# Merging Dietary Assessment with the Adolescent Lifestyle

**TusaRebecca E Schap, MSc, RD**[1], **Fengqing M Zhu, MS**[2], **Edward J Delp, PhD**[2], and **Carol J Boushey, MPH, PhD, RD**[1,3,§]

[1]Department of Nutrition Science, College of Health and Human Sciences, Purdue University, 700 W. State St, West Lafayette, IN 47907 USA

[2]School of Electrical and Computer Engineering, Purdue University, 465 Northwestern Ave, West Lafayette, IN 47907 USA

[3]University of Hawaii Cancer Center, 1236 Lauhala St, Suite 407, Honolulu, HI 96813 USA

## Abstract

The use of image-based dietary assessment methods shows promise for improving dietary self-report among children. The Technology Assisted Dietary Assessment (TADA) food record application is a self-administered food record specifically designed to address the burden and human error associated with conventional methods of dietary assessment. Users would take images of foods and beverages at all eating occasions using a mobile telephone or mobile device with an integrated camera, (e.g., Apple iPhone, Google Nexus One, Apple iPod Touch). Once the images are taken, the images are transferred to a back-end server for automated analysis. The first step in this process is image analysis, i.e., segmentation, feature extraction, and classification, allows for automated food identification. Portion size estimation is also automated via segmentation and geometric shape template modeling. The results of the automated food identification and volume estimation can be indexed with the Food and Nutrient Database for Dietary Studies (FNDDS) to provide a detailed diet analysis for use in epidemiologic or intervention studies. Data collected during controlled feeding studies in a camp-like setting have allowed for formative evaluation and validation of the TADA food record application. This review summarizes the system design and the evidence-based development of image-based methods for dietary assessment among children.

## Introduction

Assessment of diet among adolescents is problematic. Energy intake estimates are under-reported among adolescents with reported energy intakes representing 67–88% of energy expenditure as measured by doubly labeled water (Champagne et al., 1998; Bandini et al.,

§Corresponding Author: Carol J Boushey, PhD, MPH, RD, Associate Researcher, Epidemiology Program, University of Hawaii Cancer Center, 1236 Lauhala St. Suite 407, Honolulu, HI 96813, O: 808-564-5915, F: 808-586-2970, CJBoushey@cc.hawaii.edu; and Adjunct Professor, Department of Nutrition Science, Purdue University.

2003). The more accurate reporting was found among younger rather than older adolescents. The progression in age from 11 years to 14 years represents that period of time when the novelty and curiosity of assisting in or self-reporting of food intakes starts to wane and the assistance from parents is seen as an intrusion (Goodwin et al., 2001; Livingstone et al., 2004). Dietary assessment methods need to continue to evolve to meet challenges and there is recognition that further improvements will enhance the consistency and strength of the association of diet with disease risk, especially in light of the current obesity epidemic among youth.

Children and adolescents are eager in terms of adopting new technology. Among early adolescents completing a questionnaire in 2005, 59% (17/29) owned a cell phone and 66% (19/29) reported owning a digital camera (Boushey et al., 2009). When testing a prototype PDA food record tool among 31 adolescents, almost every child indicated previous experience using a PDA and readily adapted to using the tool (Boushey et al., 2009). Recent consumer reports indicate that 75% of adolescents aged 12–17 years have their own mobile telephone (Lenhart et al., 2010). This is a substantial increase since 2004 when only 45% of adolescents had a mobile telephone. Mobile phones have become "indispensable tools in teen communication patterns" (Lenhart et al., 2010). Young people desire a phone that allows for taking and sending images, playing games, social networking, and texting (Lenhart et al., 2010). Mobile devices have evolved to meet market demand for general purpose mobile computing devices and their high-speed multimedia processors and data network capability also make mobile devices ideal as a field data collection tool for dietary assessment.

