An Overview of The Technology Assisted Dietary Assessment Project at Purdue University

Nitin Khanna∗, Carol J. Boushey†, Deborah Kerr‡, Martin Okos§, David S. Ebert∗, and Edward J. Delp∗
∗School of Electrical and Computer Engineering
†Department of Foods and Nutrition
‡School of Public Health
§Department of Agricultural and Biological Engineering
∗†§Purdue University, West Lafayette, Indiana, USA
‡Curtin University of Technology, Perth, Australia

Abstract—In this paper, we describe the Technology Assisted Dietary Assessment (TADA) project at Purdue University. Dietary intake, what someone eats during the course of a day, provides valuable insights for mounting intervention programs for prevention of many chronic diseases such as obesity and cancer. Accurate methods and tools to assess food and nutrient intake are essential for research on the association between diet and health. An overview of our methods used in the TADA project is presented. Our approach includes the use of image analysis tools for identification and quantification of food that is consumed at a meal. Images obtained before and after foods are eaten are used to estimate the amount and type of food consumed.

Keywords—dietary assessment, diet record method, mobile telephone, mobile device, classification, pattern recognition, image texture, feature extraction, volume estimation.

I. INTRODUCTION

Dietary intake, what someone eats during the course of a day, provides valuable insights for mounting intervention programs for prevention of many chronic diseases such as obesity and diabetes [1], [2]. Accurate methods and tools to assess food and nutrient intake are essential for research on the association between diet and health. Conventional methods of dietary assessment, including the 24-hour dietary recall, dietary record, and food frequency questionnaire (FFQ) [3]–[7] have been described as burdensome by adolescents and adults. Children and adolescents, however, are eager in terms of adopting new technologies [8], [9]. In a survey of 29 participants in completing a questionnaire, 58.6% (17/29) owned a mobile telephone and 65.5% (19/29) reported owning a digital camera [8], [9]. When testing a prototype PDA food record tool during Camp Calcium 2005, a research study at Purdue University to determine the calcium requirements of adolescents, with 31 adolescent participants, almost every child indicated previous experience using a PDA and readily adapted to using the tool. Recent consumer reports indicate that 75% of adolescents aged 12-17 years have their own mobile telephones [8], [9]. This is a substantial increase since 2004 when only 45% of adolescents had a mobile telephone. Mobile telephones have become “indispensable tools in teen communication patterns” [8], [9]. Young people desire a mobile telephone that allows for taking and sending pictures, playing games, social networking, and texting [8], [9]. Mobile devices (e.g., a mobile telephone or PDA-like device) have evolved to meet market demand for general purpose mobile computing. Their high-speed multimedia processors and data network capability also make mobile devices ideal as a field data collection tool for dietary assessment.

The availability of “smart” mobile telephones with higher resolution imaging capability, improved memory capacity, network connectivity, and faster processors allow these devices to be used in health care applications. Mobile telephones can provide a unique mechanism for collecting...
II. REVIEW OF CURRENT DIETARY ASSESSMENT METHODS

A review of some of the most popular dietary assessment methods is provided in this section. The objective here is to analyze advantages and major limitations of these methods. This will demonstrate the significance of our mobile telephone food record (mpFR).

A. 24-Hour Dietary Recall

The 24-hour dietary recall (24HR) consists of a listing of foods and beverages consumed the previous day or the 24 hours prior to the recall interview. Foods and amounts are recalled from memory with the aid of an interviewer who has been trained in methods for soliciting dietary information. A brief activity history may be incorporated into the interview to facilitate probing (i.e. asking questions) for foods and beverages consumed. The Food Surveys Research Group (FSRG) of the United States Department of Agriculture (USDA) has devoted considerable effort for improving the accuracy of this method.

The major drawback of the 24HR is the issue of underreporting of the food consumed [16]. Youth, in particular, are limited in their abilities to estimate portion sizes accurately [3]. The most common method of evaluating the accuracy of the 24HR with children is through observation of school lunch and/or school breakfast [17] and comparing foods recalled with foods either observed as eaten or foods actually weighed. These recalls have demonstrated both under-reporting and over-reporting, and incorrect identification of foods.

