An Image Procesing Approach for Calorie Intake Measurement

Gregorio Villalobos, Rana Almaghrabi, Parisa Pouladzadeh, Shervin Shirmohammadi
Distributed Collaborative Virtual Environment Research Laboratory
University of Ottawa, Ottawa, Canada
Email: { gvillalobos, ralmaghrabi, ppouladzadeh, shervin}@discover.uottawa.ca

Abstract- **Obesity in the world has spread to epidemic proportions. In 2008 the World Health Organization (WHO) reported that 1.5 billion adults were suffering from some sort of overweightness. Obesity treatment requires constant monitoring and a rigorous control and diet to measure daily calorie intake. These controls are expensive for the health care system, and the patient regularly rejects the treatment because of the excessive control over the user. Recently, studies have suggested that the usage of technology such as smartphones may enhance the treatments of obesity and overweight patients; this will generate a degree of comfort for the patient, while the dietitian can count on a better option to record the food intake for the patient. In this paper we propose a smart system that takes advantage of the technologies available for the Smartphones, to build an application to measure and monitor the daily calorie intake for obese and overweight patients. Via a special technique, the system records a photo of the food before and after eating in order to estimate the consumption calorie of the selected food and its nutrient components. Our system presents a new instrument in food intake measuring which can be more useful and effective.**

Keywords-component: Food intake measurement, Shape recognition, Image processing, Calories measurement.

I. INTRODUCTION

Obesity has become a widespread phenomenon all over the world. The WHO defines obesity based on the Body Mass Index (BMI) of the individual. A person is considered obese when the (BMI) is greater than or equal to 30 (kg/m^2) [1]. According to WHO, in 2008 more than one in ten of the world's adult population was obese [1]. Thus, it is noticeable that obesity and overweightness are linked to a number of chronic diseases such as type II diabetes, breast and colon cancer, and heart diseases. Obesity treatment requires the patient to consume healthy food and decrease the amount of daily food intake, but in most of obesity cases, it is not easy for the patients to measure or control their daily intake due to the lack of nutrition education or self-control. Therefore, using an assistive monitoring food system is very needed and effective for obesity elimination.

In this paper we introduce a new semi-automatic system that will assist dieticians in the monitor of daily nutrient intake for the treatment of obese and overweight patients. The system is the first system expert-out-of-the-loop application for the analysis of food images to aid patients suffering from overweight or obesity. The expert-out-of-the-loop concept will enable the user/patient to obtain the analysis result of the food intake from the application that will try to simulate the calculation procedure performed by the dietician. The system makes use of a set of functionalities key to the success of the process such as the use of pictures, any type of hardware selected must be able to take and use pictures. This is because our approach is focused on image processing by segmentation and image analysis. This feature is a must because the pictures of the food are the main source for proceeding with the rest of our work. The application needs to be mobile, to address this characteristic, we developed our work over mobile applications for Smartphones; in this case, we can consider 3 different environments that will produce a similar effect, the Software Development Kit (SDK) for IPhone, Android and BlackBerry. Any of these solutions will work correctly for our purposes, due to the similar capabilities of the hardware inside these three different technologies. Any of these artifacts can take pictures, run applications, connect to internet, and apply all these features in an easy-to-use interface. In the process of transformation we need to define a simple Measurement Pattern. Once the picture is captured, at some point in the processing, we must use a measurement pattern inside of the image taken by the user, with this we can scale the portions from the image size into a real life size, in order to proceed with the measurement and calorie calculation per portion. For the measurement pattern we need to have something simple, an object practical enough to carry around, and not be aware of its existence, until it is needed. In this case we propose something as simple as the thumb of the patient, initially an image of the thumb will be captured and stored with its measurements, with this pattern, we can then perform the scale and calculations needed. As an alternative option, the user could use a coin instead of the thumb; this will add an extra degree of freedom, in the use of the application. Once the initial conditions are set, the application make use of image processing techniques. One of the ways we can perform a calorie intake measurement for the patient is to let the user take pictures of the food before and after each meal, to capture the exact type of food inside of each image, and then do a simple image subtraction to remove from the calorie consideration the leftovers of the food. This is one of the most difficult sections of our work, but it is the core process

978-1-4673-0882-3/12/$31.00 ©2012 IEEE

of the application, and a good image segmentation and analysis will produce a successful project and application.

