Development of a Mobile User Interface for Image-based Dietary Assessment

SungYe Kim†, TusaRebecca Schap‡, Marc Bosch‡, Ross Maciejewski†
Edward J. Delp†, David S. Ebert†, Carol J. Boushey‡

† Electrical and Computer Engineering, ‡ Foods and Nutrition
Purdue University, West Lafayette, IN, USA
{inside, tschap, mboschru, rmacieje, ace, ebertd, boushey}@purdue.edu

ABSTRACT
In this paper, we present a mobile user interface for image-based dietary assessment. The mobile user interface provides a front end to a client-server image recognition and portion estimation software. In the client-server configuration, the user interactively records a series of food images using a built-in camera on the mobile device. Images are sent from the mobile device to the server, and the calorie content of the meal is estimated. In this paper, we describe and discuss the design and development of our mobile user interface features. We discuss the design concepts, through initial ideas and implementations. For each concept, we discuss qualitative user feedback from participants using the mobile client application. We then discuss future designs, including work on design considerations for the mobile application to allow the user to interactively correct errors in the automatic processing while reducing the user burden associated with classical pen-and-paper dietary records.

Categories and Subject Descriptors
H.4 [Information Systems Applications]: Miscellaneous; H.5.2 [User Interfaces]: Mobile Dietary System

Keywords
Mobile user interfaces, User interface design, Mobile devices, Health monitoring tools, Dietary assessment system.

1. INTRODUCTION
As smart telephone technology continues to advance in sensor capability and computational efficiency, mobile telephones are being used for purposes other than just dialing and receiving calls. Advances in mobile technology have led to enhanced features, such as built-in cameras, games and texting on phones. Furthermore, researchers have employed mobile technologies in a variety of tools ranging in applications from personal health and wellness to biological research to emergency response [3, 6, 14, 29]. Our research focuses on employing mobile technology for dietary assessment.

Given the mounting concerns of childhood and adult obesity [20, 21], advanced tools for keeping personal dietary records are needed as a means for accurately and effectively monitoring energy and nutrient intakes. Such records provide researchers with information for hypothesis generation and assessment regarding health conditions and nutrient imbalances. In light of these issues, research has focused on a means to utilize mobile technology for dietary assessment (e.g., [10, 15, 22, 27]). Previously, we presented a dietary record assessment system [18] focused on creating a server-client architecture in which the client mobile device was used to capture a pair of images (before and after) of meals and snacks. In our system design, the images are sent to the server where image analysis and volume estimation calculations are performed and the results are returned to nutrition researchers (or dietitians). In this paper, we discuss the design and development of the mobile interface for our client-server system. The development of the mobile user interface of our dietary assessment system has been a collaboration with researchers in the Department of Foods and Nutrition.

We intersperse our discussion with lessons learned based on user feedback, providing guidelines for the development of such applications. Compared to our previous papers, this is the first time our mobile application is described in depth.

Contributions of this paper include:

- **Field specific design considerations and strategies:** We describe design considerations for our mobile user interface for dietary assessment and development strategies based on user evaluation.
- **Improved user interface:** From the point of view of an integrated system, we present details of the data that is created in the mobile application and describe the communication protocol between the mobile client to the server.
- **Flexible eating patterns:** We extend a previous design for recording food images to support flexible eating patterns such as portion refills so that each eating occasion can include not only a pair but also a series of food images.
- **Minimal data exposure and user interaction:** Through filtering eating occasions, we display only some of the images containing information that needs user confirmation. Such a process is used to minimize user interaction, thus reducing user burden.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*MUM* '10, December 1-3, 2010, Limassol, Cyprus
Copyright 2010 ACM 978-1-4503-0424-5/10/12 ...$10.00.
This paper is organized as follows: Section 2 discusses the related research. Section 3 presents our image-based mobile dietary assessment system and the design considerations of a mobile user interface. Section 4 presents the framework and functionality of our mobile dietary assessment tools. Finally, Section 5 presents conclusions and discussion.

