Information Fusion 41 (2018) 37–47

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

I sense overeating: Motif-based machine learning framework to detect overeating using wrist-worn sensing

Shibo Zhang[a,*], William Stogin[a], Nabil Alshurafa[b]

[a] EECS, Northwestern University, USA
[b] Preventive Medicine Dept. & EECS, Northwestern University, USA

ARTICLE INFO

Article history:
Received 4 February 2017
Revised 19 June 2017
Accepted 3 August 2017
Available online 4 August 2017

Keywords:
Wrist-worn sensors
Wearables
Hand-to-mouth gestures
Overeating
Inertial sensors
Motif-based segmentation
K-Spectral Centroid Clustering
Fusion
Classification
Feeding gesture

ABSTRACT

Obesity, caused primarily by overeating, is a preventable chronic disease yielding staggering healthcare costs. To detect overeating passively, a machine learning framework was designed to detect and accurately count the number of feeding gestures during an eating episode to characterize each eating episode with a feeding gesture count using a 6-axis inertial wrist-worn sensor. Moreover, detecting feeding gestures is useful to aid in end-of-day dietary recalls. It has been shown that feeding gesture count correlates with caloric intake; the more one eats, the more calories one is likely consuming. Recent research has shown promise in passively detecting feeding gestures, but this effort focuses on bridging detection of feeding gesture count and identifying overeating episodes. This paper presents results on three experiments: highly structured (participants pretending to eat), in-lab structured with confounding activities (participants eating while performing other scripted activities), and unstructured overeating (participants induced to overeat while watching television and eating their favorite foods). Our experiment successfully induced overeating in 50% of the participants, showing a correlation between feeding gesture count and caloric intake in unstructured eating (r=.79, p-value=.007). Results provide an approximate upper bound on feeding gesture classification using exact segmentation techniques, and show improvement when compared to prior sliding window techniques. Results also suggest the importance of stressing the challenge of accurate segmentation over identifying the accurate classification technique in detection of feeding gestures. Since participant-dependent models provide optimal results, a motif-based time-point fusion classification (MTFC) framework is proposed using spectral energy density, K-Spectral Centroid Clustering, symbolic aggregate approximation (SAX), a Random Forest classifier (trained on segmented motifs) and a time-point classifier fusion technique to show reliable classification of feeding gestures (75% F-measure), and a 94% accuracy of feeding gesture count in the unstructured eating experiment, resulting in a root mean square error of 2.9 feeding gestures. Mapping feeding gesture count to caloric intake, we obtain a rough estimate of whether participants overate while watching television.

1. Introduction

Eating is essential to human life, but overeating relative to need is not. Unfortunately, once bad eating habits are formed, they become challenging to overcome. People overeat for many reasons, such as loss of control [1], impulsivity due to cues [2], or heightened emotional state as a result of stress [3] or negative affect [4], or even positive affect [5]. Being able to passively detect overeating in real time will enable researchers to understand the antecedents and causes of overeating.

Existing studies have already demonstrated the feasibility of detecting eating behavior [6]. One phenomenon in feeding gesture detection is the large variability in performance of detection feeding gestures among works. One possible explanation of the variance is that performance is not only dependent on algorithm design but also significantly dependent on the quality of the dataset, including but not limited to factors such as if the participant is conducting a standard gesture, how many feeding gestures are included in an eating episode as well as how many non-feeding gestures are included.

Our objective is to use wearable sensors not only to detect eating episodes and aid in dietary recall, but also to identify problematic eating episodes, where caloric overeating is likely to occur. We hypothesize that eating duration as well as the number of feeding gestures, swallows and chews are important physical

* Corresponding author.
E-mail addresses: shibozhang2015@u.northwestern.edu, zhangshibo87@gmail.com (S. Zhang).

http://dx.doi.org/10.1016/j.inffus.2017.08.003
1566-2535/

Table 1
List of terms and descriptions.

Term	Description
FG	Feeding gestures
nFG	Non-feeding gestures
# FG	The number of feeding gestures from ground truth
# FG prediction	The number of feeding gestures from prediction
# FG accuracy	The accuracy of number of feeding gestures from prediction

features that can characterize overeating. This effort focuses on a framework for characterizing challenging eating episodes through feeding gestures using wrist-worn inertial sensors. Challenges in this effort stem from two major aspects: (1) a single person can have feeding gestures with a variety of feature representations resulting from different utensils, miscellaneous table manners, improvisational feeding behaviors and unexpected sudden events like food and utensil dropping; (2) different personal behavior habits (stroking one's hair, scratching one's face) resulting in different eating rates, which makes it challenging to fit a generalized model to all users.

