Pervasive and Mobile Computing 8 (2012) 147–163

Contents lists available at SciVerse ScienceDirect

Pervasive and Mobile Computing

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

DietCam: Automatic dietary assessment with mobile camera phones

Fanyu Kong *, Jindong Tan

Department of Electrical and Computer Engineering, Michigan Technological University, 1400 Townsend Drive, Houghton, MI, 49931, United States

ARTICLE INFO

Article history:
Received 14 May 2010
Received in revised form 28 January 2011
Accepted 18 July 2011
Available online 26 July 2011

Keywords:
Food intake assessment
Calorie estimation
Mobile phones

ABSTRACT

Obesity has become a severe health problem in developed countries, and a healthy food intake has been recognized as the key factor for obesity prevention. This paper presents a mobile phone based system, DietCam, to help assess food intakes with few human interventions. DietCam only requires users to take three images or a short video around the meal, then it will do the rest. The experiments of DietCam in real restaurants verify the possibility of food recognition with vision techniques.

1. Introduction

Mobile phones are becoming a popular and powerful platform, and many healthcare-related applications have been explored, such as remote health monitoring, SMS medical tips, fitness coaches, and diabetes guides [1]. Obesity, another possible cell phone aided healthcare problem, is becoming an epidemic phenomenon in most developed countries. In the past three decades, obesity rates for both adults and children in the US have increased significantly [2]. The fact that more than thirty-three percent of adults and sixteen percent of children are obese has proven to be one of the biggest public health challenges to the general population and social welfare. The serious consequences of obesity include severe health problems such as diabetes, stroke, and heart disease, and high-cost healthcare bills, which were estimated at $147 billion in 2008 in the US alone [3,4]. The continuing increase of overweight and obesity attributable spending has attracted increasing research interest to explore practical new technology to prevent obesity. In spite of the common sense that obesity is a complex condition caused by the interaction of many factors such as genetic makeup, secondary effects from medical treatment, and calorie imbalance, it is generally believed that obesity prevention requires individuals to foster life-long healthy food choices and regular physical activities [5]. However, the usual case is that individuals with potential obesity problems are more likely to ignore their food intakes and regular exercise. Even people who care and pay attention to nutrition information may not be sufficiently knowledgeable about the calorie content of what they are eating. Efforts have been made to record calorie contents without user awareness or knowledge by processing chewing sounds of the user with on-body sensors [6]. However, the accuracy of food content recovery from audio signals is still questionable, and it presents the users with a lot inconvenience when wearing sensors over the neck all day long.

Opportunities for novel obesity management applications arise as mobile phones are becoming more powerful for people-centric computing. The fact that mobile phones nowadays are necessary and are carried by people almost everywhere makes them perfect devices for information gathering and delivering during free living conditions. Cameras, which are equipped on most smart phones, can provide rich and reliable information. Another powerful extension of mobile technology is the combination of accelerometers, which benefit in creating valid measures of physical activities. Even though obesity and diabetes related mobile phone applications have appeared, most of them only use the mobile phone as a food

* Corresponding author.
E-mail addresses: fkong@mtu.edu (F. Kong), jitan@mtu.edu (J. Tan).

1574-1192/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.pmcj.2011.07.003

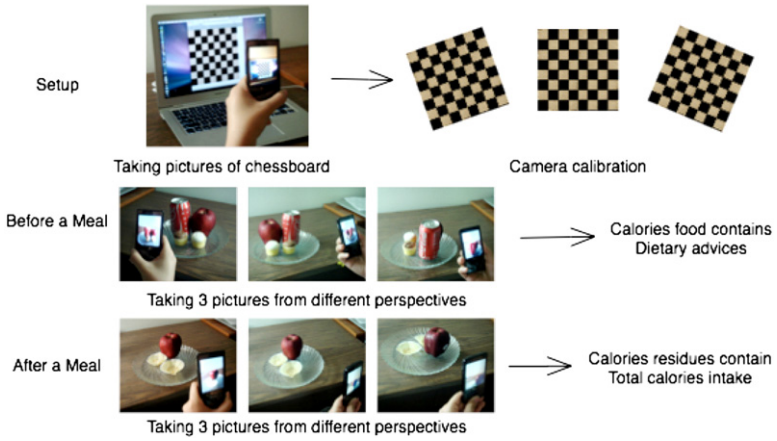

Fig. 1. Expected usage. The calorie information, which is a key to the obesity problem, will be extracted from three images or a short piece of video of the foods.

diary [7–10] or fitness diary [11–13] that requires large amounts of user input. Cameras help record dietary information automatically [14], but users still have to manually review the processed image results. We have developed a health-aware smart phone system which employs an obese prevention application utilizing the embedded camera and accelerometer. Besides extracting physical activity data through built-in accelerometer readings, it monitors food intake automatically with few user interventions.

In this paper, the automatic food calorie estimation system DietCam in a health-aware system is proposed, as shown in Fig. 1. It is able to recognize foods and calculate the calorie content of a meal automatically from images or videos with few human interventions. Before it is in use, the camera on the cell phone needs to be calibrated in a user-friendly way. When utilizing DietCam, users only have to put a credit card beside the plate and take three pictures around the dish approximately every 120° or shoot a piece of video. After that, DietCam will do the rest for the users to obtain the calorie information. Vision techniques are utilized to extract visual cues of the calorie information from images or the piece of video (if equipped with a digital compass) around the plate. Based on these visual cues, food recognition algorithms are designed to classify the food items. At the same time, three-dimensional (3D) models of visible food items will be reconstructed in order to estimate the volume of the food. The metric scale of the 3D model is inferred from the credit card. Types, volumes, and calorie densities of the food items together identify the calorie content of the meal.

The accurate measurement of food contents through vision techniques is challenging. At present, there is no technology that allows users to estimate the calorie content of a meal automatically and comfortably. The following challenges exist in this project:

1. Many different kinds of food have the same or very similar appearance that is hard to distinguish from a camera's point of view.
2. Even though some kinds of food have specific appearances, the diversity of the same kinds of food makes it impossible to recognize all these foods.
3. A meal usually has more than one food items. It is hard to segment those foods with irregular shapes, especially when occlusions exist in the image. The varying lighting conditions of restaurants make this problem even harder.
4. Even if the types of food have been recognized correctly, the amount of food is another factor affecting the calorie intake directly. Sometimes people will not eat the whole meal. It is necessary to estimate the portion consumed.
5. Even though all the above challenges are solvable by carefully designed algorithms, is it practical to implement these algorithms on a mobile phone?

Our technique addresses these challenges by utilizing a multiple-view method. The approaches are lightweight and feasible on a commercial smart phone. A prototype has been implemented on an iPhone, and the results are promising. Our main contributions are as follows.

1. *Identifying the possibility of obtaining calorie information of a meal through a camera phone.* A prototype has been implemented on an iPhone. The algorithms are under study on Windows Mobile, Android, RIM and Symbian platforms.
2. *Developing multiple-view image understanding algorithms for contents recovery.* We perform simple feature extraction on multiple images. Novel segmentation, classification, occlusion, and correspondence handling algorithms are developed for food classification. A model based volume estimation mechanism is developed.
3. *Evaluation of the scheme at home and in real restaurant locations.* We collect test samples at home, different local restaurants, and supermarkets with different combinations of food items and at different times of a day. As many as 21 business restaurants are covered. An average recognition accuracy of 92% is achieved.

The rest of this paper solves each of the challenges and elaborates on these contributions. Section 7 evaluates DietCam with field experiments. We discuss the related work in Section 9, and conclude the paper in Section 10.