A Mobile Phone User Interface for Image-Based Dietary Assessment

Ziad Ahmad[a], **Nitin Khanna**[b], **Deborah A. Kerr**[c], **Carol J. Boushey**[d], and **Edward J. Delp**[a]

[a]School Of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana

[b]Department of Electronics and Communication Engineering, Graphic Era University, Dehradun, India

[c]School of Public Health, Curtin University of Technology, Perth, Australia

[d]University of Hawaii Cancer Center, Honolulu, Hawaii

Abstract

Many chronic diseases, including obesity and cancer, are related to diet. Such diseases may be prevented and/or successfully treated by accurately monitoring and assessing food and beverage intakes. Existing dietary assessment methods such as the 24-hour dietary recall and the food frequency questionnaire, are burdensome and not generally accurate. In this paper, we present a user interface for a mobile telephone food record that relies on taking images, using the built-in camera, as the primary method of recording. We describe the design and implementation of this user interface while stressing the solutions we devised to meet the requirements imposed by the image analysis process, yet keeping the user interface easy to use.

Keywords

User interface; dietary assessment; mobile telephone food record

1. INTRODUCTION

Many chronic diseases, including obesity and cancer, are related to diet. Such diseases may be prevented and/or successfully treated by accurately monitoring and assessing food and beverage intakes. Existing dietary assessment methods, such as the 24-hour dietary recall, in which a dietitian interviews an individual about foods and beverages consumed in the last 24 hours; and the dietary record, which relies on the individual to record on paper the details of foods and beverages consumed at each eating occasion, tend to be burdensome and prone to error due to omissions and memory [1–5].

With the burden and inaccuracy associated with existing dietary assessment methods and with the enormous advances in technology, the need for new methods that rely on technology seem inevitable. Studies show that dietary records running on a mobile telephone promises to be a valuable tool for dietitians and researchers in dietary assessment [6–8].

Address all correspondence to Edward J. Delp, ace@ecn.purdue.edu or see www.tadaproject.org.

Today, mobile telephones are bundled with many features. These features include a built-in camera, network access, location information retrieval, an accelerometer, and many more. Such advancements in the mobile telephone technology allow for enhancing dietary assessment methods to reduce the major barriers to accuracy.

In this paper, we focus on describing a user interface we developed for mobile telephones for dietary assessment using images, which are captured via the built-in camera, as the primary method of recording. Such a user interface can be used as the front end for image-based dietary assessment systems such as those described in [9–15].

2. RELATED WORK

Several online and mobile telephone applications exist today that can be used to manually record food intake. Examples of such applications are Restaurant Nutrition [16] and Tap & Track [17]. These tools provide two main advantages over existing paper-based methods, they make records easier to keep and perform calculations automatically. However, such methods clearly do not make the recording process easier for the user nor do they help increase the accuracy of the food intake estimation. On the other hand, the methods presented in [9, 11, 14] rely on taking images of eating occasions as the primary method of recording, which helps address the problems mentioned above.

An online system known as FoodLog [11] relies on the users to take images of their eating occasions using a camera and then send the images to the system by email. FoodLog uses image analysis methods for food identification from which the nutrients can then be calculated. Although FoodLog shares many similarities with other systems [9, 14], it is not a dedicated mobile application but a web application. A dedicated mobile application would offer several features that cannot be available in a web-based system. For example, with a dedicated mobile application the user can be reminded to take images of eating occasions and can be better assisted to take images in a specific way that would help the automated analysis process achieve more accurate results. One of the important contributions of our work is in designing a user interface that is easy to use yet satisfies the requirements imposed by the automated analysis process. To our knowledge such issues have not been addressed in any previous work. Before closing this section, we would like to emphasize that our application is the result of several user studies involving more than 170 individuals [18–20]. Our system has images from these studies and the training images for these studies which represents about 11,000 images.

3. SYSTEM OVERVIEW

The mobile user interface described in this paper was developed as part of our complete image-based dietary assessment system presented in [9]. In this section, we briefly describe our dietary assessment system, presented in [9], which employs a client-server configuration. The architecture of this dietary assessment system is shown in Figure 1 and is generically similar to the architecture used in [14]. A user begins the process by capturing images of foods using the built-in camera in the mobile telephone, and then the images, along with information about the images, are sent to the server (step 1) for food and

beverage analysis (steps 2 and 3). The analysis part is done on the server due to its computational resource requirements. The analysis process consists of two main parts: food identification and volume estimation [9]. The analysis results are then sent back to the mobile telephone for the user to review and confirm (step 4). After the server obtains the user confirmation, dietary intakes are stored in a specific database on the server that is used for finding the nutrient information using the Food and Nutrient Database for Dietary Studies (FNDDS) database [21] (step 6). The FNDDS database contains the most common foods consumed in the U.S. and their nutrient values. Finally, the results can be made available to the research community or to healthcare professionals (step 7). In this paper we focus on the design and implementation of the client side for such a system. We have implemented the user interface on the Apple iPhone device.

4. MOBILE USER INTERFACE

In this section, we discuss in detail the design and implementation of the mobile user interface for an image-based dietary assessment system. Although the mobile telephone user interface and the server system seem to be separate, the mobile user interface cannot be designed independently of the server. Our strategy was to develop a user interface that is easy to use and intuitive given the requirements imposed by the food analysis system.