2. RELATED WORK

Many systems incorporating mobile devices (e.g., PDAs, cell phones) for data collection and management, have been proposed in various fields [7, 12, 15, 28, 29]. Denning et al. [9] proposed a wellness management system using mobile phones and a mobile sensing platform unit. In this system, users manually input food consumption information with a mobile telephone. The mobile sensing unit computes the user’s energy expenditure and transfers this data to the mobile telephone. Arsand et al. [4, 5] studied how to design easily usable mobile self-monitoring tools for diabetes.

Other dietary assessment systems also make use of mobile phones. Wang et al. [27] describes the evaluation of personal dietary intake using mobile devices with a camera. Users took food images and sent them to dietitians by a mobile telephone. Reddy et al. [24] introduced the DietSense architecture that consists of mobile phones, a data repository, data analysis tools, and data collection protocol management tools. Mobile phones with a GPS receiver were used to capture food images as well as relevant context, such as audio, time and location. The food images were collected into a data repository to be analyzed by dietitians.

Donald et al. [10] discussed the effect of mobile phones in dietary assessment by performing a user study with adolescent patients with diabetes. In their work, users were asked to complete food diaries and to capture images of all food and drink by using mobile phones. Then, food images were downloaded to a computer and a dietitian estimated portion sizes from the diaries and images.

In first developing our mobile user interface, we looked at research which focused on general guidelines and requirements for mobile interface design. Nilsson [19] presented user interface patterns for mobile applications and discussed a series of challenges regarding screen space and interaction mechanisms. Gong and Tarasewich [13] discussed design requirements for mobile user interfaces compared to those for the desktop environment based on Shneiderman’s Golden Rules of Interface Design [25] and proposed guidelines for mobile devices such as consistency for a desktop counterpart and design for top-down interaction. Chong et al. [8] focused on designing a touch screen based user interface. In this work, they proposed several key guidelines influencing a good user interface including; simplicity, aesthetics, productivity, and customizability.

We also examined research [4, 5, 16] focusing specifically on design concepts for self-health monitoring tools utilizing the capability of mobile devices. Designing field specific interfaces requires careful consideration and even field background knowledge [17]. Based on our initial system prototype and user feedback, we adopted design concepts proposed by Arsand et al. [4]. Our system design was done in conjunction with interactive feedback from nutrition researchers. Our system was developed to provide researchers with an automated analysis for the estimation of daily food intake. The dietary assessment tool uses a built-in camera on a mobile telephone to capture food images, and the user interactions of taking meal images and verifying the assessment demands careful design considerations in terms of usability and flexibility in order to reflect various situations in daily life. Specifically, we incorporate an automatic data transfer scheme (hiding the data transfer paradigm from the users) and an easy data entry scheme through the use of the mobile telephone camera.

3. SYSTEM DESIGN

The focus of our dietary assessment system is to better capture meal patterns of users using automated food and portion analysis. Our dietary assessment system utilizes a server-client configuration.

3.1 System Overview

In the record method, food images are captured using a built-in camera on a mobile device and the corresponding management data for food images is generated. The captured food images are transferred to a server for analysis where the various types of food in the images are identified and the portion sizes (volumes) of food are estimated to compute the amount of food consumed by the user. This analysis part is broken into two sub parts; food identification [31] and volume estimation [30]. These results (food type and volume) are returned to the user for review. In the review method, the user confirms or adjusts the results produced in the server. Figure 1 shows the overview of our image-based dietary assessment system. Additionally, an alternate method is considered for cases where the user cannot capture food images.

The record and review methods and the alternate method are deployed on client mobile phones. The analysis components run on the server to minimize the burden of computation on the mobile telephone. This paper focuses on the client side of the system and we present the design and development of the mobile user interfaces for an image-based dietary assessment system. For details of the analysis components, see Zhu et al. [31] and Woo et al. [30].

3.2 Design Considerations for Mobile Dietary Assessment Tools

Figure 1: Overview of an image-based dietary assessment system in a server-client architecture.
Figure 2: (Left) the framework of our mobile dietary assessment tool and (right) an eating occasion structure consisting of a series of food images, eating occasion data and food tags. The framework is broken into two tracks; one uses eating occasions with food images and the other deals with eating occasions without food images. Each food tag includes food labels and volumes. Note the dashed structure in the eating occasion diagram. This was part of our initial concept framework and was later removed based on user feedback. Please see Section 4.1 for a detailed discussion of the design concept and lessons learned.

