Practical Food Journaling

Edison Thomaz
Georgia Institute of Technology
Atlanta, GA, USA
ethomaz@gatech.edu

Abstract
Logging dietary intake has been shown to be of benefit to individuals and health researchers, but a practical and objective system for food logging remains elusive despite decades of research. My thesis is that emerging wearable devices such as life-logging cameras, the ubiquity of sensors in mobile devices, and new computational techniques such as human computation, provide the foundation for a new class of food journaling systems that are lightweight and practical in everyday settings. In this proposal I describe my research in understanding how to leverage this new landscape of mainstream ubiquitous computing towards automatic and semi-automatic food journaling.

Author Keywords
Health; Diet; Food; Dietary Intake; Food Logging; Food Journaling; Food Journal; Automatic Dietary Assessment

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

UbiComp'13 Adjunct, September 8–12, 2013, Zurich, Switzerland.
ACM 978-1-4503-2215-7/13/09.

http://dx.doi.org/10.1145/2494091.2501089

Introduction

In 2008, one third of all adults in the U.S were overweight or obese, with other countries observing similar trends [7]. It is believed that an effective method for monitoring eating habits could help researchers expand their understanding of this seriously growing problem. At the individual level, keeping track of eating habits has been shown to contribute to positive behavior change by helping individuals become more aware of their dietary intake.

Researchers have explored the domain of automated dietary assessment for decades, but a practical system for food logging has not yet been realized. The fundamental challenge in food logging is that there is not an efficient way to collect dietary information that is objective, ecologically valid and does not pose a major burden on individuals. Today, mobile phone applications represent the state-of-the-art when it comes to food journaling; there are a myriad of applications that let users take photos and notes of their meals, some of which go a step further and even display the nutritional value of a meal through crowdsourcing techniques. The key challenge with these applications is that people need to remember to use them, which has proven to be particularly hard over a long period of time. Additionally, there is a time and effort cost associated with fetching a smartphone, unlocking it, launching an app and taking a photo or typing notes. It is inevitable that even the most engaged users might forget to log a snack or meal occasionally, or grow weary of dutiful logging over the long run. The truth is, these applications are simply not practical enough for sustained use.

The thesis underlying my work is that emerging wearable devices such as life-logging cameras, the ubiquity of sensors in mobile devices and activity trackers, and the combination of computational techniques such as human computation and machine learning, provide a new foundation from which to build practical, automatic and semi-automatic food journaling systems.

For my dissertation I plan to address the following research questions:

1. Can human computation be used to recognize eating moments in first-person point-of-view images taken with wearable cameras in everyday settings?
2. How can privacy concerns be addressed when recognizing eating moments from first-person point-of-view images using human computation?
3. Can multimodal sensor data from wearable devices and mobile phones identify eating moments?
4. Can habitual eating patterns be estimated from multimodal sensor data?

One of the cornerstones of my research agenda is the identification of when an eating activity takes place, since it is the centerpiece of a number of strategies for food journaling. Once a meal activity has been identified, several courses of actions might be pursued. An automatic trigger could be sent to a wearable camera to take a picture of the food [12, 10, 15], the individual could be nudged to add an entry to a food logging mobile application, or a text message could sent to the individual later in the day requesting more details about the meal.

Related Work

Manual food journaling is the current practical paradigm when it comes to food journaling. Today, a variety of food logging smartphone applications exist, many of which are very popular such as MealSnap and MyFitnessPal. Many of these applications facilitate the journaling task by requiring people to simply take a picture of their food [8, 12]. In the realm of mobile applications, other approaches have been tried such as offering alternative entry methods in food diaries and designing notification practices that remind people to log their meals.

Research in the area of automatic food tracking and recognition dates back to the 1980s when researchers tried to detect chews and swallows using oral sensors in order to measure the palatability and satiating value of foods [14]. Other sensor-based techniques involve detecting eating and drinking actions from acoustic and inertial sensors, and monitoring caloric intake using on-body or mobile phone-based sensors [3, 1].

A recently introduced approach to dietary monitoring involves using wearable cameras such as the eButton [2] and SenseCam [6] to document people's eating behaviors. A head or chest-mounted camera is configured to take first-person point-of-view photos automatically throughout the day (e.g. every 30 seconds), and the resulting snapshots capture people performing a wide range of everyday activities, from socializing with friends to having meals with family members. This technique is particularly promising because in addition to being completely passive, the images captured truthfully reflect people's eating activities and the surrounding context of those activities.