Our current application consists of three main views, which the user can access using a tab bar. The views are titled *Record, Review*, and *More* and appear on the tab bar from left to right in this respective order, as shown in Figure 2. The *Record* view is used for capturing images of eating occasions, while the *Review* view is used for reviewing the results of the image analysis received from the server, and finally the *More* page provides general features of the application.

4.1 Recording Eating Occasions

In our system, the user records eating occasions by capturing images of foods and beverages. In Figure 2, the recording view, titled *Record*, is shown on the left, which consists of two primary buttons, one for taking an image before eating and another for taking an image after eating. For fast and efficient use, when the user launches the application, the first view is the *Record* view. The buttons are purposely large to aid the user in quickly identifying and tapping on the right buttons. When at other views, the user can go to the *Record* view by choosing the button titled *Record* from the tab bar on the bottom of the application screen. Through feedback from users in user studies, we found that limiting the record options to a simple, "before eating" and "after eating" best accommodated an eating occasion, as well as second helpings. In the latter case, taking an "after eating" image prior to adding more food and then taking another "before eating" image after adding more food was found to be a more intuitive solution from the user perspective [10].

4.1.1 Image Acquisition—A very crucial step in the recording process is the image acquisition, which has a great impact on the performance of any image-based dietary assessment system. One role of the mobile interface is to assist the user in capturing useful images. The image is considered to be "useful" if it meets certain requirements that allow the back-end processing on the server to analyze and extract food information from the

image. Particularly, the image is required to: (1) have food placed on a flat surface with all foods appearing in the view, (2) have no highlights or shadows in the view, (3) have the fiducial marker (checkerboard)* in the view and placed to the lower left of the food, (4) be captured in landscape view with a specific orientation, and (5) have a view angle between 45 and 60 degrees from the horizontal.

We developed specific solutions in our mobile interface to aid the user in meeting these image requirements and particularly requirements 3 to 5. In order to ensure that the user is aware of these requirements and to address the remaining requirements, we implemented tips in the form of an alert message that pops up before the user takes the image (Step 1 in Figure 3). These tips can be turned off in the user settings. For requirement 3, we utilized the method presented in [22]. This fast method for searching for the fiducial marker in an image executes on the mobile telephone and provides immediate feedback. If the fiducial marker is not found, the user is asked to either retake the image or save it anyway when a fiducial maker is not available (Figure 4). In addition to checking for the presence of the fiducial marker, we implemented a fast algorithm for blur detection that was also developed in [22]. This is also performed as part of the image quality checking step of Figure 4. If the image is blurry the user is asked to retake the image, similar to the case of a missing fiducial marker. For requirement 4, we display two labels with titles *Top* and *Bottom* in the camera view to indicate landscape mode and to show where the top and the bottom of the image should go (Step 2 in Figure 3). For requirement 5, we use the accelerometer of the mobile telephone to extract the angle information of the telephone, which we use to at least ensure that the user is not taking a top view or a front view of the foods and beverages. This is important for the volume estimation analysis performed on the server. To guide the user to the correct position (i.e. landscape view with correct angle) of the mobile telephone, we display guidelines on the top and bottom of the camera view that go under the *Top* and *Bottom* labels. These lines and the *Top* and *Bottom* labels are in red when the camera is not in the correct position and turns green when it is in the proper position. We added a camera assistant in the form of small-font words displayed in the middle of the camera view in order to aid the user in reaching a proper position. The camera assistant can be turned off in the user settings in order to have a clear view in the camera whenever the user feels confident and trained enough to take a proper image.

4.1.2 Saving Eating Occasion Information—For each eating occasion, we compress and save the images on the mobile telephone and, in addition to that, we write the information of each eating occasion into a specific file with a *.rec* extension (we call the *rec* file). This information includes the device ID, user ID, date and time of the eating occasion, along with the file names of the before and after eating images. The date and time of the eating occasion correspond to the date and time of the before eating image. If the user allows the application to use the location services of the mobile telephone, we save the GPS location information, when available, of each image into a specific file with *.gps* extension (we call the *gps* file).

---

*A fiducial marker is an object that is required to appear in the acquired images to aid in the automated analysis process performed on the server. It aids in geometric and color correction of the images [22].

4.1.3 Communication—Once the "after eating" image is captured and saved, the data for the specific eating occasion, including the images, the *rec* file, and the *gps* file are sent to the server. In order to avoid losing any eating occasion, we keep track of the status of the communication of each eating occasion by adding the corresponding *rec* file name to a list and use a flag to indicate whether the eating occasion data was successfully received by the server. If a network connection is available and all the required files of the eating occasion are received by the server, then an alert message is displayed on the screen notifying the user of successful communications (Figure 5 left) and the flag of eating occasions is set to indicate the successful communication of the eating occasion. If a network connection is unavailable, the user will be notified in an alert message and then a button on the top-right of the screen is displayed to indicate the number of unsent eating occasions (Figure 5). This button can be used for sending the unsent eating occasions to the server again. The button is hidden when all eating occasions have been received by the server.