B. The Food Record

The 24HR is useful in population based studies; however the preferred dietary assessment method for clinical studies is the food record. Depending on the primary nutrient or nutrients or foods of interest, the minimum number of food records needed is rarely less than two days. Training the participants, telephoning with reminders for recording, reviewing the records for discrepancies, and entering the dietary information into a nutrient database can take a large amount of time and requires trained individuals [18]. The food record is especially vulnerable to underreporting due to the complexity of recording food [19], [20].

Portion size estimation may be one contributor to underreporting. In [21] it was found that 45 minutes of training in portion-size estimation among 9-10 year olds significantly improved estimates for solid foods which were measured by dimensions or cups, and liquids estimated by cups. Amorphous foods were estimated least accurately even after training and some foods still exhibited an error rate of over 100%. Thus, training can improve portion size estimation, however more than one session may be needed and large/consistent accuracy may be unattainable. Another challenge in evaluating dietary assessment methods is comparing the results of the dietary assessment method to some measure of “truth”. This is best achieved by identifying a biomarker of a nutrient or dietary factor [20], [22]. The underlying assumption of a biomarker is that it responds to intake in a dose-dependent relationship. A biomarker does not rely on a self-report of food intake, thus theoretically the measurement errors of the biomarker are not likely to be correlated with those of the dietary assessment method.
As one can see from the above discussion, measuring accurate dietary intake is considered to be an open research problem in the nutrition and health fields. There is a tremendous need for new methods for collecting dietary information. Preliminary studies have indicated that the use of a mobile device using a camera to obtain images of the food consumed may provide a more accurate method for dietary assessment. This is the goal of Technology Assisted Dietary Assessment (TADA) mobile telephone food record (mpFR) application [12] that we have developed for a mobile device. Figure 1 shows the overall architecture of our system. The next section described the key signal processing components of this system.

III. IMAGE ANALYSIS

There has been previous work reported for automatic recognition of some types of food items. In [23] an automatic fruit recognition system is described which recognized spherical fruits in different situations such as shadows, bright areas, occlusions and overlapping fruits. This system used a three-dimensional scanner to scan the scene and generate five images to represent the azimuth and elevation angles, range, attenuation and reflectance. The position of the fruits obtained by thresholding and clustering. Circular Hough Transform was used to identify the center and radius of the fruits. A robust method to segment the food items from the background of color images was proposed in [24]. A color image was converted to a high contrast grayscale image from an optimal linear combination of the RGB color components. The image is then segmented using a global threshold estimated by a statistical approach to minimize the intraclass variance. The segmented regions were subjected to a morphological process to remove small objects, to close the binary image by dilation followed by erosion and to fill the holes in the segmented regions.

Using a single image acquired from the mobile device, our goal is to automatically determine the regions in an image where a particular food is located (segmentation) and correctly identify the food type based on its visual features (classification or food labeling) [12], [14]. The system must be easy to use and not place a burden on the user by having to take multiple images, carry another device or attaching other sensors to their mobile device. Our approach is shown in Figure 2. Each food item is segmented, identified, and its volume is estimated. “Before” meal and “after” meal images can be used to estimate the food intake. From this information, the energy and nutrients consumed can be determined.

Automatic identification of food items in an image is not an easy problem. We fully understand that we will not be able to recognize every food. Some food items look very similar, e.g. margarine and butter. In other cases, the packaging or the way the food is served will present problems for automatic recognition. For example, if the food is in an opaque container then the we will not be able to identify it.

In some cases, if a food is not correctly identified or its volume is incorrect it may not make much difference with respect to the energy or nutrients consumed. For example, if our system identifies a “brownie” as “chocolate cake” there is not a significant amount of differences in the energy or nutrient content. Similarly, if we incorrectly estimate the amount of lettuce consumed this will also have little impact on the estimate of the energy or nutrients consumed in the meal due to the low energy content of lettuce [8], [13]. Again, we emphasize that our goal is to provide a tool for better assessment of dietary intake to professional dietitians and researchers than what is currently available using existing methods.