Current Research

One of the major challenges of identifying eating moments with photos automatically captured throughout the day is that only a small portion of images depicts an eating moment. The sheer volume of images generated per day makes it impractical to annotate them manually, and despite significant progress in the field of computer vision over the years, it remains impractical to automatically identify food items and human activities in images taken in real world settings. This is the first challenge I address in my dissertation work, and I do so by applying a new form of computation that has matured in the last five years: human computation. I devised a methodology for automatically recognizing eating moments from thousands of first-person point-of-view images by leveraging one of the most popular human computation services, Amazon Mechanical Turk (AMT). The method consists of collecting and filtering images for privacy protection, formatting the images into temporal groups, presenting them to a group of human computation workers by creating a human-intelligence task (HIT), and comparing their results to those obtained by a group of trusted coders who went through the same exercise.

I evaluated this methodology in a three-day 5-participant study and the system was able to recognize eating moments in real-world settings. Overall eating moment recognition accuracy reached 89.68% accuracy in the best case scenario, with overall precision at 86.11% and overall recall at 63.26%. Privacy arouse as an important element of this work, and privacy-related constraints dictated important aspects of the methodology. One of the challenges faced was that the wearable camera setup captured a large number of photos of non-study participants. Since these individuals were not in the study, we were forced to delete all such images. Importantly, the elimination of these photos had a detrimental impact on the performance of our system. This was the impetus for my follow-up work, the second research question I address in my dissertation: a framework for reasoning about and quantifying the results of privacy-protecting measures.

I developed a formulation, the privacy-saliency matrix, to guide the understanding of removing imagery that poses a threat to privacy while retaining imagery that is salient to the analysis of the activity (e.g eating behavior). To demonstrate the use of the framework, I quantified how four simple automated image processing techniques — face detection, image cropping, location filtering, and motion filtering — address the privacy challenge. This was achieved by conducting a study in which point-of-view imagery from a different set of 5 participants over an average of 3 days each was coded for the saliency of each image with respect to eating behaviors as well as the potential for privacy concerns.

As expected, none of the image processing techniques optimized the privacy and saliency of images to desired levels, but the study exposed the need for mechanisms that support reasoning about this optimization, which I believe the privacy-saliency framework does.

Proposed Research

Thanks to advances in sensing and mobile technologies over the last decade, sensors have been employed to automatically infer many aspects of human activity [9, 11]. When it comes to dietary assessment, researchers have experimented with a number of sensor modalities [1, 13, 16]. Unfortunately, despite promising results, none of the techniques explored so far have been practical enough for real-world usage.

One of the findings of the privacy-saliency matrix research effort was the value of sensor data in the context of identifying eating moments. The location and motion filtering techniques successfully leveraged sensor data to determine the likelihood that an eating activity was taking place. My proposed research hinges on this observation. Recently, a wide range of wearable devices such as the Fitbit, the Nike FuelBand, and the Garmin Forerunner have become popular in the consumer market. I plan to address research questions #3 and #4 by combining data provided by these mainstream wearables devices with smartphone sensor data to recognize eating moments and patterns in real world settings.

To recognize eating moment from sensor data, I plan to conduct a study in Fall 2013 where 20 participants will be asked to wear an inertial and an acoustic sensor, and install a sensor data logging application on their smartphones. Participants will also be asked to wear a wearable camera that will capture a photo of their activities throughout the day every 30 seconds. The study will last a single day and will start in the morning. At the conclusion of the study I will ask participants what times they had meals that day and confirm the time of the eating activities with the first-person point-of-view images from the wearable cameras. With the knowledge of when eating moments occurred, I will train a classifier using machine learning techniques and evaluate it using cross-fold validation.

Routine characterizes human life, and these routines manifest themselves in our everyday interaction with technology [4]. The fourth research question I plan to answer in my proposed work is whether eating patterns can be recognized using opportunistic sensing and machine learning techniques. Researchers interested in discovering people's life patterns have relied a number of methods for finding discontinuous and varied-order activity patterns in an individuals behavioral data [5]. One of the challenges of these unsupervised approaches is the amount of data required. Another consideration is that once patterns have been detected, it is critical to learn what activities the patterns refer to. Interactive machine learning techniques, where end-users provide labels or features to guide the process of learning, can be used towards this end.

To address the question of whether eating patterns can be recognized, I will conduct a study with 15 participants over an entire month in Spring 2014. Participants will be provided with a sensor setup similar to the one used to answer question #3: inertial and acoustic sensors, a mobile phone application and a wearable camera. The sensor data will be automatically collected throughout the study by means of a sensor aggregation platform. At the end of the study, I will coalesce the multi-day sensor data streams for each participant and cluster them using Gaussian Mixture Models (GMM) using the EM algorithm. To evaluate whether the clusters represent actual routines in people's everyday activities, I will interview participants and ask them about their habits, attributing special emphasis to eating patterns. In a real-world scenario, where interviews are impractical, cluster labels might be obtained through SMS messaging, where users of the system might be occasionally prompted for input to guide the learning of eating pattern models.

