Journal of Food Engineering 109 (2012) 76–86

Contents lists available at SciVerse ScienceDirect

Journal of Food Engineering

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

Imaged based estimation of food volume using circular referents in dietary assessment

Wenyan Jia [a], Yaofeng Yue [b], John D. Fernstrom [c], Ning Yao [b], Robert J. Sclabassi [d], Madelyn H. Fernstrom [c], Mingui Sun [a,b,]*

[a] Department of Neurosurgery, University of Pittsburgh, PA, United States
[b] Department of Electrical and Computer Engineering, University of Pittsburgh, PA, United States
[c] Department of Psychiatry, University of Pittsburgh, PA, United States
[d] Computational Diagnostics Inc., Pittsburgh, PA, United States

ARTICLE INFO

Article history:
Received 6 January 2011
Received in revised form 22 April 2011
Accepted 28 September 2011
Available online 7 October 2011

Keywords:
Dietary assessment
Volume estimation
Circular referent
Image processing

ABSTRACT

Measuring food volume (portion size) is a critical component in both clinical and research dietary studies. With the wide availability of cell phones and other camera-ready mobile devices, food pictures can be taken, stored or transmitted easily to form an image based dietary record. Although this record enables a more accurate dietary recall, a digital image of food usually cannot be used to estimate portion size directly due to the lack of information about the scale and orientation of the food within the image. The objective of this study is to investigate two novel approaches to provide the missing information, enabling food volume estimation from a single image. Both approaches are based on an elliptical reference pattern, such as the image of a circular pattern (e.g., circular plate) or a projected elliptical spotlight. Using this reference pattern and image processing techniques, the location and orientation of food objects and their volumes are calculated. Experiments were performed to validate our methods using a variety of objects, including regularly shaped objects and food samples.

1. Introduction

Accurate dietary assessment is an important procedure in both clinical and research studies of many diseases and disorders, such as obesity and diabetes (Franz et al., 2002; Mokdad et al., 2004; Knowler et al., 2009; Thompson and Subar, 2001). However, a major problem exists: the assessment is often inaccurate for individuals in a real life setting (Thompson and Subar, 2001; Wacholder, 1995; Trabulsi and Schoeller, 2001; Johansson et al., 1998). The self-reporting of food intake (e.g., food record, 24-h dietary recall and food frequency questionnaire) has been the primary method for acquiring information about consumed food and the portion size (Thompson and Subar, 2001; Wacholder, 1995; Trabulsi and Schoeller, 2001; Johansson et al., 1998; Goris and Westerterp, 1999; Goris et al., 2000; Most et al., 2003; Pietinen et al.,1988a,b; Chu et al., 1984; Martin et al., 2007). However, self-report methods are known to be biased, due to significant underreporting (Trabulsi and Schoeller, 2001; Johansson et al., 1998; Goris and Westerterp, 1999; Goris et al., 2000). In order to improve assessment accuracy, many portion-size measurement aids (PSMAs) have been developed (Posner et al., 1992; Faggiano et al., 1992; Kuehneman et al., 1994; Nelson et al., 1994; Lucas et al., 1995; Cypel et al., 1997; Godwin et al., 2006; Nelson and Haraldsdottir, 1998; Frobisher and Maxwell, 2003; Williamson et al., 2003; Turconi et al., 2005; Ershow et al., 2007; Foster et al., 2005; Martin et al., 2009). Common PSMAs include household measures (e.g., rulers), food models (real food samples or abstract models), and food photographs. While such PSMAs are useful in estimating portion size, they provide only a rough volumetric measure and are not always suitable or convenient in practice. Recently, camera-enabled cell phones and other mobile devices (e.g., iPad) have become very popular. These devices can be used to take food pictures conveniently, which can then be remotely transmitted to dietitians for dietary analysis (Kikunaga et al., 2007; Boushey et al., 2009; Zhu et al., 2010; Ngo et al., 2009).

Although digital images provide an excellent tool to document foods consumed, it is difficult to estimate food portion size from a single image without a dimensional referent, except in special cases (Du and Sun, 2006; Wang and Nguang, 2007). Accordingly, several approaches have been studied that include a dimensional referent in the field of view, such as a checkerboard (Boushey et al., 2009; Zhu et al., 2010; Subar et al., 2010; Sun et al., 2008, 2010a,b; Jia et al., 2009; Yue et al., 2010). However, it can be inconvenient and annoying for subjects constantly to have to remember

* Corresponding author at: Department of Neurosurgery, University of Pittsburgh, PA, United States. Tel.: +1 412 802 6481; fax: +1 412 802 6785.
E-mail address: drsun@pitt.edu (M. Sun).