4.2 Reviewing Eating Occasions

In our application, we implement a dedicated view for reviewing eating occasions. The user can always navigate to this view, titled *Review*, by choosing the button titled *Review* from the tab bar on the bottom of the application screen. The *Review* view is shown in Figure 2 in the middle. This view consists of a list of eating occasions that are ready to be reviewed and a button on the top-right of the screen that is used to check and download any new results available on the server. When the user taps this button, which we call the *Refresh* button, the application checks for new results on the server. If no new results are available on the server, the user is then notified about this, otherwise, the results are downloaded in a file with extension *.tag* (we call this file a *tag* file) and the eating occasion is then added to the list indicating that it is ready to be reviewed. We list the eating occasions by showing a thumbnail of the before eating image and the date and time of the eating occasion. The *tag* file has the same name as its corresponding before eating image and includes the suggested foods, with their names and FNDDS database codes, and the pixel location for each identified food in the image.

To begin reviewing an eating occasion, the user can choose the eating occasion from the list in the *Review* view. The before eating image is then displayed in landscape view with the food labels placed on the image according to the results received from the server (Figure 6(1)). The labels are in the form of pins pointing to the food and a bubble that includes an abbreviated food name of 15 characters or less, due to the small screen space constraints. We use different colors to differentiate labels, to help the user in understanding bubble-pin correspondence. However, we reserve the green color for confirmed labels. The user can zoom in to have a better view of the location of the pins (Figure 6(2)), which is important when labels happen to be overlapping.

To edit the labeling of the food, there are four main operations: change the food shown in the bubble, remove the label, confirm the label, or add a label. To do any of the first three operations the user has to tap on the bubble of the label and then a list consisting of three different sections with a search bar on top is displayed as shown in Figure 6(3). The sections have the titles, and also appear in the respective order, *Confirm/Remove Pin, Suggested*

Food, and *Complete Food List*. Under the *Confirm/Remove Pin*, the options to confirm and remove pin are listed. When the label is confirmed, the colors of both the bubble and its corresponding pin are turned to green. In the *Suggested Food* section, four suggested foods are listed in order of likelihood as received from the server. The third section, which is the *Complete Food List*, lists all foods available in the database on the telephone. Within the *Suggested Food* and the *Complete Food List* sections, each food occupies two lines of text. The first line corresponds to the abbreviated food name visible on the screen. The second line of text contains more detailed information about the food, for example, in the case of "cake", the detail indicates "cake, cupcake, chocolate, w/ icing or filling" (Figure 6(3)). In Figure 6(4), the abbreviated "diet coke" is listed twice in the search result. However, the second line reveals two choices to be distinct, i.e., "soft drink, cola-type, decaffeinated, sugar-free" or "soft drink, cola-type, sugar-free." Thus, a user could specify whether the cola in the image was decaffeinated or not, which is a feature that cannot be determined from image analysis. We have as well implemented an incremental searching mechanism that would help the user in quickly searching the list. The user can search by using the search bar that is located above the list, as show in Figure 6(4). The user can type-in a new food name if it is not found in the list. We reserve the last item in the list of the *Complete Food List* section and the last item in the list of results of searching (Figure 6(4)) to be used to go to the screen where the user can type-in a new food. This screen, which we call food not listed, is shown in Figure 6(5).

For the fourth operation, which is adding a label, the user has to choose the *Add* button from the tool bar to start adding a label. Then, the toolbar displays two buttons; one with a pencil image and the other one to cancel adding a pin. After zooming to the food that is to be labeled, the user can choose the pencil button to start drawing, using the touch screen, a contour around the food. The information about the contour represented in points is saved in a file with extension *.pts* (we will call the *pts* file). After drawing the contour, the user can either choose to clear the contour or use it. If the user chooses to use the contour, then a list, consisting of the *Complete Food List* section only, is shown. After choosing a food, a pin, in green color indicating that the pin is confirmed, is added to the image at the appropriate location with the name of the chosen food appearing in the bubble. The whole process of adding new labels is shown in Figure 7. We consider this method of adding labels by drawing contours around the food to be a significant contribution to the implementation of user interfaces of such applications, because it allows the user to add a food label in a simple way, and more importantly, allows for description retrieval using the contour of the added food, which is made accurate by the zooming capability. The information about the contour of the food is sent to the server to help the automatic process performed on the server to learn and hence reach better results in the future.

After the user completes the labeling of the image, i.e. all the labels are in green color, the user can choose the *Confirm* button placed in the middle of the toolbar (Figure 6(1)). Then, the updated *tag* file is sent to the server along with a single *pts* file that is created by appropriately merging all the *pts* files of the user-added labels. If a network connection is available and the communication with the server was successful, the user is then notified using an alert message with an *OK* button. After the user taps on the *OK* button, the application goes back to the review screen with the confirmed eating occasion removed from

the list. On the other hand, if a network connection is unavailable and hence communication of the files to the server is unsuccessful, the user is informed about that, and whenever the user chooses to review this specific eating occasion in the future, he/she will be asked to try and resend the data to the server by choosing the *Confirm* button again.

5. CONCLUSIONS

In this paper, we presented a mobile telephone user interface for an image-based dietary assessment system. We described the design and implementation of the user interface with focus on the solutions we devised to meet the requirements imposed by the image analysis system. Our strategy was to develop a user interface that is easy and intuitive given these requirements. We developed the mobile application for the Apple iPhone and tested it in various user studies (presented in [18–20]) involving more than 170 users that acquired just over half of the 11,000 images in our system's database.

