

# An Exploratory Study on a Chest-Worn Computer for Evaluation of Diet, Physical Activity and Lifestyle

Mingui Sun, Ph.D.<sup>1,6,8\*</sup>, Lora E. Burke, Ph.D., R.N.<sup>2</sup>, Thomas Baranowski, Ph.D.<sup>3</sup>,  
John D. Fernstrom, Ph.D.<sup>4</sup>, Hong Zhang, Ph.D.<sup>5</sup>, Hsin-Chen Chen, Ph.D.<sup>1</sup>,  
Yicheng Bai<sup>6</sup>, Yuecheng Li<sup>1</sup>, Chengliu Li<sup>6</sup>, Yaofeng Yue<sup>6</sup>, Zhen Li<sup>1</sup>,  
Jie Nie<sup>1</sup>, Robert J. Sclabassi, M.D., Ph.D.<sup>7</sup>, Zhi-Hong Mao, Ph.D.<sup>6,8</sup> and  
Wenyan Jia, Ph.D.<sup>1</sup>

*Departments of <sup>1</sup>Neurosurgery, <sup>2</sup>Health and Community Systems, <sup>4</sup>Psychiatry,  
<sup>6</sup>Electrical and Computer Engineering, <sup>8</sup>Bioengineering, University of Pittsburgh,  
Pittsburgh, PA, USA*

*<sup>3</sup>Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA*

*<sup>5</sup>Center for Image Processing, Beihang University, Beijing, China*

*<sup>7</sup>Computational Diagnostics, Inc., Pittsburgh, PA, USA*

Submitted January 2014. Accepted for publication December 2014.

## ABSTRACT

Recently, wearable computers have become new members in the family of mobile electronic devices, adding new functions to those provided by smartphones and tablets. As “always-on” miniature computers in the personal space, they will play increasing roles in the field of healthcare. In this work, we present our development of eButton, a wearable computer designed as a personalized, attractive, and convenient chest pin in a circular shape. It contains a powerful microprocessor, numerous electronic sensors, and wireless communication links. We describe its design concepts, electronic hardware, data processing algorithms, and its applications to the evaluation of diet, physical activity and lifestyle in the study of obesity and other chronic diseases.

**Keywords:** Wearable computer, mobile technology, diet, physical activity, lifestyle, health

## 1. INTRODUCTION

It has been well recognized that lifestyle and environmental factors, including diet, physical activity, living environment, social interaction, etc., play an essential role in human health. It has been estimated, for the chronic diseases as a whole, between 70% and 90% of the disease risks are associated with lifestyle and environmental factors [1-3]. It has also been reported that environmental pollutions, indoor and outdoor types combined, account only for 7%-10% of all disease risks [4, 5]. In contrast, some lifestyle related factors, such as diet, weight control, and tobacco smoking, comprise risks of stroke, colon cancer, coronary heart disease and Type 2 diabetes by 70%-90% [4, 5].

---

\*Corresponding author: Mingui Sun, Department of Neurosurgery, 3520 Forbes Ave, Suite 202, Pittsburgh, PA 15260, USA. Phone (412) 802-6481. Email: drsun@pitt.edu. Other authors: lbu100@pitt.edu; Thomas.Baranowski@bcm.edu; fernstro@pitt.edu; dmrzhang@gmail.com; hsinlovehsin@gmail.com; baiyicheng@gmail.com; ycleo202@gmail.com; yueyf2006@gmail.com; lizhen0130@gmail.com; niejie117@126.com; bobs@cdi.com; zhm4@pitt.edu; jiawenyan@gmail.com.

Despite the high importance of lifestyle in disease risks, in recent years, unhealthy lifestyle has been adopted by an increasing portion of the population. For example, more than 60% of adults in the United States are overweight, and over one-third are obese [6]. Obesity causes approximately 300,000 deaths each year in the U.S. [7]. A recent study indicates that the estimated direct and indirect costs of obesity to the U.S. economy are at least \$215 billion annually [8]. Over the years, unhealthy foods with an increasingly large portion size have been among the best selling products of fast food chains; the abundance of “effort-saving” facilities, the increased use of smartphones, and extended TV watching have reduced physical activity of many people, notably children and adolescence, to the minimum. These problems have led to a steady rise of chronic diseases, such as cardiovascular disease, cancer, lung disease, and diabetes.

Although the link between lifestyle and health has been well established, there is no objective method to evaluate an individual’s lifestyle except direct observation. Currently, self-reporting has been the primary tool for lifestyle and environmental exposure assessment which depends on the memory and willingness of individuals to report their personal data and experience. This subjective method is not only unreliable and inaccurate, but also extremely burdensome and time-consuming for both the subject being assessed and the assessor. As pointed by Rappaport [5], “we have only a sketchy understanding of the culprits because information about exposures is gleaned almost exclusively from indirect and uncertain questionnaire data.”

In recent years, there have been rapid advances in information technology, microelectronics, and mobile systems. These technological advances provide new hope to evaluate people’s diet, physical activity and lifestyle, providing accurate information for healthcare professionals to reduce the risk of chronic diseases and keep people healthy. There has been considerable research in this field. Over the years, webpages or mobile phone apps have become widely available to evaluate diet. Although this type of evaluation is facilitated by using convenient human-computer interface, the method is still self-report subject to bias and memory errors as stated previously. Recently, smartphones have been used for dietary evaluation. An interesting app called Meal Snap provides an estimate of calories based on a single food picture uploaded by the app user [9]. However, the estimate is not done computationally, but provided by a random group of people on the web (called Mechanical Turk) who give estimates for a small amount of payment. More technologically, the smartphone has been used to take pictures of food before and after it is consumed [10-12]. These pictures are either saved for off-line processing or transmitted wirelessly to a data processing center where food volumes and nutrients/calories are estimated using image processing algorithms and an established nutrient/calorie database. Although the smartphone can produce high-quality food pictures, this method depends on human attention during an eating event, which is inconvenient and may modify the person’s behavior.

For assessment of physical activity, accelerometers have been utilized by wearing it at different locations to monitor body motion [13]. In this domain, there are numerous commercial products in forms of armbands, wristbands, shoe attachments and waist belt attachments. Examples of these include the Actigraph (a waist belt attachment by Actigraph, Inc., Pensacola, FL, USA) and the Fitbit (a wristband featured with a slim

design by Fitbit Inc., San Francisco, CA, USA). Although these devices provide objective measurements, they all have difficulties in distinguishing certain specific physical activities, such as normal walking or a treadmill walking. Therefore, these wearable devices cannot provide sufficient information to evaluate an individual's lifestyle.