## Development of a mobile telephone food record

The Technology Assisted Dietary Assessment (TADA) program aims to leverage technology to improve dietary assessment methods. One of the systems under development is a food record application for real time recording of dietary intakes in which images of eating occasions will be captured using a mobile hand-held computing device with an integrated camera (e.g. iPhone, iPod Touch). The food record application deployed on a mobile telephone is referred to as the mobile telephone food record (mpFR). Image processing for identification and quantification of foods and beverages consumed will be employed and indexed with the USDA Food and Nutrient Database for Dietary Studies (FNDDS) for computation of macro- and micro-nutrient intakes. Automated image-driven methods hold the most potential to reduce the burden of many aspects of recording dietary intake for the users and reduces burden of analysis for the researchers.

Given the rapid evolution of mobile technology, design decisions for the mpFR must consider cross-platform usability as much as possible. Initial development of the mpFR was conducted on HTC mobile telephones running Windows for Mobile (Mariappan et al., 2009; Six et al., 2010). For the purpose of developing a product for testing, current platforms' consumer popularity is a major consideration. Platforms currently being deployed include iPhone and smartphones running Google Android (e.g., Motorola and HTC). Schools often prohibit mobile telephone use in schools. Therefore, small handheld mobile devices such as the iPod Touch offer a viable non-telephone alternative that could be suitable for use in schools. All of these mobile handheld devices have high-speed multimedia processors, integrated cameras and data network capability that are necessary components for use with the TADA food record application together with the functions desired by young people.

The TADA food record applications will provide a detailed diet analysis for use in epidemiologic or intervention studies. The primary focus of development is an image-based system in which images captured by the user during eating occasions are sent to the server

for automated food identification and volume estimation (Zhu et al., 2010a; Zhu et al., 2011a). The analyzed image is then returned to the user for confirmation before being indexed with FNDDS for nutrient analysis. The integrated system design is outlined in Figure 1. Other methods of data entry complement the automated image-based method. If an image cannot be used for automated analysis, a user can aid in identification by scanning the bar code of the product when available and aid in volume estimation by manually selecting from a list the portion of the product consumed using multiple measurement descriptors, such as 1 cup, 1 piece (Schap et al., 2011). The user may also manually identify the food item in the image using a type and search mechanism that will search the FNDDS and then the user will estimate the portion consumed. When an image is not available, the user may record foods and beverages or scan the bar code of the product wrapper and estimate the portion consumed.

All of the dietary intake data are sent to a server for indexing with the FNDDS for daily estimates of energy and nutrient intake (Bosch et al., 2011a). All versions of the TADA food record application are designed so that any data entered can be sent to a central server using available networks such as 3G or WiFi. The advantages of storing data on a central server include: 1) allowing the researchers to periodically ensure participants in the study are properly recording food intake, 2) keeping the image analysis within a reasonable timeframe, and 3) reducing the likelihood of lost data that can occur when a mobile device malfunctions.

## Image Analysis

**Segmentation and Identification—**The first step in the image analysis process is to locate each food in the image. This is called image segmentation. Various approaches to segment food items in an image that have been successfully used in many computer vision and image analysis applications have been investigated (Zhu et al., 2008; Zhu et al., 2010a; Zhu et al., 2010b). Results from the segmentation step are used for food labeling and volume estimation. Thus, the accuracy of this step plays a crucial role in the overall performance of the system.

After segmentation, the next step is to extract visual features from each segment that can be used for food identification. A digital image is different from a photograph in that useful information, called metadata, is captured that is not visible, such as the time stamp and digital information. Image analysis methods use visual characteristic features such as color and texture from an image to automatically identify a food. Methods for automatic identification of food using image analysis has been previously published (Mariappan et al., 2009; Mariappan et al., 2005; Zhu et al., 2008; Zhu et al., 2011b). Initially, synthetic plastic food models (e.g., apple, baked beans, roast beef, scrambled eggs) were used and the methods developed resulted in 94% of 32 food objects being correctly identified (Mariappan et al., 2009). The success of this system with synthetic foods provided the foundation for progress to real food images captured by adolescents.