Acknowledgments

References

1. Klesges R, Eck L, Ray J. Who underreports dietary intake in a dietary recall? Evidence from the Second National Health and Nutrition Examination Survey. Journal of Consulting and Clinical Psychology. 1995 Jun.63:438–444. [PubMed: 7608356]

2. Johnson R, Soultanakis R, Matthews D. Literacy and body fatness are associated with underreporting of energy intake in US low-income women using the multiple-pass 24-hour recall: a doubly labeled water study. Journal of the American Dietetic Association. 1998 Oct.98:1136–1140. [PubMed: 9787719]

3. Tooze J, Subar A, Thompson F, Troiano R, Schatzkin A, Kipnis V. Psychosocial predictors of energy underreporting in a large doubly labeled water study. American Journal of Clinical Nutrition. 2004 May.79:795–804. [PubMed: 15113717]

4. Bathalon G, Tucker K, Hays N, Vinken A, Greenberg A, McCrory M, Roberts S. Psychological measures of eating behavior and the accuracy of 3 common dietary assessment methods in healthy postmenopausal women. American Journal of Clinical Nutrition. 2000 Mar.71:739–745. [PubMed: 10702167]

5. Sawaya A, Tucker K, Tsay R, Willett W, Saltzman E, Dallal G, Roberts S. Evaluation of four methods for determining energy intake in young and older women: Comparison with doubly labeled water measurements of total energy expenditure. American Journal of Clinical Nutrition. 1996 Apr. 63:491–499. [PubMed: 8599311]

6. Six B, Schap T, Zhu F, Mariappan A, Bosch M, Delp EJ, Ebert DS, Kerr D, Boushey CJ. Evidence-based development of a mobile telephone food record. Journal of American Dietetic Association. 2010 Jan.110:74–79.

7. Boushey CJ, Kerr D, Wright J, Lutes K, Ebert DS, Delp EJ. Use of technology in children's dietary assessment. European Journal of Clinical Nutrition. 2009; 63:S50–S57. [PubMed: 19190645]

8. Kerr D, Pollard C, Howat P, Delp E, Pickering M, Kerr K, Dhaliwal S, Pratt I, Wright J, Boushey C. Connecting Health and Technology (CHAT): protocol of a randomized controlled trial to improve nutrition behaviours using mobile devices and tailored text messaging in young adults. BMC Public Health. 2012 Jun.12:477. [PubMed: 22726532]

9. Zhu F, Bosch M, Woo I, Kim S, Boushey CJ, Ebert DS, Delp EJ. The use of mobile devices in aiding dietary assessment and evaluation. IEEE Journal of Selected Topics in Signal Processing. 2010 Aug; 4(4):756–766. [PubMed: 20862266]

10. Kim, S., Schap, T., Bosch, M., Maciejewski, R., Delp, EJ., Ebert, DS., Boushey, CJ. Development of a mobile user interface for image-based dietary assessment. Proceedings of the International Conference on Mobile and Ubiquitous Multimedia; Limassol, Cyprus. 2010 Dec. p. 13:1-13:7.

11. Kitamura, K., Yamasaki, T., Aizawa, K. Foodlog: capture, analysis and retrieval of personal food images via web. Proceedings of the ACM Multimedia Workshop on Multimedia for Cooking and Eating Activities; Beijing, China. 2009 Oct. p. 23-30.

12. Arab L, Estrin D, Kim D, Burke J, Goldman J. Feasibility testing of an automated image-capture method to aid dietary recall. European Journal of Clinical Nutrition. 2011 May; 65(10):1156–1162. [PubMed: 21587282]

13. Sun M, Fernstrom JD, Jia W, Hackworth SA, Yao N, Li Y, Li C, Fernstrom MH, Sclabassi RJ. A wearable electronic system for objective dietary assessment. Journal American Dietetic Association. 2010 Jan.110:45–47.

14. Puri, M., Zhu, Z., Yu, Q., Divakaran, A., Sawhney, H. Recognition and volume estimation of food intake using a mobile device. Proceedings of the IEEE Workshop on Applications of Computer Vision; Snowbird, UT, U.S.A. 2009 Dec. p. 1-8.

15. Shang, J., Sundara-Rajan, K., Lindsey, L., Mamishev, A., Johnson, E., Teredesai, A., Kristal, A. A pervasive dietary data recording system. Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops; Seattle, WA. 2011 Mar. p. 307-309.

16. Restaurant Nutrition. https://itunes.apple.com/us/app/restaurant-nutrition/id285180322. Foundation HealthCare Network Products;

17. Tap & track. http://nanobitsoftware.com. Nanobit Software Products;

18. Six B, Schap T, Kerr D, Boushey C. Evaluation of the food and nutrient database for dietary studies for use with a mobile telephone food record. Journal of Food Composition and Analysis. 2011 Dec; 24(8):1160–1167. [PubMed: 22389554]

19. Daugherty BL, Schap TE, Ettienne-Gittens R, Zhu F, 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 Apr.14(2):e58. [PubMed: 22504018]

20. Six B, Schap T, Kerr D, Boushey C. Importance of user interaction with automated dietary assessment methods. European Journal of Clinical Nutrition. 2012 May.66(5):648. [PubMed: 22333872]

21. United States Department of Agriculture; Agricultural Research Service Nutrient Data Laboratory Home Page. USDA Food and Nutrient Database for Dietary Studies. http://www.ars.usda.gov/Services/docs.htm?docid=12089