A. Image Segmentation

Our system uses various approaches to segment the food items in the image. In particular, we use connected component analysis, active contours, and normalized cuts. Since we are interested in measuring the amount of food in the image, we have developed a very simple protocol for users of our system [8], [13]. This protocol involves the use of a calibrated fiducial marker consisting of a checkerboard (color checkerboard) that is placed in the field of view of the camera. This allows us to do geometric and color correction to the images so that the amount of food present can be estimated.

B. Feature Extraction and Classification

Two types of features are extracted/measured for each segmented food region: color features and texture features. As noted above, as part of the protocol for obtaining food images the subjects are asked to take images with a calibrated fiducial marker consisting of a color checkerboard that is placed in the field of view of the camera. This
allows us to correct for color imbalance in the mobile
device's camera. For color features, the average value of
the pixel intensity (i.e. the gray scale) along with two
chrominance components are used. For texture features, we
use Gabor filters to measure local texture properties in the
frequency domain [25]. In our implementation, we divide
each segmented food item into \( N \times N \) non-overlapped blocks
and use Gabor filters on each block. We use the following
Gabor parameters: 4 scales (S=4), and 6 orientations (K=6).

Once the food items are segmented and their features are
extracted, the next step is to identify the food items using sta-
tistical pattern recognition techniques. For classification of
the food item, we use a support vector machine (SVM) [26].
A classification task usually involves training and testing
data. Each element in the training set contains one class
label and several “attributes” (visual features). The feature
vectors used for our system contain 51 values, 48 texture
features and 3 color features. The labeled food type along
with the segmented image are sent to the automatic portion
estimation module where camera parameter estimation and
model reconstruction are utilized to determine the volume
of food.

C. Volume Estimation

One of the challenging problems of image-based dietary
assessment is the accurate estimation of food portion size
from a single image. As we have indicated above this is
done to minimize the burden on the user. We have developed
a method to automatically estimate portion size of a variety
of foods through volume estimation [15]. These “portion
volumes” utilize camera parameter estimation and model
reconstruction to determine the volume of food items, from
which nutritional content is then determined. Two images
are used as inputs, one is the food image taken by the user,
the other image is the segmented image described in the pre-
vious section. The camera calibration step estimates camera
parameters, comprised of intrinsic parameters (distortion, the
principal point, and focal length) and extrinsic parameters
(camera translation and orientation). We use the fiducial
marker discussed above as a reference for the scale and pose
of the food item identified. The fiducial marker is detected in
the image and the pose is estimated. The system for volume
estimation partitions the space of objects into “geometric
classes,” each with their own set of parameters. Feature
points are extracted from the segmented region image and
unprojected into the 3D space. A 3D volume is reconstructed
by the unprojected points based on the parameters of the
geometric class.

D. Calorie and Nutrient Estimation

Once the volume estimate for a food item is obtained,
it must be converted to a mass for calorie and nutrient
estimation. In order to do so, the densities of the food
items must be either known or have an acceptable prediction
method so that food intake can be appropriately estimated.
Presently, there are more than thousand main foods, such
as granola bar, with no volume information in the USDA
Food and Nutrient Database for Dietary Studies (FNDDS
database) [27]. The FNDDS database contains the most
common food items consumed in the U.S., their nutrient
values, and weights for typical food portions. In addition,
for a number of foods, such as plain yogurt, cake, apple and
potato, a range of densities are reported in the literature.
To address these challenges, we are developing predictive
methods to accurately determine the density of foods using
techniques of computed tomography (CT), magnetic res-
one imaging (MRI) and laser scanning [28], [29]. By
using these techniques, we are adding density information
to the FNDDS. There are three main densities that we are
interested in estimating: true density, apparent density and
bulk density [30]. Techniques for estimating these densi-
ties include: dimension measurement, liquid displacement,
buoyant force determination, solid displacement (Rapeseed
method) and gas pycnometer [31]. We are also developing
effective techniques to predict densities of foods and food
mixtures given the composition and process conditions [30].

IV. SYSTEM ARCHITECTURE

We are developing two different configurations for the
mpFR: a standalone configuration and a client-server con-
figuration. Each approach has potential benefits depending
on the operational scenario.