Currently, texture descriptors, local and global features, and a "voting" based classifier to identify food items have lead to enhanced final decisions for food identification (Bosch et al., 2011b; Bosch et al., 2011c). Figure 2 shows a Confusion Matrix for 19 food items in images taken by adolescents. A *confusion matrix* (Schonhoff & Giordano, 2006) is commonly used to assess the accuracy of image classification. In a confusion matrix, classification results are compared to the correct or true results. Dots are used to indicate a match between the results of the automatic classification analysis and the correct or true result. Dots that are aligned on a diagonal straight line represent an accurate (correct) match. In the case of Figure 2, the addition of global features improved the classification system as more matches (dots) are on the diagonal line than in the space above or below the diagonal

line. This simple visual representation allows quick identification of the errors, as well as the distribution of the errors. Figure 2 (a) shows categorization results using global features, while Figure 2 (b) illustrates categorization results by decision fusion of the global and local features. The latter decision fusion approach improves the categorization rate of the classifier considerably by reducing the number of misclassified foods, i.e., fewer non-diagonal elements in Figure 2(b) compared to 2(a).

**Volume Estimation**—A context dependent, automated volume estimation technique is used for approximating food volumes using 3D primitive shapes reconstructed from a single image. Recent work has demonstrated the efficacy of this approach in generating repeatable, low variance volume estimates (Chae et al., 2011). Improved accuracy was achieved by minimizing the false-segmented regions and smoothing the segmentation boundaries of foods.

The input to the volume estimation method consists of the original meal image, the segmented regions, and food identification information (food label and FNDDS food code number) obtained from the image segmentation and classification methods described above (Zhu et al., 2010b). In the camera calibration step, the camera parameters are estimated using the fiducial marker in the original image and this information is used to reconstruct the 3D shape of a food. Given the food name and food code from the food classification process, each food is associated with an appropriate template shape. For example, a glass of milk would correspond to a generalized cylindrical shape and an orange to spherical shape. Once the best-matched template shape is assigned, errors in the segmented region are minimized to improve the next step, 3D volume feature extraction. The shape template is then used to determine geometric information for a food, such as the height, radius and area. The 3D shape is reconstructed to estimate the food volume using the geometric information. Currently, spherical, cylindrical, and arbitrary extruded solids as shape templates are used.

### The User interface

Initial testing of the TADA food record has been done among adolescents (Schap et al., 2011; Six et al., 2010). As previously reported one sample of adolescents was recruited from summer camps held on the university campus and a second sample was recruited from the community. The two samples of adolescents used the TADA food record during meals and then provided feedback about their experience during interactive sessions (Six et al., 2010). 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 meals (08:30, 12:30, and 18:00 hrs) and snacks for a 24-hour period. All foods served were familiar to adolescents and each food was matched to a food code in the FNDDS 3.0.

Observation of adolescents interacting with the TADA food record and weighing of all foods served and plate waste provides an ideal study design for the development of components for the TADA food record. Automated food identification (i.e. automated assignment of an FNDDS food code) can easily be compared to the known food code. The error associated with automated volume estimates can be determined by comparing automated estimate to the true value using a Pearson's correlation coefficient statistic, paired *t*-tests, or Bland Altman plots. Errors detected can be addressed through improvements in the user instruction, the user interface, programming, database linkages, nutrient analysis, device memory, or hardware. Using known foods and carefully recording gram weight of served portions and plate waste allows for development of the most accurate tool possible.