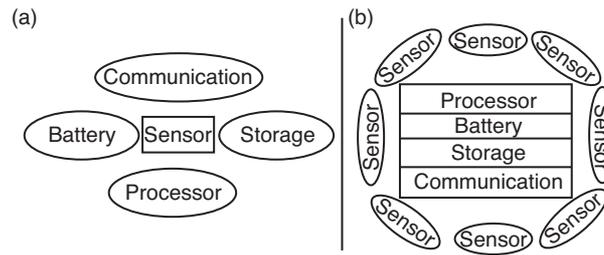
With a research grant from the National Institutes of Health (NIH) in the U.S., we have developed a wearable computer, eButton, for the study of personal health and wellbeing. The major goal of this research is to provide the research community with a convenient, unobtrusive and multifunctional tool for data collection in the real world. This tool is in the form of a miniature wearable computer. When this computer, in the shape of a disc, is pinned to the clothing in front of the chest or other locations, it operates automatically in the personal space to acquire objective data for the evaluation of diet, physical activity and lifestyle. The design and implementation of this device span several fields, including structural and electronic hardware design, computational algorithms for data management and analysis, and implementation strategies in different application domains. Although we have previously published a number of papers focusing on specific technical issues with detailed experimental data, the rationale, concepts and applications of the wearable computer and its relevance to personal health and wellness have not been discussed in depth. In this paper, we will present these rationale, concepts and applications. Experimental data or real-world examples are provided after each methodological presentation. Due to the length limit of this paper, we emphasize on design concepts without great details on specific technical issues. However, we provide references containing detailed descriptions. In the final sections, we discuss the current limitations of the wearable computing approach and then conclude this work.

## **2. DESIGN CONCEPTS AND HARDWARE DESIGN**

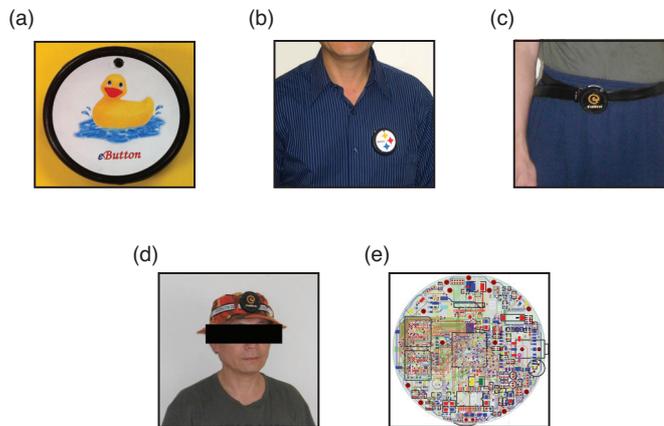
Our primary concept is to design a wearable computer with integration of a powerful data processor, massive data storage, wireless communication links, and a large array of sensors. Our secondary concept is to design a new form of computer which is convenient, unobtrusive, passive, attractive and decorative in the personal space so it can help its owner anywhere, anytime, naturally. Our immediate goal of this design is to construct a multi-purpose, unified electronic system to acquire various forms of data for the evaluation of diet, physical activity and lifestyle.

### **2.1. Architecture**

Traditional electronic sensor systems are mostly focused on acquiring one or several physical variables, such as an accelerometer for motion measurement and a wearable camera for taking pictures. The architecture of these systems is depicted in Figure 1(a) where the sensor is the central component of the system. It is served by other necessary system components such as a power supply, a storage, a communication channel, and a processor (e.g., a microcontroller running a pre-stored program in its memory). In contrast, the architecture of a typical computer, shown in Figure 1b, features a central processing unit (CPU) running an operating system (e.g., Linux or Android), a power



**Figure 1.** (a) Architecture of a typical sensor system where the available resources are dedicated to the sensor. (b) Architecture of a typical computer where an array of sensors share the same system resources.



**Figure 2.** (a) eButton designed as a pleasant wearable object with a replaceable face image of personal preference; (b) - (d) Examples of wearing options; (e) Mother board design inside eButton.

supply (e.g., a rechargeable battery), several types of memories, and a communication system. These central resources are shared by peripherals which, in our case, are clusters of sensors.

## 2.2. Electronic Hardware

The eButton adopts the architecture of Figure 1b. Therefore, it is a wearable computer. This device has about the same size as a chest pin (60 mm diameter), and weighs approximately one quarter of a smart phone. The face of eButton is covered by a removable sticker which can be personally designed (Figure 2a). Despite its simple and personalized appearance (Figure 2b), eButton contains a powerful CPU running the Linux or Android operating system with a 32-bit ARM Cortex A9 architecture and four processing cores. The CPU can be clocked at a maximum of 1.4 GHz and is associated with a 2GB RAM, an 8GB NAND flash memory, and a set of wireless communication

ports supporting both Bluetooth and Wi-Fi. These system resources are shared by many sensors, including two video cameras, a light sensor, a 3-in-1 inertial measurement unit (IMU) IC (one 3-axis accelerometer, one 3-axis gyroscope and one magnetometer), an audio processor, a proximity sensor, a barometer, and a global positioning system (GPS) receiver. The hardware and software resources make eButton a smart computer, yet it is small enough to be worn conveniently and naturally (Figure 2b through 2d). Unlike the smartphone which is utilized occasionally during the day, the eButton is able to work continuously for 4-8 hours (longer with a larger or an extension battery), performing a variety of personal computing tasks in real-life settings.

### **3. DATA PROCESSING**

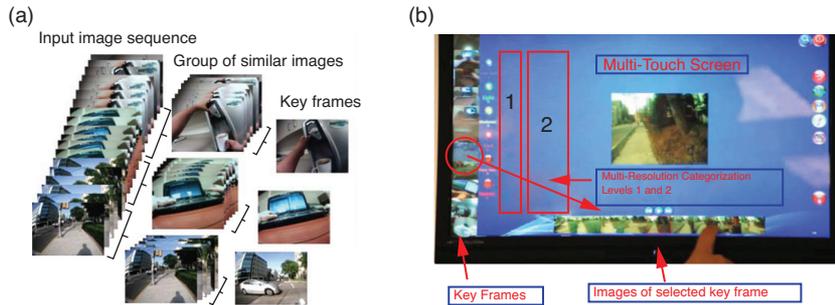
The group of sensors in the eButton produces multiple streams of data at high rates. Therefore, this device faces a so-called “big data” problem. It is impractical to process these data entirely by manual examination. Automatic processing is highly desirable. Unfortunately, the state-of-the-art of current algorithms cannot support fully automatic processing, such as recognizing a large variety of human foods and activities. As a temporary solution, we use computer to condense data produced by eButton before examination by humans. These data can be categorized into two types: time series in multiple channels and image sequences.

#### **3.1. Multichannel Time Series**

In many cases, the time series data can be processed automatically according to their physical properties. For example, the sequence of latitude and longitude coordinates produced by the GPS sensor can be processed by incorporating them into the Google map forming continuous paths, while the atmosphere pressure data produced by the barometer can be converted into height values relevant to a selected reference (e.g., the ground level or the first floor of a building). Detailed discussion can be found in [14].

#### **3.2. Image Sequence**

In order to save storage space, the eButton is often programmed to take a picture every 1 to 5 seconds. We call the resulting data an “image sequence” which may be considered to be the third type of pictorial data, different from the commonly known types of “image” and “video”. Despite the extremely low picture-taking rate in the image sequence, eButton still produces thousands of images per hour. In order to reduce data processing burden, we utilize an automatic data pre-processing scheme as shown in Figure 3a. Each image in the input sequence is compared with its neighbors to measure similarities with respect to intensity, color, texture, edge profile and other features [15]. Features from other data, such as body motion and GPS data, may also be utilized in this segmentation process [16]. Then, similar images are grouped together called “shots”. Finally, within each shot, a representative image, which is either the centrally located image or the one having the minimum “distance” to all images in the same group, is selected as the “key frames” of the group for next-step computer processing or human examination. Although this algorithm usually produces meaningful results, incorrect event segmentation occasionally takes place. We



**Figure 3.** (a) Automatic segmentation of image sequence (left) into groups of similar images (middle) and key frames (right). For privacy protection, we used image processing algorithms to recognize human faces and block them (orange circles) before images are observed. (b) Interactive software tool with red arrows and blocks indicating different components. Using this tool, the key frames produced by computer can be examined and the events observed can be categorized into two levels of resolutions.