Based on design discussions with our nutrition research partners and analysis of various eating patterns in real life, we derived several design considerations for our mobile dietary assessment tools.

1) **Easy and fast collection of daily food images**: Capturing food images by using a mobile telephone with a built-in camera provides users with easier, faster and potentially more accurate ways to record meals in comparison to methods of remembering or noting what they ate [1]. Thus, we chose to utilize the built-in cameras on mobile phones as a data collection tool.

2) **Minimum user interaction**: Since users of a dietary assessment system are as various as users of mobile phones, we need to consider a wide range of users from adolescents to senior citizens. Hence, the user interface should be easy to use and the interaction needed to record or review eating occasions should be minimal.

3) **Flexible eating patterns**: The system should be flexible enough so that various eating patterns in real life, such as multiple eating occasions per day and the addition of foods during an eating occasion, can be collected and managed in the system.

4) **Protection of personal data**: Users’ food images and additional private data should be kept private. Hence, food images on the mobile telephone should be hidden from other people.

5) **Automatic data processing**: A server-client architecture reduces the burden of analysis and computation on the mobile telephone and increases the overall system performance. While employing this, the data transfer between a mobile telephone and a server should be hidden from the users. Moreover, users do not need to know details of server processing.

6) **Exceptional situations**: Additional methods should provide users with tools to manually save or modify eating occasions. For instance, in cases that users cannot capture food images due to situational or environmental conditions, users should be able to create their eating occasion data without food images. Moreover, if the results from the analysis components contain errors, users should also be able to correct them and update the server.
4. MOBILE USER INTERFACE FOR IMAGE-BASED DIETARY ASSESSMENT SYSTEM

Based on the design considerations in Section 3.2, we describe the framework of our mobile dietary assessment tool. In this section, we describe our system tools and implementations. Furthermore, we discuss lessons learned and present qualitative user feedback on each of the deployed components. User studies have included 117 adolescents (10∼18 years) and 57 adults (21∼65 years).

Figure 2 shows the framework for our mobile interface and the basic structure of an eating occasion. As shown in Figure 2 (left), our mobile tool includes three main methods: the record, review and alternate method. The record and review methods share eating occasion data with food images, whereas the alternate method supports users in creating eating occasions even when they cannot capture food images. Each eating occasion may contain several food images, eating occasion data created by the record method, and food tags including food labels and volumes as shown in Figure 2 (right).

Due to programming constraints (i.e., no multi-threading) with iOS 3.x and earlier, we separated the record and review methods into two different applications. Eating occasion data is created in the record method in one application and transferred to the server. The review method in the other application receives the eating occasion data from the server after completing the analysis on the server. User feedback indicated that such a separation was sub-optimal. Fortunately, the new iOS 4.x allows an application to remain in the background by providing multi-threading so that data communication between a mobile phone and a server will be available without such separation. Figure 3 shows a diagram for data communication between the server and the separated two record and review applications and an example database on the server. In the following sections, we present the detailed functionality, development, lessons learned and user feedback of each method for our mobile dietary assessment.

4.1 Recording the Eating Occasion

By utilizing the record method, the user captures food images with the built-in camera on the mobile device. While recording food images, the eating occasion data for the food images (*.rec) (see Figure 2 (right)) is generated for each eating occasion containing the device ID, user ID, date/time and file names of all food images. The date and time represent when the first image was captured. All food images and eating occasion data are transferred to the server (step 1 in Figure 3). The guide lines denoting the top and bottom of the image, superimposed on the camera view were key to helping users fit foods into the screen. More importantly, the top and bottom notation insured correct orientation of a mobile phone. Prior to implementing this feature, images were captured in a variety of orientations, confounding image segmentation and volume estimation.