I feel strongly that the availability of models that can predict eating moments and patterns from multimodal, opportunistic sensor data will serve as the foundation for a new class of food journaling systems that are lightweight, practical and usable in everyday settings. This is especially the case because the devices from which the sensor data will originate will be consumer products that individuals have already incorporated into their lives, such as smartphones and activity trackers. This is in contrast to custom sensing approaches for dietary assessment that have been used in previous research.

Biographical Sketch

In August 2013, I will start my fourth year as a Ph.D. student at the Georgia Institute of Technology, in the Human-Centered Computing program. My advisors are Dr. Gregory D. Abowd, Regents' and Distinguished Professor in the College of Computing, and Dr. Irfan Essa, Professor in the College of Computing. We are affiliated with the School of Interactive Computing and the GVU Center. I have a S.M. in Media Arts and Sciences from the MIT Media Lab, awarded in 2002, and a B.A. in Computer Science from The University of Texas at Austin, awarded in 1999.

Acknowledgements

I would like to thank the Intel Science and Technology Center for Pervasive Computing (ISTC-PC) for supporting this work.

References

[1] Amft, O., and Tröster, G. On-Body Sensing Solutions for Automatic Dietary Monitoring. *IEEE pervasive computing 8*, 2 (Apr. 2009).

[2] Bai, Y., Li, C., Yue, Y., Jia, W., Li, J., Mao, Z.-H., and Sun, M. Designing a wearable computer for lifestyle evaluation. In *Bioengineering Conference (NEBEC), 2012 38th Annual Northeast* (2012), 93–94.

[3] Chen, S., Lach, J., Amft, O., Altini, M., and Penders, J. Unsupervised Activity Clustering to

Estimate Energy Expenditure with a Single Body Sensor. *marcoaltini.com*.

[4] Consolvo, S., McDonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., and Landay, J. A. Activity sensing in the wild: a field trial of ubifit garden. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, ACM Request Permissions (Apr. 2008).

[5] Eagle, N., and Pentland, A. S. Eigenbehaviors: identifying structure in routine. *Behavioral Ecology and Sociobiology 63*, 7 (2009), 1057–1066.

[6] Gemmell, J., Williams, L., Wood, K., Lueder, R., and Bell, G. Passive capture and ensuing issues for a personal lifetime store. *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences* (2004), 48–55.

[7] Kimokoti, R. W. R., and Millen, B. E. B. Diet, the global obesity epidemic, and prevention. *YJADA 111*, 8 (Aug. 2011), 1137–1140.

[8] Kong, F., and Tan, J. DietCam: Automatic dietary assessment with mobile camera phones. *Pervasive and Mobile Computing 8*, 1 (Feb. 2012), 147–163.

[9] Lane, N., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., and Campbell, A. A survey of mobile phone sensing. *Communications Magazine, IEEE 48*, 9 (2010), 140–150.

[10] Liu, J., Johns, E., Atallah, L., Pettitt, C., Lo, B., Frost, G., Wearable, G.-Z. Y., and Implantable Body Sensor Networks BSN, . N. I. C. o. An Intelligent Food-Intake Monitoring System Using Wearable Sensors.

[11] Lu, H., Yang, J., Liu, Z., Lane, N., Choudhury, T., and Campbell, A. The Jigsaw continuous sensing engine for mobile phone applications. *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems* (2010), 71–84.

[12] Martin, C. K., Han, H., Coulon, S. M., Allen, H. R., Champagne, C. M., and Anton, S. D. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. *British Journal of Nutrition 101*, 03 (July 2008), 446.

[13] Passler, S., and Fischer, W. Acoustical method for objective food intake monitoring using a wearable sensor system. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on* (2011), 266–269.

[14] Stellar, E., and Shrager, E. E. Chews and swallows and the microstructure of eating. *The American journal of clinical nutrition 42*, 5 (1985), 973–982.

[15] Sun, M., Fernstrom, J. D., Jia, W., Hackworth, S. A., Yao, N., Li, Y., Li, C., Fernstrom, M. H., and Sclabassi, R. J. A wearable electronic system for objective dietary assessment. *Journal of the American Dietetic Association 110*, 1 (2010), 45.

[16] Yatani, K., and Truong, K. N. BodyScope: a wearable acoustic sensor for activity recognition. 341–350.