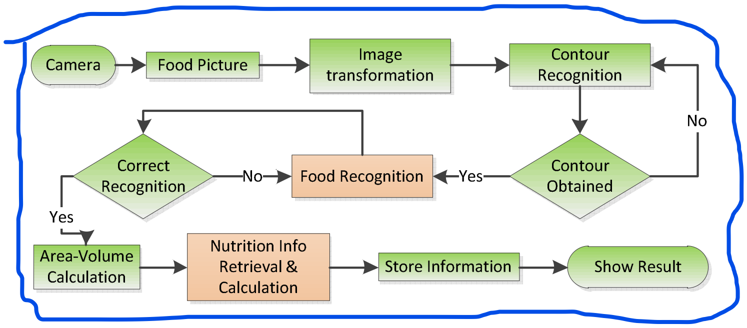

Figure 1: Diagram of the application

Figure 1 shows the diagram of the application proposed by this paper. Though the following sections in this paper the image processing approach is explained and how the segmentation approach is defined, to obtain successfully the resulting images, and the entire set of analysis to support the dietician work.

II. RELATED WORK

This section will present a number of the most common food intake measuring methods, plus their advantages and disadvantages. As a result, we will show the importance of our calorie-intake measurement system which can be used for normal people and clinical purposes in order to enhance the treatment methods for people who suffer from obesity and overweightness.

The first work is the 24-Hour Dietary Recall (24HR) [2]. This procedure is the listing of the daily food intake by using a special format for a period of 24 hours. A brief activity history may be incorporated into the interview to facilitate probing for foods and beverages consumed. The main disadvantage of the 24HR is the delay of reporting the eaten food. This delay in the recording of the daily food intake comes from several factors, such as, age, gender, education, credibility and obesity. Harnack, et al. [3] found important underreporting of large food portions when food models showing recommended serving sizes were used as visual aids for respondents. There are also other studies conducted that focus on underreporting of food intake, in order to define a pattern in this problem present in obesity treatment [4] [5] [6]. Another method is the food frequency questionnaire (FFQ), which uses an external verification based on double labeled water and urinary nitrogen [7]. A brief activity may be included into the interview to facilitate probing for foods and beverages consumed [8]. There are approaches focused on the analysis on how hunger and gender can be deterministic in the level of calories consumed, while images are used to store these behaviors [9]. The use of image processing to analyze food content has been proposed in [8] [10]. In both cases there is a set of pictures for the food before and after its consumption in order to recognize and classify the food and its respective size. In the different methods mentioned, the existence of a premeasured and predefined measurement pattern is used inside of the images, to translate the size in pixels of each portion [10] [11]. All these conditions can generate difficulties, to overcome this, Martin et al. [12] introduced a method where the system will capture the images and send them to a research facility where the analysis and extraction will be performed, producing a considerable delay and an offline processing over the images. There are other methods where the food is weighted before and after eating, or a modified set of kitchen appliances containing an internal scale will evaluate the plate and portions before and after the food intake [13] [14]. But these kinds of approaches generate inconvenience to the users, increasing underreporting generated by the proneness of the user to forget or the unwillingness of the patient to use these kinds of procedures.

It is very important to note that exact or even high level of accuracy is not the main purpose of any such system. The reason is that in general, when it comes to calorie measurement, accuracy is not possible. Even if we put a plate of food in front of a nutrition specialist, s/he will not be able to give us an accurate measure of the calorie inside that food by simply looking at it. There are too many food items and ingredients that cannot be visually inspected. For example, does the food contain animal oil, vegetable oil, or olive oil? Is there soya sauce in it, or hot sauce? Any cream, or sugar, or any other hidden ingredient cannot be considered present in the preparation of the food. Therefore, accuracy is not possible in practice with any mobile or personal assistive system. Rather, the aim of any such system is to estimate the calorie in-take.

III. SYSTEM OVERVIEW

The object of our system is to help people who are suffering from obesity to measure the daily calories they consume, without causing an inconvenience in recording this data. To accomplish this, we use a mobile device with camera that supports wireless connection (such as any of today's mobile phones); the system will enable the mobile device to take pictures of the food for analysis and immediate response to the user. All the analysis will be supported by a set of data stored in a centralized database that will be also connected with the hospital or clinic database where the patient is receiving the treatment. The image processing part of our application takes relevance at this point, because we will use segmentation, analysis and shape recognition to apply inside the complete set of images related to the same food and to measure the amount of calories present in each portion of the images. In the following sections the system is described in more detail.