0260-8774/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jfoodeng.2011.09.031

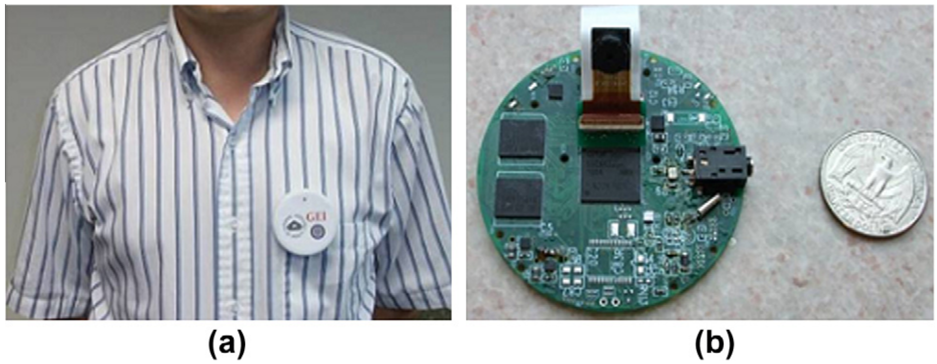

Fig. 1. (a) Subject wearing the device; (b) electronic circuit board within the device.

to carry a checkerboard, and to place it beside the food when eating (Boushey et al., 2009; Zhu et al., 2010; Subar et al., 2010). Further, eating behavior changes when subjects are asked to document food intake. For example, they may reduce intake significantly, an effect frequently observed in obese subjects (Goris and Westerterp, 1999; Goris et al., 2000). In order to obtain objective data on food intake, we have developed a low-cost electronic device containing a miniature camera and a number of other sensors to record food intake automatically (Sun et al., 2008, 2009, 2010a,b). This device is designed to be almost completely passive to the subject, and thus hopefully not intrude on or alter the subject's eating activities. This button-like device can be worn using a lanyard around the neck or attached to clothes using a pin or a pair of magnets (see Fig. 1). When utilized, it takes pictures of the scene in front of the subject at a pre-set rate (e.g., 2–4 s/frame). All the images are stored on a memory card. After the dietary study is complete (e.g., 1 week), the data are transferred to the dietitian's computer for further analysis. Since images of the entire eating period are recorded (see Fig. 2), a pair of images can always be selected to represent each food item before and after eating (Fig. 2(a) and (d)). Once each food item is identified by the dietitian and its consumed volume is determined from the "before" and "after" images, our software system calculates calorie and nutrient values for each food using data available from the United States Department of Agriculture (USDA) Food and Nutrient Database for Dietary Studies (FNDDS) database (USDA, 2010). The energy intake during the study period can be accumulated and a detailed report on calorie and nutrient intake created for dietary assessment.

In FNDDS, the conversion between measurable/observable volumes (e.g., cup, cubic inch, piece, slice, large/medium/small) and weight is available for most foods, such as cheeses, breads, vegetables and meats. For example, the weights of one cubic inch cheese and a cup of stewed pork chop are 17.3 and 134 g, respectively. Therefore the energy and nutrient contents of a specific food can be estimated based on its volume, as calculated from an image. In the current version of the FNDDS database, certain foods are listed without volumetric measures. In order to solve this problem, several research groups, including us, have been collaborating with USDA to conduct research on food density measures.

In the current study, we focus on image-based food volume estimation and present two novel methods for calculating food volume from a single image. Previous studies have shown that, under certain conditions, the elliptical image of a known circular object can provide sufficient information about the location and orientation of any planar object (Shin and Ahmad, 1989; Safaee-Rad et al., 1992). In our study, we have proposed two types of approaches based on the elliptical patterns extracted from a single input image (Jia et al., 2009; Yue et al., 2010). The first type is based on a circular object (e.g., a circular plate) which is commonly present during eating. We call the corresponding portion size estimation method the "plate method". The second type is based on a conic spotlight projected by a light emitting diode (LED) installed on our wearable electronic device. We call this method the "LED method". In either method, we first determine the location and orientation of an object plane by interactive elliptical pattern extraction and computation using our algorithms. Then, dimensional variables (e.g., length and thickness) are measured by selecting several feature points on the food image. The measured dimensions provide an estimate of food volume. In the following, we will describe our methods in detail and discuss our experimental results.

2. Methods

We model the camera using the pinhole model which is commonly utilized in photography (Sonka et al., 2007; Hartley and Zisserman, 2003). For simplicity, we assume that the camera has a fixed focal length, which is true in our wearable device and most cell phones; and that the physical image plane of the camera is perpendicular to its optical axis, which is almost universally true. Under these assumptions, the intrinsic parameters of the camera (e.g., the focal length) can be obtained either from manufacturer's

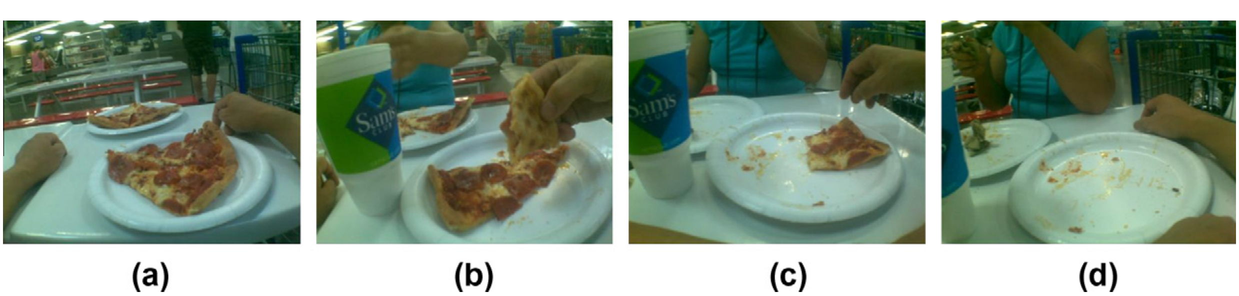

Fig. 2. Typical pictures before (a), during (b–c) and after eating (d).