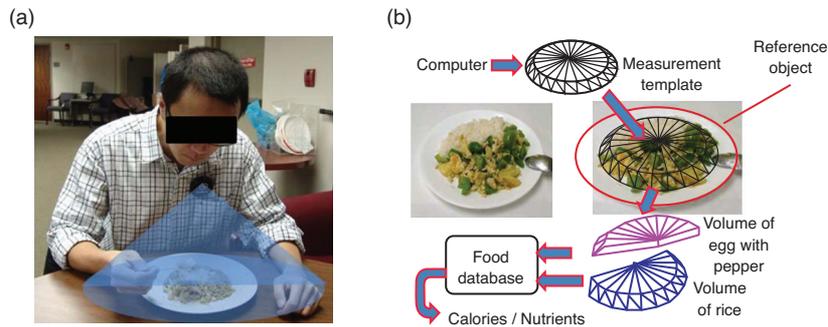
developed an interactive software tool to make necessary adjustments. In order to facilitate human-computer interaction, a large multitouch screen is utilized allowing convenient operation by sliding fingers (Figure 3b).

#### 4. APPLICATIONS OF EBUTTON AND EXPERIMENTAL DATA

As a wearable computer with opening eyes, the eButton can serve as a personal aid for health and wellbeing with functionalities beyond those currently provided by the smartphone. Its healthcare related applications, including the assessments of diet, physical activity and lifestyle which are essential to health maintenance and wellbeing, are described below. All research studies involving human subjects were reviewed and approved by the Institutional Review Board (IRB) before experiments were conducted.

##### 4.1. Dietary Evaluation

In contrast to physical activity evaluation for which a variety of commercial wearable devices are available, there is currently no wearable device on the market that can monitor human diet automatically and objectively in daily life. The available dietary evaluation tools, such as mobile phone apps and web-based software tools are based on self-reporting which, as stated previously, are time-consuming, subjective and inaccurate. The eButton has a potential to become the first electronic “eating sensor” that enables objective dietary evaluation in the real-world setting. The mechanism of this evaluation is quite simple. When the eButton is worn as a chest pin (Figure 2b and Figure 4a), a person’s meal is constantly and passively recorded by the cameras on the eButton. The field of view, between  $105^\circ$  and  $120^\circ$ , is indicated by the blue conic field in Figure 4a. After the data are downloaded or wirelessly transmitted (depending on



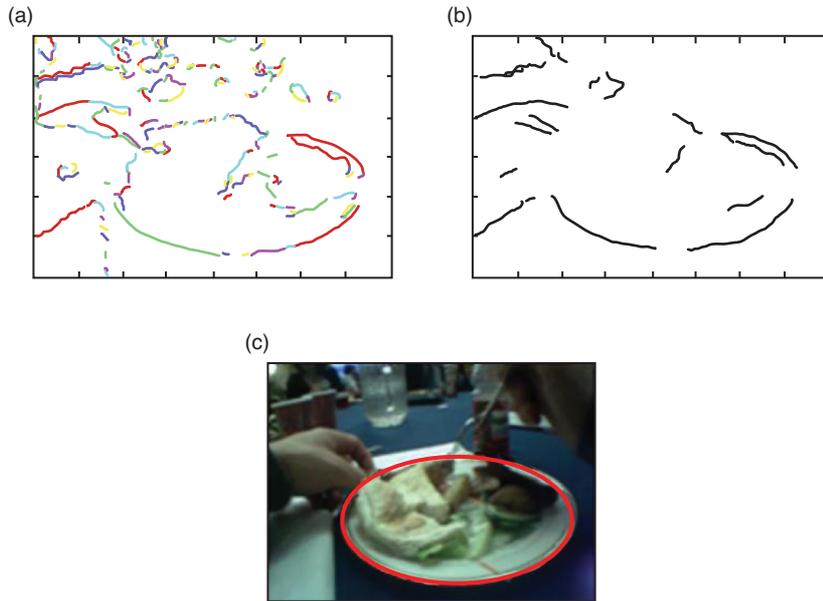
**Figure 4.** (a) During a meal, the camera(s) on the eButton take(s) pictures of the food continuously and passively [17] (used with permission). (b) The pictures are processed using a wire mesh to measure volume(s), and a food database to determine calorie and nutrient values.

whether a real-time evaluation is necessary) to a computer, we perform the following computational procedure to estimate the calories and nutrients in the food as shown in Figure 4b: 1) eating events are detected from thousands of images by automatically detecting the presence of dining utensils, such as plates (marked by red ellipse); 2) food within plates are segmented automatically (e.g., rice and peppers with eggs in Figure 4b); 3) the volume of food is measured using a computer generated 3-D wire mesh of an appropriate shape (here the shape is a truncated conic mesh); 4) different food items, if any, are separated (the purple and blue portions of the mesh); 5) the food name and portion size are submitted to a food database (e.g., the public-domain FNDDS database developed by the U.S. Department of Agriculture [18]) from which the calories and nutrients are determined. Since food recognition in the last step is performed manually (as stated previously, automatic recognition from images is currently an unsolved problem) and the food database is already established, we will describe the first four steps and demonstrate the effectiveness of our computational method in each step.

We must point out that not all eating scenarios can be monitored using the five computational steps. “Finger foods” (foods eaten without using utensils, such as candies, snacks and some fast foods) are difficult to be segmented directly from images automatically. While this problem is still being investigated, human help is currently needed in these cases during the data processing stage.

#### 4.1.1. Plate Detection

In terms of image processing, finding plates is equivalent to detecting ellipses (which could be partial ellipses due to occlusion or illumination effects). We have developed the following algorithm for ellipse detection: 1) detect edges in the image using the well-known Canny detector [19]; 2) connect edge pixels to form continuous curves using a line drawing algorithm [19] (Figure 5a); 3) each curve  $y = f(x)$  is split into arcs when: a) there is a non-differentiable hard corner ( $f'(x_c)$  does not exist), b) the tangent

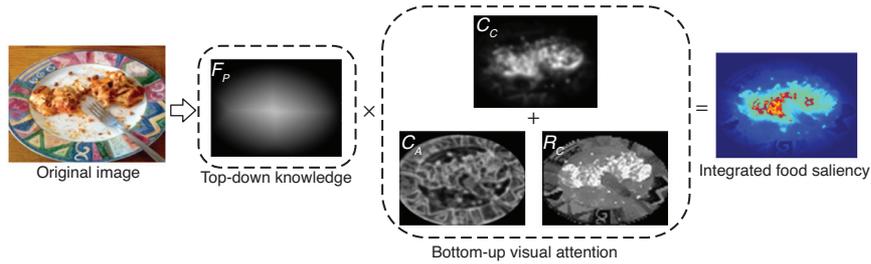


**Figure 5.** (a) Edge image after connecting edge pixels and splitting into arcs where colors are used to identify splitting results. (b) Edge image after removing small arcs. (c) Detected plate in the input image.

line is horizontal ( $f'(x_c) = 0$ ), or c) there is a reflection point ( $f''(x_c) = 0$ ); 4) remove small arcs according to an arc length threshold derived from statistics of plate and background edges (Figure 5b); 5) arcs are grouped according to the property that a convex object is always on the same side of the tangent line through any of its boundary; 6) an upper triangular conjunction matrix is formed which describes arc connectivity and convexity; 7) ellipses are identified by fitting convex groups of arcs with the goodness of fitting above a threshold; 8) plates are determined from ellipses which satisfy a set of size and orientation criteria (Figure 5c). The details of this algorithm are reported in [20-22].