22. Xu, C., Khanna, N., Boushey, CJ., Delp, EJ. Low complexity image quality measures for dietary assessment using mobile devices. Proceedings of the IEEE International Symposium on Multimedia; Dana Point, CA. 2011 Dec. p. 351-356.
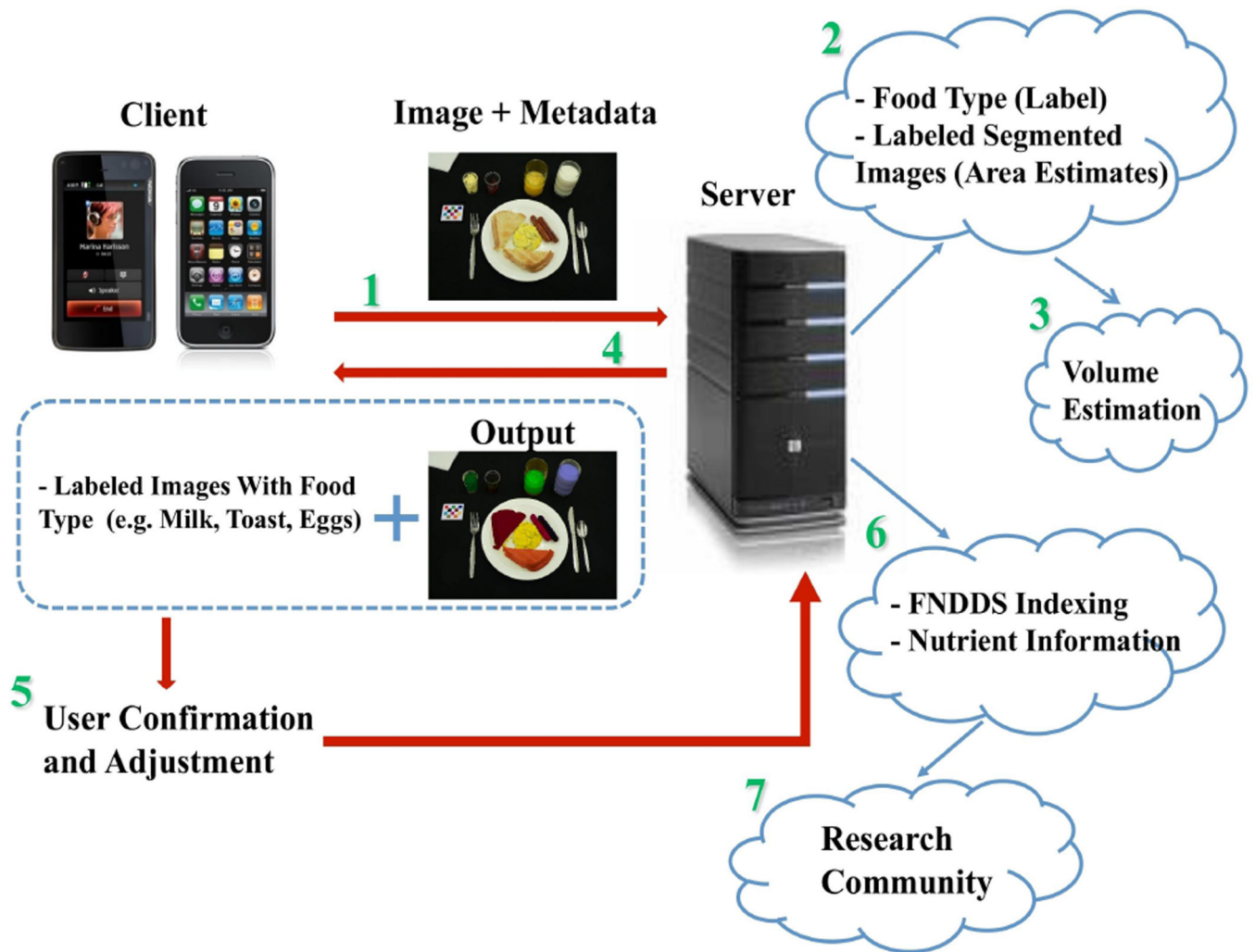
Figure 1.
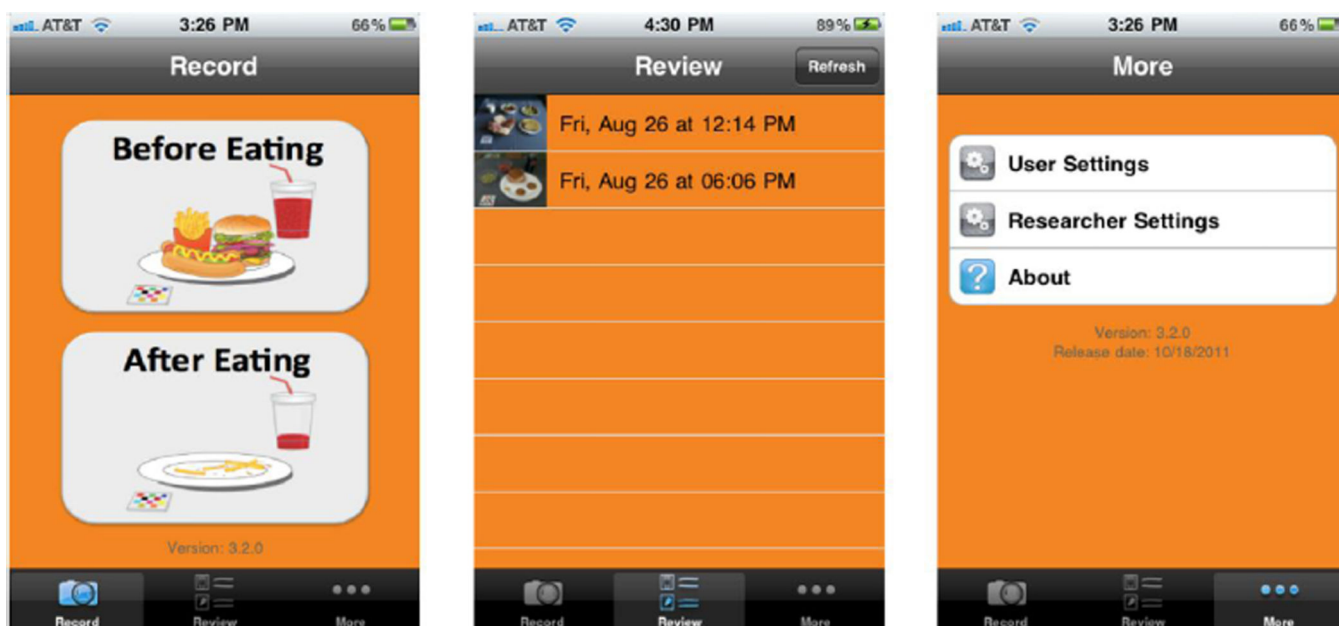Architecture of