Due to the challenges described above, and since performance of feeding gesture classification depends on the test protocol and the behavior of each participant we explore the classification ability of machine learning algorithms to distinguish feeding gestures from feeding-like activities. We designed and conducted three different experiment protocols, including a highly structured test (participants pretending to eat), in-lab structured test with confounding activities (participants following an eating protocol), and an unstructured test (participants induced to overeat while watching T.V. and eating their favorite foods, after being full).

To detect and count feeding gestures, we designed a framework in two steps. First, we extracted motifs from exact segments (defined by ground truth) to build a database of motifs. Secondly, we perform K-Spectral Centroid Clustering to extract motif templates and perform motif matching to search for candidate feeding gesture segments. Thirdly, motif matching is performed using a symbolic aggregate approximation (SAX) method, followed by feature extraction, Random Forest classification, and a unique decision-level fusion method combining majority voting [7] and multiple overlapping window segments to detect and count feeding gestures.

The main contributions of this paper comprise:

1. a new time-point classifier fusion framework using motif-based templates with varying window lengths to count feeding gestures;
2. an upper bound of feeding/non-feeding gesture classification using exact segmentation;
3. a comparison between a motif-based segmentation technique and a previous sliding window technique;
4. promising results for predicting overeating behavior with only one wrist-worn motion sensor. Through detecting and counting the feeding gestures we can effectively predict overeating intervals; and
5. a well-labeled and video-recorded dataset containing a range of three experiment protocols.

Table 1 summarizes notations used in reporting results in this paper.

2. Related works

Sensing devices that use three-axis accelerometers and three-axis gyroscopes to define eating gestures and activities come in many forms and in many body locations including the wrist, neck, arm, and trunk [8–16]. Table 2 summarizes existing literature on detecting eating using feeding gesture counts along with the sensors used, segmentation methods and test data set specifications. Our results improve on existing literature by testing on more confounding gestures in lab and induction of T.V. watching-linked overeating episodes, while not compromising accuracy.

The widespread availability of embedded wearable accelerometers and gyroscopes has enabled a new area of research to detect eating passively through on-body inertial sensors. The focus of this research effort is on detecting and characterizing eating through feeding gestures.

The problem of identifying hand-to-mouth gestures has also been studied to detect smoking activities in order to predict smoking relapse [17–19]. PuffMarker successfully detected the timing of a relapse using multiple sensors, where detecting hand-to-mouth gestures and hand orientation (using roll and pitch angles) are part of the smoking lapse system [17]. RisQ also leveraged multiple inertial measurement units (IMUs) placed on a person's body together with 3D animation to detect hand-to-mouth gestures, and while the focus of these systems is primarily smoking, their systems have shown preliminary success in detecting feeding gestures.

Dong et al. showed correlation between bites and caloric intake and measures intake via automated tracking of wrist motion [20,21]. Some of their limitations include requiring the user to turn the device on and off, and they focus primarily on detecting the start and end of an eating episode throughout the day, as opposed to characterizing a given eating episode.

Thomaz et al. presented a framework for detecting eating episodes using wrist-worn accelerometer data [9] and density-based spatial clustering of applications with noise to identify eating episodes throughout the day. The effort focused more on detecting eating episodes that are at-least five minutes long, but our effort focuses on characterizing these eating episodes to obtain a more fine-grained understanding of eating behavior. We also compare our method with the sliding window technique to show improvements in feeding gesture classification.

Amft et al. [8,14] integrated commercial motion sensors in a jacket to capture the movement of the participants wrists, lower arms, upper arms and upper back. Then they built a feeding gesture detection system with the collected data. In comparison, the system that we used for our study is more accessible using a single wrist-worn sensor mounted on the participant's dominant hand.

Although many studies focus on eating gesture detection with various methods and systems, most of those studies are based on fixed-length sliding window segmentation and eating/non-eating gesture classification. Fixed-length sliding window segmentation benefits from simplicity, speed, and robustness. For example, in [9], the author presented presents a fixed-length sliding window machine learning framework to detect and count feeding gestures. In [10,12,13,15], the fixed-length sliding window segmentation framework was adopted. However, there are other segmentation methods. In [22], the author developed an interesting method to extract candidate windows. A fast (0.8 s window) and a slow (8 s window) sliding window were used for calculating the average gyroscope magnitude. Then segments where the fast-moving average lies below the slow-moving average were selected. In these regions, the magnitude is lower than the average magnitude of the neighborhood.

To distinguish our project from existing works, we present a flexible length sliding window approach. Our approach is based on capturing the complete hand up and hand down gesture, resulting in robust and intelligent segmentation and classification of feeding gestures.

Motif matching is the process of finding a specific motif within a signal. It is a wide topic of time series data mining which has been intensively studied in the academic community of bio-