#### 4.1.2. Food Detection and Segmentation

Using a computer to detect edible objects in images is a very difficult problem. It requires large amounts of knowledge about human foods and their appearances. However, when dining plates are detected, food detection is greatly simplified although the problem is still non-trivial. We approach this problem by attacking three common problems that affect food identification within a plate: 1) complex and highly variant shapes, colors and sizes of food components; 2) decorative patterns on some plates; and 3) presence of non-food objects, such as a fork, chopsticks, or a spoon. We have developed an effective algorithm for food segmentation by computing a salient map consisting of four masks,  $F_p$ ,  $C_A$ ,  $C_c$  and  $R_c$  [23, 24]. Mask  $F_p$  is simply a Gaussian



**Figure 6.** Process of salient map construction: The center of the plate in the input image (the left-most panel) is used to define a Gaussian mask  $F_p$ . This mask is multiplied by the sum of three masks  $C_c$ ,  $C_A$  and  $R_c$  representing, respectively, the color contrast, color abundance, and spatial arrangement. The resulting salient map is shown in the right-most panel which represents the area where food is likely located.

function centered at the plate center, representing the likelihood of food location within the plate. Mask  $C_A$  is computed by evaluating the abundance of different colors within the plate. Mask  $C_c$  represents color contrasts which characterize the attractiveness of a region against its surroundings. Finally, mask  $R_c$  measures spatial arrangement of food in a single or a number of compact regions within plates. Each mask is normalized with values between 0 and 1. Then, a saliency map  $F_s$  is constructed by  $F_s = \frac{1}{3}F_p(C_c + C_A + R_c)$ .

Figure 6 exemplifies the process of computing the salient map  $F_s$ . Once  $F_s$  is obtained, we utilized a registration-assisted deformation algorithm to define the food region according to the boundary of  $F_s$  [25]. An iterative approach is implemented using an ellipse as the initial primitive contour. An energy function is then minimized to obtain the final boundary in a way similar to that utilized in the segmentation algorithm called “snakes” [26]. Figure 7 shows several examples of food segmentation. It is demonstrated that our algorithm performs well even with difficulties.

#### 4.1.3. Volume Measurement

It would be ideal to reconstruct the food in 3D based on a set of 2D images so that the volume can be measured based on a complete 3D representation. Unfortunately, this is impractical in our case. Although a camera can take multiple pictures, the variation of the camera’s view is small, as seen in Figure 4a. Effectively, there is only one (if a single camera is used) or two (if two cameras are used) effective views of the food, which is (are) not enough for a 3D reconstruction. To solve this problem, based on the fact that most foods have known shapes, e.g., an apple is spherical and a piece of pizza is triangular, we developed a virtual reality (VR) approach to measure food volume [17, 25].

We utilized a mathematical model (Figure 8, upper panel) to relate three coordinate systems, a real-world coordinate system where the 3D food objects are present (labeled as “real object”), a virtual coordinate system in which a 3D mesh for volume

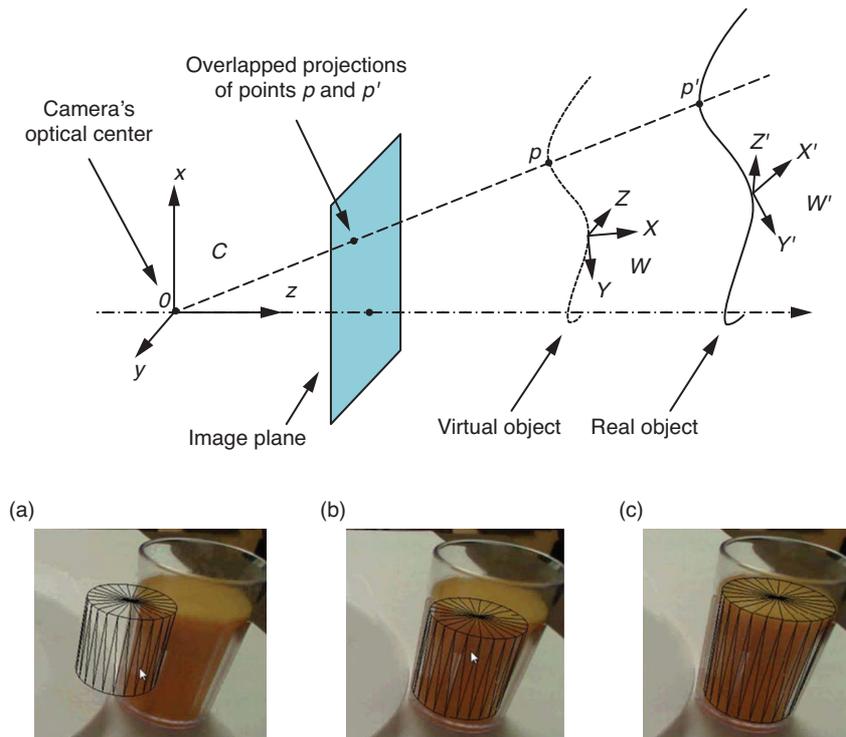


**Figure 7.** Results of automatic food segmentation using the salient map approach. The red arrows indicate previously stated difficulties (presence of chopsticks and decorative plate patterns in the left and middle panels, respectively).

measurement is defined (“virtual object”), and an image coordinate system for the sensor plane of the camera (“image plane”). Our idea is to superimpose the real world scene of the food on the image plane with the projection of a 3D wire mesh of known shape and volume. Since the 3D mesh can be rotated, deformed, and scaled, food volume can be estimated from the volume encapsulated by the mesh when the optimal fit between the food and mesh is achieved (Figure 8, bottom panels).

The process of determining coordinate correspondence is called calibration which is usually performed based on a pinhole model [27-29]. Several matrices and vectors are determined to provide coordinate transformations. If a single camera is utilized, a referent of known shape and size is required to be present in the image. Particularly, if a circular dining plate is used as the referent, the diameter of the plate must be known. In the two-camera case, a referent is not required because the known distance between the two cameras can be used to gauge the food size.