**Handling various eating patterns:** To record and manage multiple helpings (portion refills), we allowed users to capture a series of images for a given eating occasion. Each image in the sequence was a before or after image except for the first image (which can only be a before image) and the last image (which can only be an after image). All these images were then saved in the order of their capture time to reflect the time progression of the eating occasion. Figure 4 shows the food image recording interface that was composed of three buttons. Capturing the first image creates a new eating occasion and capturing the last image sends the current eating occasion to the server for analysis.

To implement this concept, we included three buttons: First, In-Between and Last. Adolescent users found the In-Between button to be confusing. Many users interpreted it to mean they should take an image mid-way through their meals. Therefore, the button was disabled and a series of First/Last images were used to accommodate portion refills and multiple courses. This sequencing was well received.

**Recommendations for training and recording:** The user interface development and available technology can provide many solutions to aid in the automated analysis. However, some issues must be addressed through user training. Suggestions to the users for improved training and recording include:

1) **Full image acquisition:** Foods should be captured large enough and around the center area in the images so that they can easily be recognized during analysis.

2) **Highlight and shadow condition:** Highlights and shadows hinder the automatic analysis by making the recognition of food areas difficult.

3) **Spatial consistency:** When several foods exist in the images, keeping them at the same position helps the analysis part to automatically recognize foods in different images of the same eating occasion.

4) **Camera setup:** Significant perspective distortion can occur in the image according to the distance and the angle of view. We experimentally found angles between 30∼45° to be satisfactory.

4.2 Reviewing the Eating Occasion

After recording eating occasions, the users review the analyzed results from the server to either confirm or change food types (food names) and volumes (the amount of intake for each food). As shown in Figure 1, the analysis is automatically decided by the food identification and saved into food label data (*.tag), and the food volume is algorithmically estimated in the volume estimation and saved into food volume data (*.vol) to be returned to the users. The purpose of this review method is to provide user interfaces on mobile phones for reviewing the data transferred from the server. When the user starts the review method, an application requests the eating occasion data of food images generated in the record method and analysis results to the server (steps 4-1 and 4-2 in Figure 3).

**Filtering eating occasions:** First, a list of eating occasions are displayed so that the users can select one eating occasion to review as shown in Figure 5. At this moment, we delay loading the food images and related context information (i.e., food tag data) on the memory of the mobile device until the user selects one to review among the eating occasions. This lazy loading [11] strategy is applied to all the other methods to reduce the memory consumption on the mobile device. Conceptually, users view images they recorded so that they can review them and make adjustments or corrections as necessary to the foods’ type or volume.
The dot indicators were proposed to remind users of which image has been taken. In this example, four red dots indicate that the first image and three in between images have been taken. The right most grey dot among the dots shows that the last image has yet to be taken.

If the list of eating occasions displayed includes all eating occasions recorded in the record method, it could be extremely long after continuous usage. Although performance issues of handling long lists could be averted with efficient coding of the application, the concern is that the users would have to scroll through the long list of eating occasions in order to look for a specific one. Hence, we filter eating occasions to display on the review method using three strategies:

1) When all the food tags in an image have been confirmed by the user, the corresponding image should be hidden from the eating occasion.

2) When all the images in the eating occasion have been confirmed by the user, the corresponding eating occasion is hidden from the user. (The eating occasion is not displayed on the list anymore.)

3) Eating occasions whose images are yet to be processed by the server are not displayed on the list.

**Displaying food tags:** Once the user selects an eating occasion, food images with food tags (food labels) are displayed in order. Figure 6 (left) shows the image with food labels. Each label is transparent to reduce occluding food areas and is matched to the food where the pin icon with the same color is placed. In the case that there are several food labels, the positions of the labels are rearranged to minimize cluttering.

**User confirmation and adjustment:** In the case that the analysis results from the server contain errors, we engage the confirmation and adjustment of the food tags by the users (step 5 in Figure 3). The confirmation reports to the server that all food tags in the image are correct and updates the server when food tags are changed (step 5-1 in Figure 3), and the adjustment allows users to locally make changes to food labels and volumes (step 5-2 in Figure 3).