an image-based dietary assessment system [9].
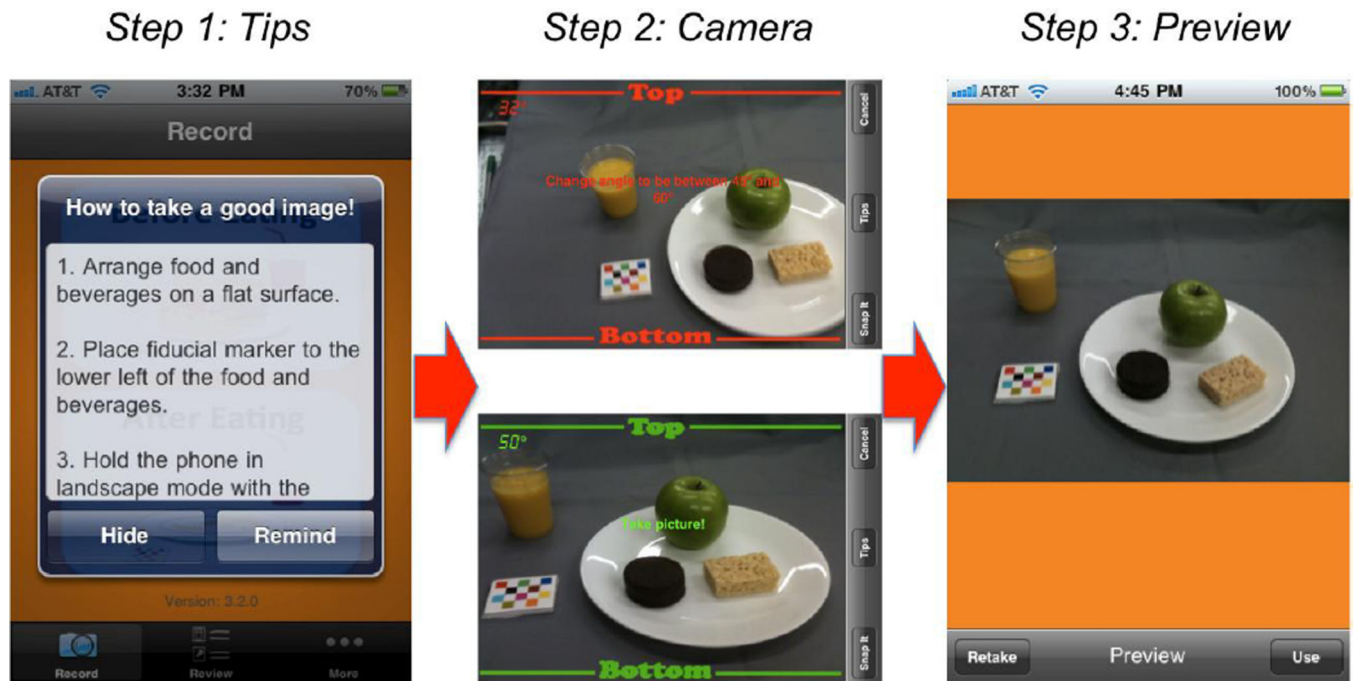
Figure 2.
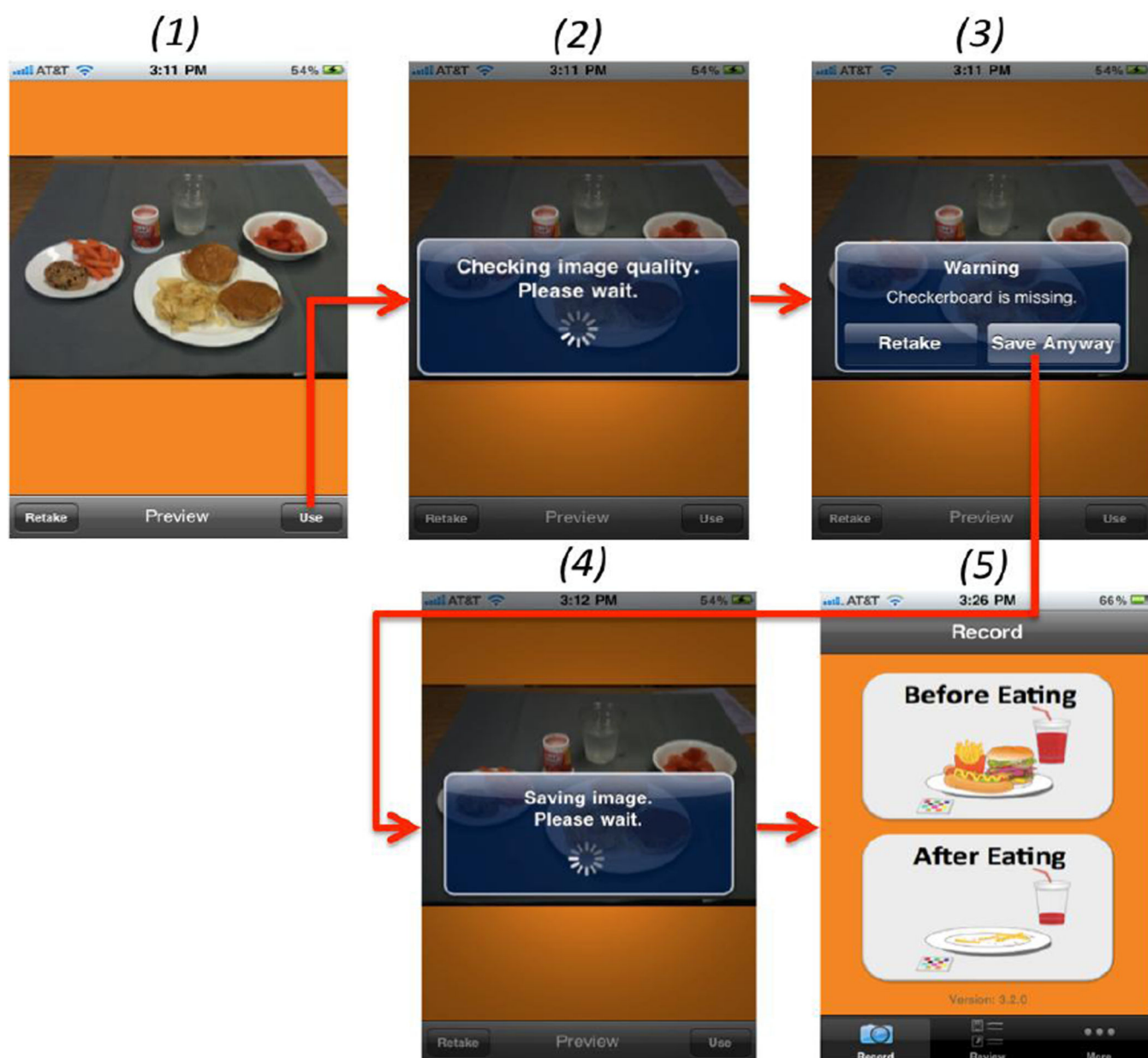Main views of the application.