Once the correspondences between the real and virtual world coordinates are established, the next task is to select and manipulate the wire mesh. We first established a mesh prototype library, including a whole spheroid, a portion of a spheroid, a rectangular box, a cylinder, a triangular prism, a conical frustum, and an irregular shape that the user can self-define. We pre-associate each common food with a mesh prototype (e.g., an orange is associated with a sphere, and a piece of pizza is associated with a prism). Therefore, as long as a food name is given (either automatically or manually), a wire mesh prototype is ready to use. During our computational matching procedure, the prototype mesh is first placed on the base of the plate centered within the food contour. Then, the mesh is scaled and deformed iteratively so that the projected 2D outer boundary of the mesh best matched the segmented food boundary described



**Figure 8.** Procedure of food volume measurement taking the virtual reality approach. Upper panel: Three coordinate systems including the 3D real-world coordinates, 3D virtual coordinates and 2D image plane coordinates. The three coordinates are related by a calibration process. Lower panel: An example of measuring orange juice volume: (a) a cylindrical template is selected; (b) the template is placed within the target; (c) the template is adjusted to fit the volume of orange juice.

previously. This automatic procedure generally produces acceptable results for food with simple shapes. However, in cases where a food has a complex shape, is unfavorably oriented, or is closely connected with other foods in the image, human adjustment becomes necessary. We have developed a software platform to perform this adjustment manually on a computer screen. Figure 9 shows several examples of matching result using our virtual reality method.

#### 4.1.4. Experimental Result of Food Volume Measurement Accuracy

As discussed earlier, the eButton provides two pieces of information allowing dietary evaluation in real-world settings. The first piece is a pictorial record from which foods can be identified or recalled. The second piece is food volume (or portion size)



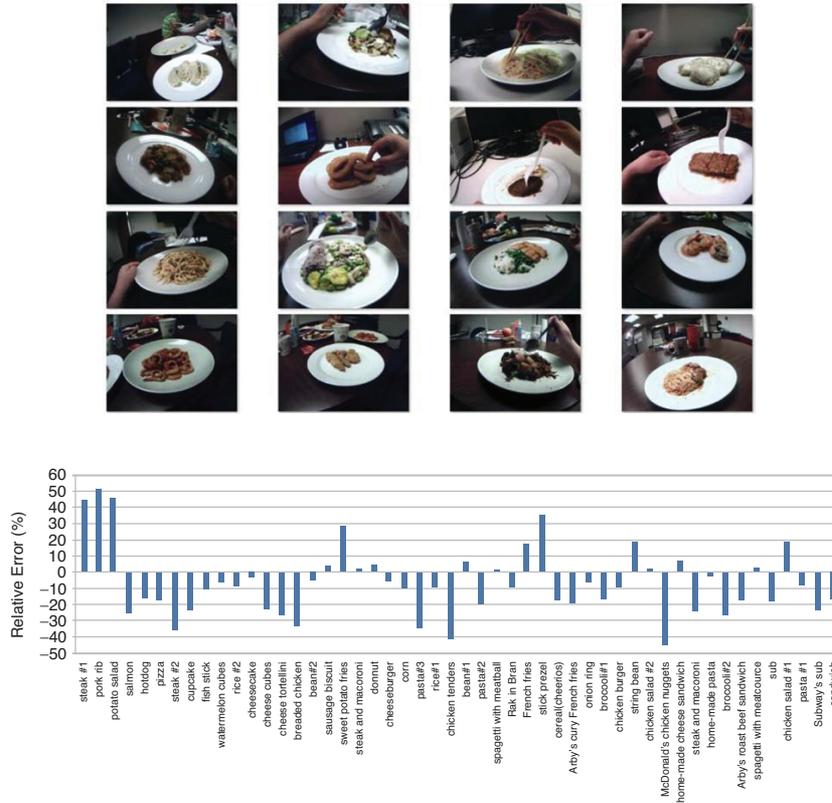
**Figure 9.** Examples of final stages of virtual reality portion size measurement.

measured from images. Calories and nutrients can then be determined from a food database based on the two pieces of information. Since food recognition is currently performed manually and the database is already established, the performance of eButton is mainly determined by the accuracy in food volume measurement. In order to evaluate this accuracy, we studied 100 real-life foods, including 50 western foods and 50 Asian foods. Seven human subjects participated in this study who ate lunch normally wearing an eButton. Part of the foods recorded is shown in the upper row of Figure 10. Before each study, the food brought in by the participant from his/her home or purchased from a local restaurant was physically measured by wrapping it in a plastic film and then submerging it in a pool of millet seeds. The increase in volume before and after submerging was used as the true food volume. Our results for Western foods are shown in the bottom panel of Figure 10 (those for Asian foods are similar). While 85 out of the 100 foods have errors within 30%, some errors are quite large. The foods with large errors mostly had irregular shapes or small sizes, or were recorded in images of poor quality. The mean absolute relative error (MARE) for all foods was 16.4% and the root mean square error (RMSE) was 20.5%, where the MARE and RMSE were defined as the mean of the absolute volumetric error divided by the true volume, and the square root of the mean square of the relative error, respectively, both expressed in percentage values.

In order to compare the objective method using eButton and the subjective method for portion size estimation, we presented the same 100 food images to three raters and asked them to estimate portion sizes visually. When compared with the ground truth, the MAREs for the three raters were 74.3%, 34.2%, and 51.5%, respectively, and the RMSEs were 100.5%, 42.2 %, and 58.1%, respectively. Therefore, the eButton method has a much better agreement with the ground truth than the traditional subjective estimation method.

#### **4.2. Physical Activity Evaluation**

Traditionally, physical activity was evaluated in a way similar to diet, i.e., self-report the activity type, such as aerobic exercise, strength exercise, and sports participation [30, 31]. In recent years, as miniaturized accelerometers based on microelectromechanical systems (MEMS) technology become popular, various wearable devices have been developed and commercialized for physical activity evaluation. The most widely used wearable device is perhaps the pedometer (e.g., Omron pedometers) which counts the number of walking steps. More sophisticated



**Figure 10.** Part of the eButton recorded food images, including both Western and Asian foods (top photos). Relative errors of volumetric measurements with the physical volumetric measurements using millet seeds as the ground truth (bottom chart) [17] (used with permission).

wearable devices integrate a number of functions into a single device, such as the Fitbit Flex which counts the number of steps, walking distance, calories burned and active minutes. The latest devices feature more functions, such as sleep monitoring, perspiration, and heart rate, as well as wirelessly transmitting data to a computer or a smartphone using the Bluetooth technology (e.g., the Fitbit Flex Wireless, Jawbone UP, Motorola MotoActv, Basis B1, etc.). Many devices are designed in convenient user-friendly forms such as wristband or belt clip. Therefore, if the goal is to monitor body motion and estimate calorie expenditure, the current commercial devices provide ample choices. However, in many cases, it is desirable to know specifically what type of activity, especially sedentary activity, is performed at a given time, such as distinguishing between TV watching and web surfing. Such function is important since,

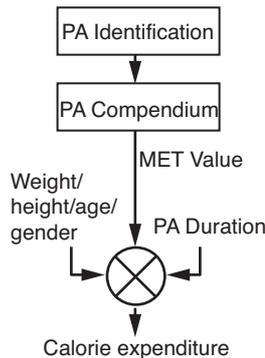
for most people, sedentary events dominate their daily lives [32], and eButton provides a powerful tool to this end. As an intelligent computer equipped with more sensors, such as cameras and a GPS sensor, it can bring physical activity evaluation to a higher level. Three methods for different applications are presented below.

***Method 1: Accelerator-Based Method***

The accelerometer data have been processed to obtain “counts” which measure the strength of an activity [33]. Our algorithm includes the following steps: 1) applying a band-pass filter to suppress noise; 2) within each of a set of non-overlapped windows (e.g., one-minute windows), calculating the number of times (counts) the amplitude of the normalized waveform is above a pre-set threshold. Counts have been related to the Metabolic Equivalent (MET) which is used to compute energy expenditure of a physical activity in population studies [34, 35].

