vo. 33 no. 6, 2012, pp.1073-1093.

[28] E. Rush, M. Valencia and L. Plank, “Validation of a 7-day physical activity diary against doubly-labelled water”, in *Annals of Human Biology*, vol. 35 no. 4, 2008, pp. 416-421.

[29] E. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. Melanson and M. Neuman, “Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior”, in *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, 2010, pp. 626-633.

[30] E. Sazonov and S. Schuckers, “The Energetics of Obesity: A Review: Monitoring Energy Intake and Energy Expenditure in Humans”, in *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 1, 2010, pp. 31-35.

[31] E. Sazonov and J. Fontana, “A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing”, in *IEEE Sensors Journal*, vol. 12 no. 5, 2012, pp. 1340-1348.

[32] D. Schoeller, D. Thomas, et. al., “Self-report–based estimates of energy intake offer an inadequate basis for scientific conclusions”, in *Am. J. of Clinical Nutrition*, vol. 97 no. 6, 2013, pp. 1413-1415.

[33] M. Sonka, V. Hlavac and R. Boyle, “Image Processing, Analysis, and Machine Vision”, second edition, PWS Publishing, 1999.

[34] STMicroelectronics, “MEMS Inertial Sensor, LIS344ALH Accelerometer”, http://www.st.com/internet/analog/product/207281.jsp, retrieved Dec 5, 2011.

[35] STMicroelectronics, “MEMS Inertial Sensor, LPR410AL Gyroscope”, http://www.st.com/internet/analog/product/248621.jsp, retrieved Dec 5, 2011.

[36] F. Thompson, A. Subar, C. Loria, J. Reedy and T. Baranowski, “Need for Technological Innovation in Dietary Assessment”, in *Journal of the American Dietetic Association*, vol. 110, no. 1, 2010, pp. 48-51.

[37] J. Tooze, A. Subaret, F. Thompson, R. Troiano, A. Schatzkin, and V. Kipnis, “Psychosocial Predictors of Energy Underreporting in a Large Doubly Labeled Water Study”, in *American Journal of Clinical Nutrition*, vol. 79, 2004, pp. 795-804.

[38] J. Waterhouse, P. Buckley, B. Edwards and T. Reilly, “Measurement of, and Some Reasons for, Differences in Eating Habits Between Night and Day Workers”, in *Chronobiology Int’l*, vol. 20 no. 6, 2003, pp. 1075-1092.

[39] N. Wellman and B. Friedberg, “Causes and Consequences of Adult Obesity: Health, Social and Economic Impacts in the United States”, in *Asia Pacific Journal of Clinical Nutrition*, vol. 11, no. 667-751, 2002, pp. 705-709.

[40] K. Westerterp, “Assessment of physical activity: a critical appraisal”, in *European J. of Applied Physiology*, vol. 105 no. 6, 2009, pp. 823-828.

[41] World Health Organization Media Center, ”Obesity and Overweight”, http://www.who.int/mediacentre/factsheets/fs311/en/index.html, retrieved January 29, 2013.