The expertise of a multi-disciplinary team allows for the development of novel tools for improving accuracy of dietary assessment. However, it is critical to use evidence-based development in order to design the TADA food record from the perspective of the user

(Hertzum & Simonsen, 2004; Six et al., 2010). Interaction design is one form of evidence based development that focuses on designing interactive products that support the way people communicate and interacts in their everyday lives (Sharp et al., 2007). Interaction design is an iterative cycle of usability testing in which user feedback is applied to future versions of device development, in this case the TADA food record application. Input from use among adolescents improves accuracy to acceptable levels, enhances usability, minimizes burden, and improves analytical output. The remainder of this review describes the interaction design of individual components of the research including the user interface, automated food identification, and automated volume estimation.

**Capturing an Image—**The primary method of collecting dietary intake data using the mpFR is image-based. Of major concern is an adolescent's ability and willingness to capture images of all eating occasions. Organized meal sessions were ideal for observing adolescents' skills necessary for following directions and capturing images useful for analysis. During meals, adolescents in Samples 1 and 2 used the mpFR running on HTC p4351 mobile telephones (HTC Corp, Taoyuan, Taiwan) running Windows Mobile 6.0 (Microsoft Corp, Redmond, WA). The adolescents were instructed to include all foods and beverages and the fiducial marker, an object of known dimensions and markings, in the image. All foods and beverages appeared in 61 (78%) of the before and after meal images taken for the first meal. In the second meal, 59 (84%) of the before and after meal images included all food and beverages (Six et al., 2010). Similarly, the fiducial marker was completely visible in 54 (69%) of the images taken during the first meal and 53 (76%) in the second meal (Six et al., 2010).

Following the first meal, participants took part in an interactive session where they received additional training on capturing images in various snacking situations (Six et al., 2010). The instruction improved the adolescents' perceptions with regard to the ease of capturing images with the mpFR. Prior to the session, only 11% (n=78) of participants agreed taking images before snacking would be easy. After the session, the percent increased significantly to 32% (p<.00001). For taking images after snacking, there was also improvement (21% before and 43% after, p<.00001) (Six et al., 2010). Adolescents readily adopt new technologies and images useful for analysis can be captured by adolescents. However, creative training sessions will likely enhance cooperation among adolescents.

During interactive feedback sessions, forced choice and open-ended questions were used to collect preferences and perceptions of adolescents in regards to their experience with the mpFR. Input from adolescents clarified their reactions to using the mpFR and was informative for revisions to the TADA food record application. After using the mpFR one time, 79% of adolescents agreed that the software was easy to use (Six et al., 2010). A majority of the adolescents (78%) agreed that they would be willing to use a credit card sized fiducial marker. Therefore, the current version of the fiducial marker is similar in dimension to a credit card.

The TADA food record does not include voice recognition because adolescents have repeatedly responded that they would not feel comfortable using voice recognition to record the names of foods into a mobile device (Boushey et al., 2009; Six et al., 2010). The individuals that consumed all foods and beverages recommended the addition of a touch-option, "ate all food and beverages"; rather than capturing an image of completely empty plates and glasses (Six et al., 2010). The bottom line, adolescents want a tool to have as few steps as possible. Updated versions of the TADA food record have fewer screens to interact with and a more streamlined approach to capturing images. Repeated use, focused training, and timely reminder messages will alleviate errors associated with remembering to take images of all eating occasions that include all foods and beverages. To maintain cooperation

of users, our system relies on a single image. Thus, it is important to assist the user in taking a quality image by providing immediate feedback about image quality. Work is underway to incorporate quick image feedback to users.

**User confirmation and adjustment—**After foods have been automatically identified, the server returns the image with corresponding food identification labels to the user for review. The user can confirm the labels or correct them when mistakes have been made in the automated process. Although the mpFR relies primarily on automated systems, there may be times when the image is not useful for analysis. For example, when an image is blurry the user may be able to identify foods that the system cannot (Schap et al., 2011). In this instance the user can complete the image-assisted record method during which the user can manually type the names of the foods and beverages present in the image.