***Method 2: Camera-Based Method***

We have utilized the following camera-based method to document people’s daily activities, summarize them, and compute their energy expenditure values. This method, illustrated in Figure 11, begins with visually observing the images and other data acquired by eButton using the automatic data segmentation and interactive categorization methods described previously (Section 3.2 and Figure 3). Although images are recorded in the “first person”, the activity being performed is usually recognizable from the image content, clock time, degree of motion, and the activity location. Once the activity is determined, the physical activity compendium, an extensive database updated recently [35], is utilized to compute the MET value for the activity. Finally, the energy expenditure (in calories) is obtained using the MET value and some demographical information of the individual, including gender, age and bodyweight.



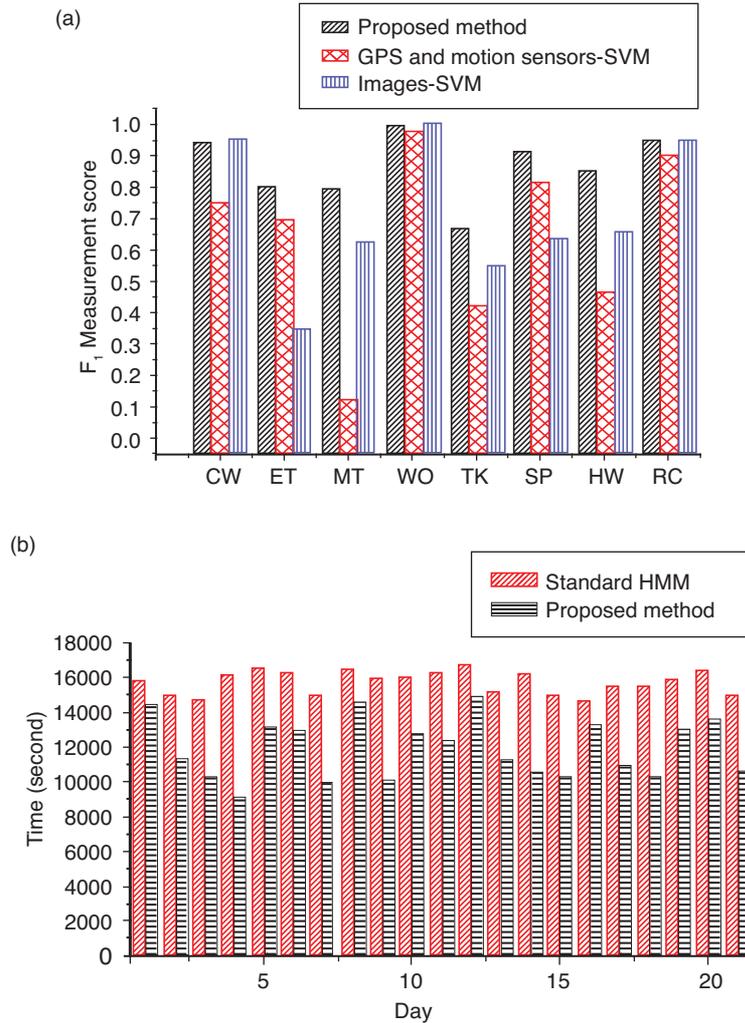
**Figure 11.** Flow diagram of camera-based method for physical activity evaluation.

### ***Method 3: Automatic Activity Recognition***

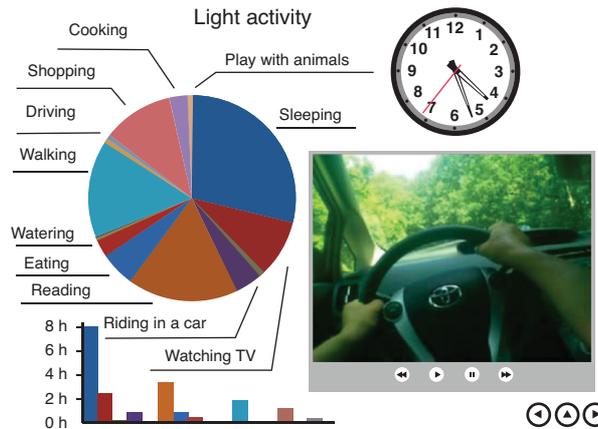
In general, it is very difficult to recognize human activities automatically since they are complex and numerous. However, if we reduce the recognition task to targeting only frequent activities performed by a specific individual, the recognition problem can be simplified greatly. Since eButton is a personalized device, this simplification is meaningful. Based on this concept, we have developed an automatic method for physical activity recognition [36]. In this method, we first ask the individual to complete a table specifying the likelihood of activities (e.g., breakfast, driving, office work, lunch, etc.) at any given time during the day. In order to reduce variance, we ask the individual to use different tables for workdays and weekends/holidays for people who follow regular daytime working hours during weekdays [16]. A prior probability table of activities is then formed based on the information provided. Next, we associate each activity with certain sensors which are most effective in recognizing the activity. For example, a camera is most effective in recognizing TV watching, a GPS sensor works best to recognize driving. We also use the adaptive Hidden Markov Model (HMM) to choose suitable features from which the activity is recognized based on the use of the support vector machines (SVMs) [37]. In this way, only a small, but relevant part of the data are examined for each activity, thus greatly reducing the computational cost. Experiments were performed in which eight frequent activities were recognized [36]. The results of methods with and without using prior information were compared. Our results indicated that the proposed method performed better in identifying physical activities (Figure 12a) at a lower computation cost, i.e., shorter computation time (Figure 12b).

### **4.3. Lifestyle Evaluation**

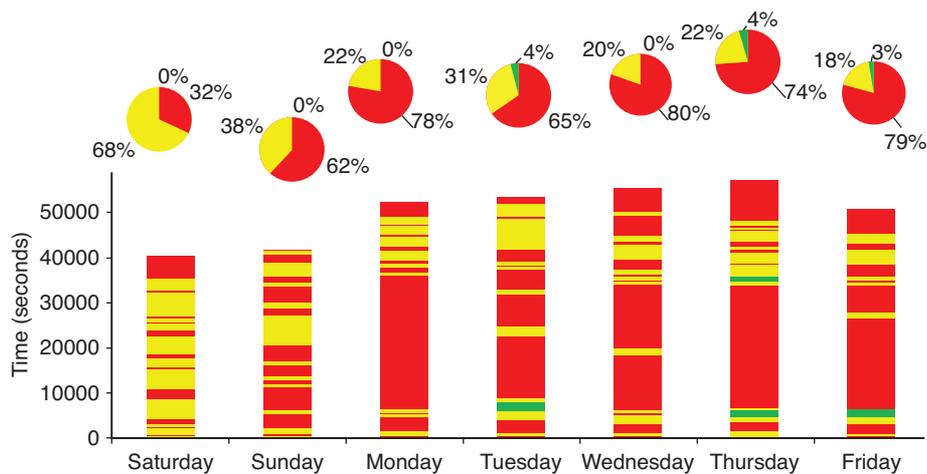
In the context of health, lifestyle includes diet, physical activity, hygiene, living environment, social interaction, mental health, etc. It is clear that the eButton cannot manage all the information about lifestyle. However, it is capable of documenting most daily events experienced by an individual. The result can be used to evaluate whether the activity type and duration benefit health. We performed a study on nine adult research participants who wore eButton for a total of 33 days [39]. The average wearing time was 11.2 hours per day. The data were first partitioned into “shots” using the method described in Section 3.2. Then, with our software, the events in these “shots” were labeled manually and condensed into an interactive video summary shown in Figure 13. The size of each piece of the pie graph represents the energy intake/expenditure of the corresponding activity. A mouse click on each piece triggers a video play of the image sequence, allowing observation and extraction of the lifestyle information. Finally, we integrate the daily information into a weekly summary in the form shown in Figure 14 where the horizontal and vertical axes represent, respectively, days and times spent in three categories of activities: sedentary, light and moderate. Despite the rough categorization, it can be observed that this research participant (a software engineer) had a problem of physical inactiveness, since a majority of her daytimes was in red regions (working on computers), sometimes for a very long period of time continuously (e.g., approximately 7 hours on Monday). The non-sedentary