The Client-Server configuration is shown in Figure 1. In
most applications this will be the default mode of operation.
The process starts with the user sending the image and
meta-data (e.g. date, time, and GPS location information,
when available) to the server over the network (step 1) for
food identification and volume estimation (step 2 and 3), the
results of step 2 and 3 are sent back to the client where the
user can confirm and/or adjust this information if necessary
(step 4). Once the server obtains the user confirmation, food
consumption information is stored in another database at
the server, and is used for finding the nutrient information
using the FNDDS [27] (step 6). Finally, these results can be
sent to dietitians and nutritionists in the research community
for further analysis (step 7). A prototype system has been
deployed on the Apple iPhone as the client and we have
verified its functionality with various combination of foods.
A prototype of the client software has also been deployed
on the Nokia N810 Internet Tablet.

It is important to note that our system has two modes
for user input. In the “automatic mode” the label of the
food item can be confirmed/changed by the user using
the touch screen on the mobile device, after the automatic
analysis. Based on these corrections, the energy intake is
estimated using the FNDDS. The other mode addresses the
problem when no image is available. For some scenarios
it might be impossible for users to take meal images. For
example, the user may not have their mobile telephone with them or may have forgotten to take meal images. To address these situations, we developed an Alternative Method in our system that is based on user interaction and food search using the FNDDS. With the help of experts from the Foods and Nutrition Department at Purdue University, the Alternative Method captures sufficient information to perform food and nutrient analysis, including date and time, food name, measure description, and the amount of intake.

As part of the TADA project, we have designed three unique interconnected databases: an image database (I-TADA) that contains data generated by an image including all the available meta-data, an experiments database (E-TADA) that contains data related to each study and results from the image analysis experiments, and finally an enhanced version of the FNDDS database (T-FNDDS) which includes both nutritional and visual descriptions of each food. These databases provide unique tools in aiding the overall system and can be used by the health/diet/research community for finding, predicting, and extracting useful information for dietary assessment.

V. EXPERIMENTAL RESULTS

Several controlled diet studies were conducted by the Department of Foods and Nutrition at Purdue University whereby participants were asked to take pictures of their food before and after meals [13]. These meal images and their corresponding ground truths were used for our experiments [12].

Initial testing of the mpFR has been done among adolescents. In one of the studies two samples of adolescents used the mpFR during meals and then provided feedback about their experience during interactive sessions [13]. Sample 1 included 63 adolescent boys and girls who participated in one lunch and 55 (87%, 55/63) returned for breakfast the next morning. Adolescents in Sample 2 (n=15) received all foods served familiar to adolescents and each food was matched to a food code in the FNDDS. Further details of this study and results can be found in [13]. The user studies and image analysis experiments are ongoing efforts of this project.

In another experiment, the images taken by the adolescents in Sample 2 described above were considered. These images were from their three different meal events and contained 19 unique food items [12]. By using randomly chosen 50% of the images for training and the rest of them for testing, an average accuracy of 96% (with a variance of 1.8%) was obtained for classifying these 19 food items. Some foods are inherently difficult to classify due to their similarity in the feature space. Examples of such errors are scrambled eggs misclassified as margarine and Catalina dressing misclassified as ketchup. We have also shown from our experiments that the performance of the image segmentation plays a crucial role in achieving correct classification results.

To measure the accuracy of volume estimation for both spherical and prismatic objects, we used 7 food items (5 spherical and 2 prismatic objects) for the experiment. The details of these experiments and results can be found in [15].

The energy intake measured from the known food items for each meal was used to validate the performance of our system. Based on the the number of images used for training, we estimated the mean percentage error of our automatic methods compared to nutrient data collected from the studies. With 10% training data, the automatic method reported within 10% margin of the correct nutrient information. With 25% training data, the automatic method improved to within 3% margin of the correct nutrient information. With 50% training data, the improvement was within 1% margin of the correct nutrient information. Our experimental results indicated that the use of a mobile device using a camera to obtain images of the food consumed is a valid and accurate tool for dietary assessment.

VI. CONCLUSIONS AND DISCUSSION

In this paper we summarized the development of a dietary assessment system using mobile devices. As we indicated, measuring accurate dietary intake is considered to be an open research problem in the nutrition and health fields. We feel we have developed a tool that will be useful for replacing the traditional food record methods currently used. We are continuing to refine and develop the system to increase its accuracy and usability.

REFERENCES