There may also be times when technical error (e.g. dead battery, software malfunction) or situation (e.g. driving a car) could prevent an image from being taken, or there may be times when the user simply forgets to take an image. An alternative method for self-reporting foods and beverages as well as portions consumed is being developed.

The image assisted record method and the alternate method rely on the users' abilities to correctly identify foods and beverages consumed. To distinguish knowledge from memory, a task was designed to test the abilities of adolescents in Sample 1 to identify foods at the time of consumption (Schap et al., 2011). For Sample 2, the adolescents were provided with a printed image of their meals and were asked to identify the foods in the image approximately 10–14 hours after the meals were consumed. All adolescents wrote down the food and beverage identifications on a blank worksheet.

Adolescents in Sample 1 correctly identified thirty of the thirty-eight foods. For Sample 2, eleven of the thirteen foods were identified correctly. All misidentified foods were identified within the same major food group. For example, Coca-Cola® was misidentified as root beer which is still a cola drink. These results provide evidence that adolescents can correctly identify familiar foods at the time of the meal and that they can look at an image of their meal up to 14 hours post-prandial and correctly identify foods in the image. Thus, providing evidence that delaying the user confirmation step to a time convenient to the user is feasible among adolescents.

To complete the alternate method, the user will need to provide not only the name but also the amount of food consumed. Portions size estimation among adolescents is problematic, and the small screen of a mobile device is a limiting factor for on screen estimation aids. Fourteen hours after the breakfast meal, Sample 2 participants were asked to estimate the amount of each food consumed at breakfast and at day time snacks using one of two estimation aids (i.e. 2 dimensional estimation aid or multiple descriptors) (Schap et al., 2011). The results were congruent with other studies indicating the challenge of estimating portions. Two foods were estimated within ±10% of the true amount consumed at least once when the 2 dimensional estimation aid was used, while six foods were estimated within ±10% of the true amount consumed at least once when the multiple descriptors estimation aid was used (Schap et al., 2011). A single estimation aid does not work well with all foods, however the use of technology may allow for using portion size estimation aids tailored to an individual food. Careful consideration must be used when choosing portion size estimation aids to be displayed during the alternate method on the mpFR.

NIH-PA Author Manuscript

## Summary and Conclusions

The TADA mpFR system takes advantage of the technology present in recent mobile devices in order to integrate digital images, image processing and analysis, and a nutrient database to allow an adolescent user to discretely "record" foods eaten. Further advancements in mobile device technology as well as advances in image analysis techniques can be leveraged to further the success of the TADA research. While technology development often emphasizes the requirements of the system as determined by the engineer or software programmer, the goal of the TADA research aims to focus on the interaction of the user with the system (Sharp et al., 2007; Hertzum & Simonsen, 2004). Involving adolescents has been key to developing a mpFR that will fit into the lives of the users. The adolescents' ideas have been assessed qualitatively and quantitatively to aid in the refinement of the mpFR (Six et al., 2010). Future studies to test the feasibility of implementation and full integration of the mpFR in clinical and population studies of adolescents will further drive the refinement of the application. Dietary assessment methods need to evolve to meet challenges and burdens faced by adolescents. The use of properly designed mobile device applications that work through the paradigm of how young people live and interact in the 'digital' age may address many of the issues outlined as barriers to recording food intake among adolescents. There is recognition that further improvements in dietary assessment methods will enhance the consistency and strength of the association of diet with disease risk, especially in light of the current obesity epidemic among youth.

## Acknowledgments

## Reference List

1. Bandini LG, Must A, Cyr H, Anderson SE, Spadano JL, Dietz WH. Longitudinal changes in the accuracy of reported energy intake in girls 10–15 y of age. The American Journal of Clinical Nutrition. 2003; 78:480–484. [PubMed: 12936932]

2. Bosch, M.; Schap, T.; Khanna, N.; Zhu, F.; Boushey, CJ.; Delp, EJ. Integrated database system for mobile dietary assessment and analysis. Proceedings of the 1st IEEE International Workshop on Multimedia Services and Technologies for E-health; 2011a.