**Figure 12.** (a) Accuracy measured by the  $F_1$  score (defined as  $(2 \times \text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$  [38]) for eight physical activities (CW: working on computer; ET: eating; MT: meeting; WO: outdoor working; TK: talking/socializing; SP: shopping; HW: housework; RC: riding in a car as a passenger; OT: other activities); (b) Computational time spent for activity recognition.



**Figure 13.** Snapshot of an interactive video summary of a single day of events experience by an individual.



**Figure 14.** Summary of daily activities for a female subject during one week. The activities are roughly categorized into sedentary (red), light (yellow), and moderate (green). For this individual, computer use is the predominant daily activity. The horizontal and vertical axes represent day of the week and time (in seconds with 0 being the starting time of wearing the eButton), respectively.

times during the week were almost all spent on light activities (yellow regions), such as walking and cooking. There was only a tiny portion of the time spent on moderate activities. This graph thus provides objective data for the individual to improve lifestyle, i.e., taking breaks during work hours and being more physically active.

## 5. DISCUSSION

We believe that the mobile technology is entering a new phase consisting of not only the smartphone, but also a variety of handheld, wearable and implantable smart devices. As this technology advances, there will be a profound transition from the current form of symptom-based healthcare to a new form emphasizing on lifestyle modification and disease prevention. Our investigation on eButton aims at accelerating this transition. Although some encouraging results have been achieved as described above, there are still many unsolved problems and unanswered questions.

We think that the issue of privacy is very important which requires not only continuous technical study, but also new policies, regulations and consensus in the public. The most sensitive data recorded by eButton are GPS locations and images. Technically, we must find ways to acquire only the data relevant to the intended study with the least invasion to the privacy of the device wearer and other people. Another technical approach is to automate data processing and avoid human observation of the recorded data. However, both approaches are difficult to implement, even with the current state-of-the-art computational algorithms.

Another significant challenge is the power supply for the eButton. The battery can currently last from 4 to 8 hours, depending on the peripheral sensors and the capacity of the rechargeable battery. A larger battery is not feasible since it is inconvenient to wear. Therefore, power management is necessary to reduce power consumption. The most power-hungry components in the eButton at present are the central microprocessor and the cameras. We have been using power management techniques, such as lowering the clock rates and putting the sensors to sleep at certain times. However, the effects are limited since the eButton is often required to work continuously. Power-efficient cameras that fit in the device with satisfactory image quality and field of view are presently unavailable. As the microelectronic technology advances, more efficient processors and cameras suitable for use in wearable computers are expected to be developed in the near future.

The structural design of the wearable device is a critical issue for its practical application. Many forms of devices with different wearing methods have been explored, such as an eyeglass frame (e.g., Google Glass), a wristwatch or a wristband (i.e., Samsung Smartwatch), an armband (e.g., Bodymedia Fit), a waistbelt clip (e.g., Actigraph), a shoe insert (i.e., Nike-iPot sensor), a net worn device with a lanyard (e.g., SenseCam), and a chest worn device (e.g., eButton). Among them, we believe that the head and upper chest locations in the body are the most suitable locations for camera. When compared to the eyewear design, the chest pin design appears to be more comfortable for long-term wearing, more passive in its operation, more natural in appearance, and more tolerable to device size and weight. However, the eyewear has a distinct advantage of automatically steering the view of the camera by natural head

rotation. In addition, the eyewear allows the use of a display unit which projects an image to the retina directly from the eyewear. In all forms of wearable device designs, however, the lack of a convenient, unobtrusive human-computer interface is still an unsolved problem. This problem represents a significant bottleneck in wearable computer applications to the healthcare field.

An important question arises whether the wearable computer technology is ready to be used in the real world. At this stage, we have already conducted small-scale studies on diet, physical activity, and real-life events. Other applications are also being explored by our group, including those for the blind [14, 40], elders and children [41, 42], either being studied in our laboratory or initiated for real-world data collection. We believe that, in terms of technology, the wearable computer is close to real-world adoption. However, its readiness depends on a number of practical issues, including the attractiveness and convenience in the structural design, streamline of unitization procedure, improvement in energy efficiency, and enhancement of automatic data processing. There are also other critical factors such as reimbursements, regulatory issues, privacy concerns by the public, and commercial profitability of the manufacturers. These important factors affect the adoption of the new technology tremendously. However, they are beyond the scope of this paper.

## **6. CONCLUSION**

In this paper, we have presented our research on wearable computing for applications to personal health. A miniature wearable computer called eButton has been designed as a circular button worn on the chest or other locations. This device contains a powerful microprocessor and an array of sensors to acquire multiple forms of data which support research studies on diet, physical activity and lifestyle. We have described the hardware design, key computational methods, and the performance of the eButton for real-world data collection with respect to specific applications. We have also discussed the current limitations of the device and potential future improvements.

## **ACKNOWLEDGEMENTS AND DISCLAIMER**

The authors would like to acknowledge the valuable contributions from numerous faculty members, research staff members and undergraduate/graduate students who were involved in this research project. This work was supported by National Institutes of Health grants No. U01HL91736, R01CA165255, and R21CA172864. The opinions expressed in this article are the author's own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government. The fifth author acknowledges research support by China National Science Foundation Grant No. 61272351.

## **CONFLICT OF INTEREST**

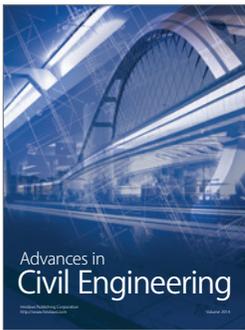
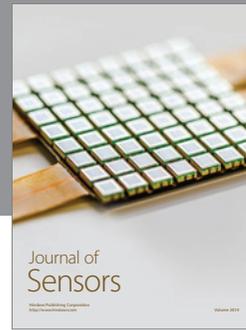
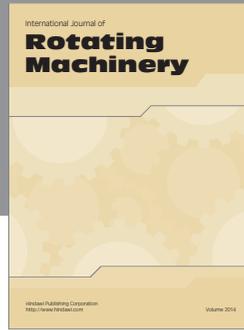
The authors declare that there is no conflict of interest.