To correct wrong food labels in the food tags, we employ the stemming algorithm [23] for keyword search, simplified for the Food and Nutrient Database for Dietary Studies (FNDDS) 3.0 [2]. However, utilizing general keyword search for food data often produces undesired results. For example, keyword search for ‘apple’ returns a list of 158 food names from the FNDDS 3.0 and the raw apple is buried at least 50 foods deep. Hence, another potential way to search foods is to use a set of predefined keywords. Since defining search keywords for foods requires domain specific knowledge and user evaluation, our nutrition colleagues have been helped to define the keywords. In future versions of our user interface, new search method will be incorporated using these predefined keywords. Further, feedback from adults and adolescents also indicated that a search mechanism needs to accommodate misspellings and synonym names for foods. Integration of these concepts will improve the efficiency of the tool used to search for foods.

**4.3 Alternate Method**

The main concept of our dietary assessment system is to utilize food images recorded by the users during eating occasions. However, it may, at times, be impossible for users to take food images. For instance, the user may not have had his or her mobile telephone or may have forgotten to take food images. To support such situations, we also provide an alternate method that is based on user interaction and a manual food search.

This alternate method is developed as an independent track in the framework of the client side as shown in Figure 2 (left) and provides users with tools for creating eating occasions containing enough information for nutrient analysis, including date and time, food name, measure description and the amount of intake as shown in Figure 7. In this method, food names are searched by using the same method employed for the user confirmation in the review method.
The portion units for the amount consumed depend on the food that the user selects and come from FNDDS 3.0. From user feedback, users have expressed that the portion units need to be polished (i.e., the number ‘1’ in front of each portion unit is confusing (Figure 7 (right))). This is an issue linked to modification of the FNDDS rather than user interface design. Our nutrition colleagues have also been working to update the FNDDS. Thus, in our future versions of the alternate method, this issue will be improved.

5. CONCLUSION AND DISCUSSION

We have described the design and development of a mobile user interface for an interactive, image-based dietary assessment system using mobile devices. As a self-monitoring system for dietary assessment, our system uses food images captured using a built-in camera on mobile devices and analyzes the images on the server in order to automatically label foods and estimate the amount of food consumed. For flexibility, our system also considers an alternate method based on user interaction and food search instead of food images. Hence, our mobile client system using mobile phones is composed of the record, review and alternate methods.

The current user interface design considerations are used to handle capturing meal images and reporting mistakes during the automatic food identification. We have begun work on implementing user controls for also adjusting errors in the volume estimation. As an initial idea, we are currently proposing a system that will provide 3D primitive template shapes and interactions for translating, rotating and scaling the overlayed volumes. Figure 6 (right) shows the image with cylindrical food volume. A transparent 3D shape is superimposed on the food area to provide users with tools to either confirm or adjust the food volumes estimated on the server. User interfaces for translating and scaling the 3D shape are displayed when the user selects a specific food tag in order to minimize cluttering the screen.

While designing and developing our mobile user interface, we collaborated with researchers in the Department of Foods and Nutrition to collect feedback on the interaction design through several user studies [26]. Based on the feedback, we tailored the user interface of our client mobile system for improving usability as well as system flexibility. To demonstrate, we developed and deployed our client mobile tools on the iPhone 3GS devices running iOS 4.0.

6. ACKNOWLEDGMENTS

We would like to thank our collaborators and colleagues, Bethany Six of the Department of Foods and Nutrition and Deb Kerr of the School of Public Health, Curtin Institute of Technology for feedback about mobile user interfaces through user studies, Insoo Woo, JungHoon Chae, Fengqing Zhu of the School of Electrical and Computer Engineering for their work in developing food volume estimation and food image analysis on the server and Ahmad M. Razip of the School of Electrical and Computer Engineering for his help in revising the user interface. Support for this work comes from the National Cancer Institute (1U01CA130784-01) and the National Institute of Diabetes, Digestive, and Kidney Disorders (1R01-DK073711-01A1). More information about our project can be found at www.tadaproject.org.

7. REFERENCES

[1] [Diet History Questionnaire: Web-based DHQ, National Cancer Institute]
[2] [The Food and Nutrient Database for Dietary Studies]