3. Bosch, M.; Zhu, F.; Khanna, N.; Boushey, CJ.; Delp, EJ. Food texture descriptors based on fractal and local gradient information. Proceedings of the European Signal Processing Conference (Eusipco); 2011b.

4. Bosch, M.; Zhu, F.; Khanna, N.; Boushey, CJ.; Delp, EJ. Combining global and local features for food identification in dietary assessment. Proceedings of International Conference on Image Processing (ICIP); 2011c.

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

6. Chae J, Woo I, Kim S, Maciejewski R, Zhu F, Ebert DS, et al. Volume estimation using food specific shape templates in mobile image-based dietary assessment. Proceeding of the SPIE Computational Imaging IX. 2011; 7873:78730K-1–8.

7. Champagne CM, Baker NB, DeLany JP, Harsha DW, Bray GA. Assessment of energy intake underreporting by doubly labeled water and observations on reported nutrient intakes in children. Journal of the American Dietetic Association. 1998; 98:426–433. [PubMed: 9550166]

8. Goodwin RA, Brule D, Junkins EA, Dubois S, Beer-Borst S. Development of a food and activity record and a portion-size model booklet for use by 6- to 17-year olds: a review of focus-group testing. Journal of the American Dietetic Association. 2001; 101:926–928. [PubMed: 11501871]

9. Hertzum, M.; Simonsen, J. Evidence-based development: A viable approach?. New York: ACM; 2004. p. 385-388.

10. Lenhart, A.; Ling, R.; Campbell, S.; Purcell, K. Teens and mobile phones (Rep No Pew Research Center: Pew Internet & American Life Project). Pew Research Center; 2010.

11. Livingstone MBE, Robson PJ, Wallace JMW. Issues in dietary intake assessment of children and adolescents. British Journal of Nutrition. 2004; 92:S213–S222. [PubMed: 15522159]

12. Mariappan A, Igarta M, Taskiran C, Gandhi B, Delp EJ. A low-level approach to semantic classification of mobile multimedia content. Proceedings of the 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Medcical Technologies. 2005; 2005:111–117.

13. Mariappan A, Ruiz MB, Bosch M, Zhu F, Boushey CJ, Kerr DA, et al. Personal dietary assessment using mobile devices. Proc IS&T/SPIE Conference Imaging VII. 2009; 7264:72640Z.1–12.

14. Schap TE, Six BL, Delp EJ, Ebert DS, Kerr DA, Boushey CJ. Adolescents in the United States can identify familiar foods at the time of consumption and when prompted with an image 14h postprandial, but poorly estimate portions. Public Health Nutrition. 2011; 14:1184–1191. [PubMed: 21324224]

15. Schonhoff, TA.; Giordano, AA. Detection and Estimation: Theory and its Applications. 1. New Jersey: Pearson Education, Prentice Hall; 2006.

16. Sharp, H.; Rogers, Y.; Preece, J. Interaction Design: Beyond Human-Computer Interaction. 2. Chichester, West Sussex: John Wiley & Sons, Ltd; 2007. What is interaction design?; p. 1-40.

17. Six BL, Schap TE, Zhu FM, Mariappan A, Bosch M, Delp EJ, et al. Evidence-based development of a mobile telephone food record. Journal of the American Dietetic Association. 2010; 110:74–79. [PubMed: 20102830]

18. Zhu F, Mariappan A, Boushey CJ, Kerr D, Lutes KD, Ebert DS, et al. Technology-assisted dietary assessment. Proceedings SPIE-The International Society for Optical Engineering. 2008; 6814:1–10.

19. Zhu, F.; Bosch, M.; Delp, EJ. An image analysis system for dietary assessment and evaluation. Proceedings of the IEEE International Conference on Image Processing; 2010a. p. 1853-1856.