Figure 3.
Image acquisition process.
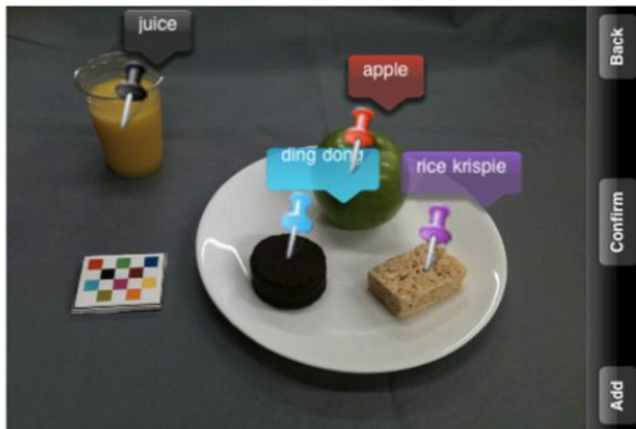
Figure 4.
Image quality checking.
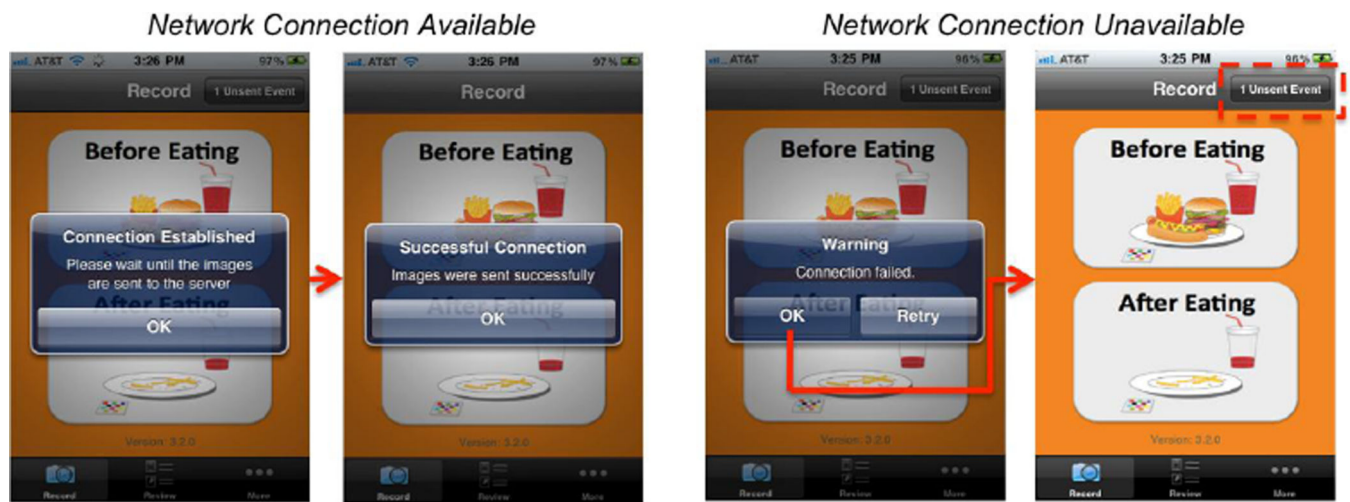
Figure 5.
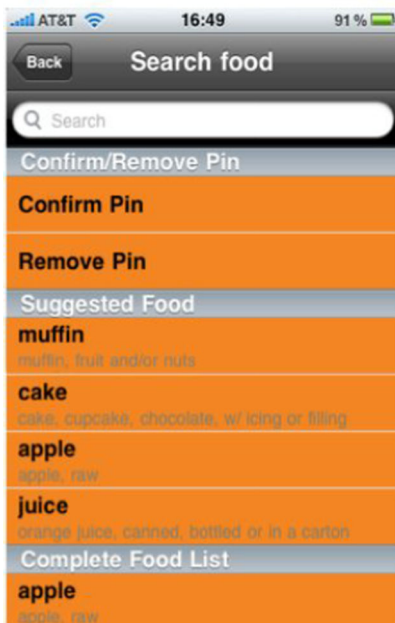Communication with the server after taking image pair.
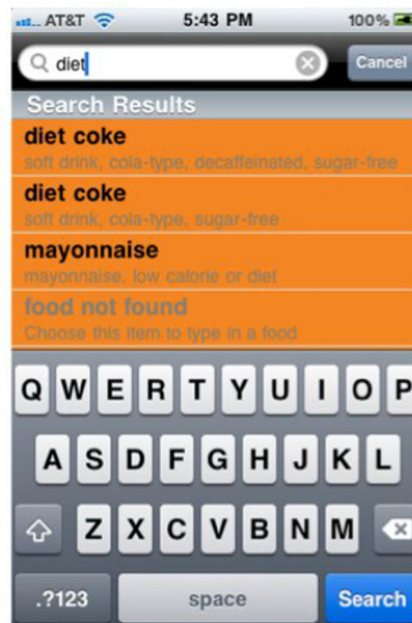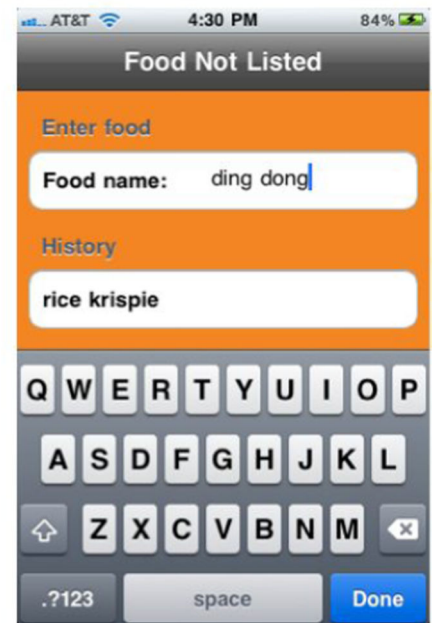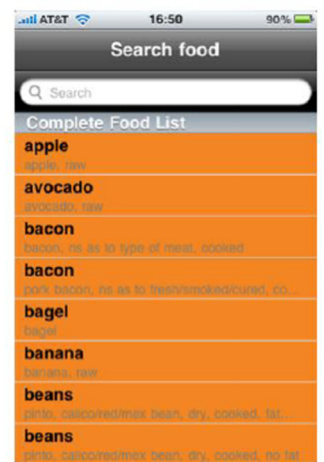
Figure 6.
Image labeling in the review process.
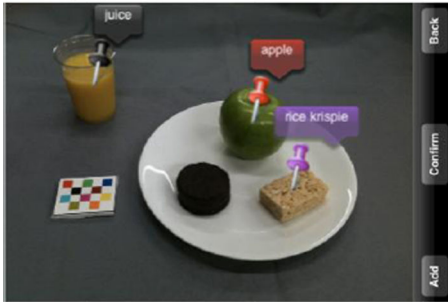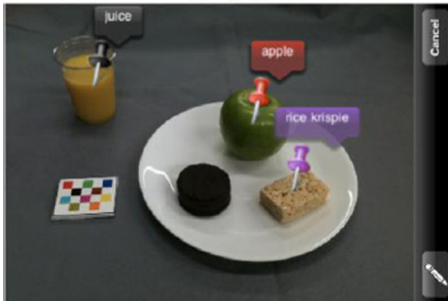
Figure 7.
Adding a new food label to the image.