## REFERENCES

- [1] Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences*, 2009, 106(23):9362–9367.
- [2] Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *The New England Journal of Medicine*, 2000, 343(2):78–85.
- [3] Willett WC. Balancing life-style and genomics research for disease prevention. *Science*, 2002, 296(5568):695–698.
- [4] Rappaport SM, Smith MT. Epidemiology. environment and disease risks. *Science*, 2010, 330(6003):460–461.
- [5] Rappaport SM. Implications of the exposome for exposure science. *Journal of Exposure Science and Environmental Epidemiology*, 2011, 21(1):5–9.
- [6] Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. *JAMA*, 2010, 303(3): 235–241.
- [7] Overweight and obesity: a major public health issue, U.S. Department of Health and Human Services. 2001. Available: <http://odphp.osophs.dhhs.gov/pubs/prevrpt/01fall/pr.htm>
- [8] Hammond RA, Levine R. The economic impact of obesity in the United States. *Journal of Diabetes, Metabolic Syndrome and Obesity*, 2010, 3:285–295.
- [9] Magic. Pure Magic. <http://mealsnap.com/>. Accessed July 29, 2014.
- [10] Boushey CJ, Kerr DA, Wright J, Lutes KD, Ebert DS, Delp EJ. Use of technology in children’s dietary assessment. *European Journal of Clinical Nutrition*, 2009, 63 (Suppl 1):S50-57.
- [11] Lee CD, Chae J, Schap TE, Kerr DA, Delp EJ, Ebert DS, Boushey CJ. Comparison of known food weights with image-based portion-size automated estimation and adolescents’ self-reported portion size. *Journal of Diabetes Science and Technology*, 2012, 6(2):428–434.
- [12] Daugherty BL, Schap TE, Ettienne-Gittens R, Zhu FM, Bosch M, Delp EJ, Ebert DS, Kerr DA, Boushey CJ. Novel technologies for assessing dietary intake: evaluating the usability of a mobile telephone food record among adults and adolescents. *Journal of Medical Internet Research*, 2012, 14(2):e58.
- [13] Lee IM, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. *British Journal of Sports Medicine*, 2014, 48(3):197–201.
- [14] Bai Y, Jia W, Zhang H, Mao ZH, Sun M. Helping the blind to find the floor of destination in multistory buildings using a barometer. *Proceedings of 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, 4738–4741.
- [15] Zhang W, Jia W, M. Sun. Segmentation for efficient browsing of chronical video recorded by a wearable device. *Proceedings of 36th Northeast Biomedical Engineering Conference*, 2010.
- [16] Li Z, Wei Z, Jia W, Sun M. Daily life event segmentation for lifestyle evaluation based on multi-sensor data recorded by a wearable device. *Proceedings of 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, 2858–2861.
- [17] Jia W, Chen HC, Yue Y, Li Z, Fernstrom J, Bai Y, Li C, Sun M. Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. *Public Health Nutrition*, 2014, 17(8):1671–1681.
- [18] Ahuja JKA, Montville JB, Omolewa-Tomobi G, Heendeniya KY, Martin CL, Steinfeldt LC, Anand J, Adler ME, LaComb RP, Moshfegh AJ. [Online]. USDA Food and Nutrient Database for Dietary Studies, 5.0..U.S. Department of Agriculture, Agricultural Research Service, Food Surveys Research Group, Beltsville, MD.
- [19] Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6):679–698.

- [20] Nie J, Wei Z, Jia W, Li L, Fernstrom JD, Sclabassi RJ, Sun M. Automatic detection of dining plates for image-based dietary evaluation. *Proceedings of 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010, 4312–4315.
- [21] Nie J, Fernstrom JD, Sclabassi RJ, Fernstrom MH, Wei Z, Li L, Zhang W, Jia W, Mao Z-H, Sun M. Automatic detection of dining plates in digital video. *Proceedings of 36th Northeast Biomedical Engineering Conference*, 2010.
- [22] Sun X, Yao H, Jia W, Sun M. Eating activity detection from images acquired by a wearable camera. *Proceedings of ACM SenseCam and Pervasive Imaging*, 2013, 80–81.
- [23] Perazzi F, Krähenbühl P, Pritch Y, Hornung A. Saliency filters: contrast based filtering for salient region detection. *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, 2012, 723–740.
- [24] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11):1254–1259.
- [25] Chen HC, Jia W, Yue Y, Li Z, Sun YN, Fernstrom JD, Sun M. Model-based measurement of food portion size for image-based dietary assessment using 3D/2D registration. *Measurement Science and Technology*, 2013, 24(10):105701.
- [26] Levi Z, Gotsman C. D-Snake: image registration by as-similar-as-possible template deformation. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(2):331–343.
- [27] JHaO S. A four-step camera calibration procedure with implicit image correction. *Proceedings of Computer Vision and Pattern Recognition*, 1997, 1106–1112.
- [28] Tsai R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 1987:323–344.
- [29] Zhang Z. Flexible camera calibration by viewing a plane from unknown orientations. *Proceedings of IEEE International Conference on Computer Vision*, 1999, 666–673.
- [30] Janz KF. Physical activity in epidemiology: moving from questionnaire to objective measurement. *British Journal of Sports Medicine*, 2006, 40(3):191–192.
- [31] Wareham NJ, Rennie KL. The assessment of physical activity in individuals and populations: Why try to be more precise about how physical activity is assessed? *International Journal of Obesity*, 1998, 22: S30-S38.
- [32] Kerr J, Marshall SJ, Godbole S, Chen J, Legge A, Doherty AR, Kelly P, Oliver M, Badland HM, Foster C. Using the SenseCam to improve classifications of sedentary behavior in free-living settings. *American Journal of Preventive Medicine*, 2013, 44(3):290–296.
- [33] ActiGraph Service: What are counts? <https://help.theactigraph.com/entries/20723176-what-are-counts>. Accessed July 29, 2014.
- [34] Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, O'Brien WL, Bassett DR, Jr., Schmitz KH, Emplaincourt PO, Jacobs DR, Jr., Leon AS. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise*, 2000, 32(9 Suppl): S498-504.
- [35] Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr., Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. Compendium of physical activities: a second update of codes and MET values. *Medicine & Science in Sports & Exercise*, 2011, 43(8):1575–1581.
- [36] Li Z. Study on analytic methods for human physical activity recognition based on wearable systems. PhD dissertation, Ocean University of China, 2013.
- [37] Burges C. A tutorial on support vector machines for pattern recognition. *Journal of Data Mining and Knowledge Discovery*, 1998, 2:121–167.
- [38] Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2011, 2(1):37–63.
- [39] Jia W, Sun M. An event-based approach to daily physical activity evaluation using mobile devices. presented at the 2011 mHealth Summit, Washington DC, 2011.

- [40] Bai Y. A wearable indoor navigation system for the blind and visually impaired individuals. PhD Dissertation, University of Pittsburgh, 2014.
- [41] PandaCare: Demo Preparing. <https://blogs.cornell.edu/cornellcup2013pandacare/>. Accessed July 29, 2014.
- [42] Baranowski T, Islam N, Baranowski J, Martin S, Beltran A, Dadabhoy H, Adame SH, Watson KB, Thompson D, Cullen KW, Subar AF. Comparison of a Web-Based versus Traditional Diet Recall among Children. *Journal of the Academy of Nutrition and Dietetics*, 2012, 112(4):527–532.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