20. Zhu F, Bosch M, Woo I, Kim SY, Boushey CJ, Ebert DS, et al. The use of mobile devices in aiding dietary assessment and evaluation. IEEE Journal of Selected Topics in Signal Processing. 2010b; 4:756–766. [PubMed: 20862266]

21. Zhu, F.; Bosch, M.; Khanna, N.; Boushey, CJ.; Delp, EJ. Multilevel segmentation for food classification in dietary assessment. Proceedings of the 7th International Symposium on Image and Signal Processing and Analyais (ISPA); 2011a.

22. Zhu F, Bosch M, Schap TE, Khanna N, Ebert DS, Boushey CJ, et al. Segmentation assisted food classification for dietary assessment. Proceedings of the IS&T/SPIE Conference on Computational Imaging IX. 2011b; 7873:1–8.
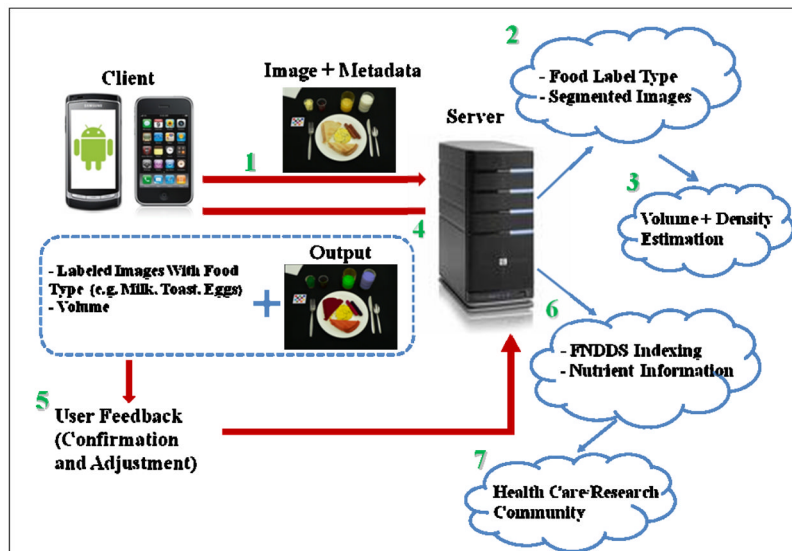
**Figure 1.**
System overview of the mobile telephone food record. Numbers refer to distinct steps in the system.
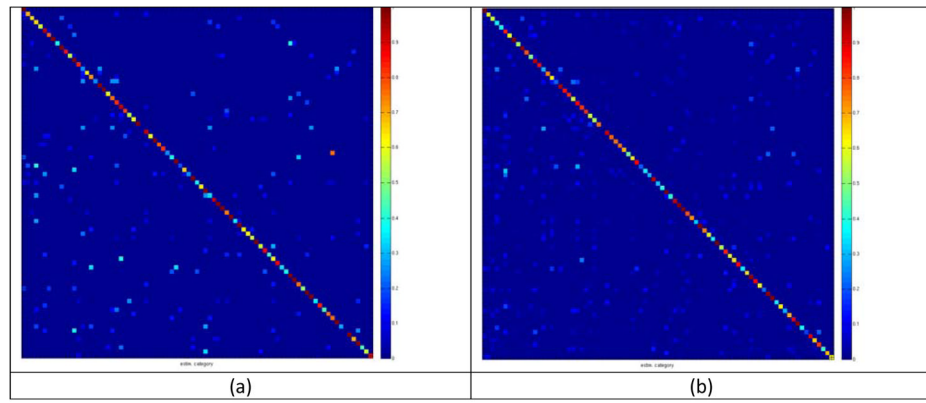
**Figure 2.**
Confusion Matrix (a) Using only global color and texture features and (b) using local and global features. The nearly straight-line performance indicates accurate classification (See text for more information about the confusion matrix).