IV. USER INTERFACE

The user captures three pictures of the food with his/her thumb on a suitable position on the dish so the picture will not only contain the food item, but also the user's thumb, which is used for size calibration. The first two images are taken before the food consumption, one from the top view that will enable us to extract the portions and its corresponding areas, the other from the side of the dish, to analyze the height of the food items inside the dish. With these two measurements, we can obtain a better approximation for the volume, and its translation to calories and nutritional facts. The third picture must be taken at the end of food intake, to subtract from the calculations the food not consumed by the patient. The technique of using the thumb in a photo captured has an important usage in our system, because the thumb is considered as a standard for calculating the dimensions of the food items. Compared to the previous measuring method such as PDAs and the calibration card, thumb is more flexible, controllable and stable standard, giving to the patient the freedom to use the application without the need to carry around uncommon equipment or in this case measurement patterns. As an alternative to the thumb (for disabled patients who might not have a thumb), the user can pace a coin inside of the image, so the system will use this coin instead of the finger, to translate the portions of the food from the picture size into real life size. The system is designed to store the patient's thumb size during its one-time calibration process. Once the food is recognized and the application suggests the type of food, the user is responsible to accept or correct the type of food from the application in the mobile device.

Our Food Recognition System (FRS) is an application which has a user-friendly Interface (GUI) in a mobile device and makes use of the camera built-in the device. The function of our method is to calculate the amount of calories and the nutrients information by using image processing. FRS is a combination of image analyses, where the image is initially corrected and the noise present inside the picture is removed. Once the image is ready, segmentation is applied, with this we extract the different portions of food present in the dish, and for each portion a set of characteristics is obtained based on the average color, size and shape of the food portion. All these characteristics are used to feed the classification procedure that makes use of the Support Vector Machine (SVM) method, and a nutrient database, to semi-automatically detect the food type. This last part if semi-automatic because the user is always prompted to correct the detected food type in case it is incorrect.

V. SYSTEM WORKFLOW

This section will specify the interaction between the user and the system, and the system itself. First, the food must be identified, for this the user must start by placing the food and the thumb close to each other, so he/she will be able to take the appropriate image. Once the application performs the complete analysis of the image and the food portions, the user will get a suggestion of the type of food for each portion, and the user must then accept the suggestion made by the application, or correct the type of the food. The image will be shown to the user at the end of the recognition procedure. For instance, if the user captures a photo of a chicken leg, the system will extract the characteristics of this portion, and based on the shape, size and color the system will apply the SVM to recognize and suggest to the user what kind of food is present in this portion. Once the user accepts or corrects the system's suggestions, the system calculates the photo's dimensions by using built in image processing algorithms to deliver the final result which is the amount of calories and nutritional facts of the captured food item. In order to provide precise results, the previously mentioned procedure is repeated twice before and after eating (in case the patient didn't finish the meal, and we need to subtract food leftovers). When all the measurement and user interaction is done, the calculations and results are shown to the user on top of the image and the system stores the information in the database. The application can perform further analysis, using historical data from the previously stored data in the database, to present graphical interpretation of the information, so the patient and the dietitian can analyze if the calorie intake is according to the patient's intake allowance made by the doctor.

VI. PROOF-OF-CONCEPT IMPLEMENTATION

We have developed a proof-of-concept of the application for iPod Touch and iPhone. We have performed successfully the definition of the images and the contour of the portions inside of the dish. Once the portions and its corresponding contours are defined, we move to extract one by one each portion to analyze one portion at a time. This way, we can focus on several characteristics that correspond to the same portion.

We tested our image processing approach against different other approaches. The images used were obtained with devices similar to what the patient will be using, to produce similar conditions and generate a more realistic analysis. In the image segmentation area, there are different techniques to perform object extraction. One of the first approaches is semi-automatic contour definition, but this option was rejected right away due to the high interaction needed from the user to perform this process. Another method is Watershed Transformation, which is based solely on the gradients analysis, where the pixels with the higher gradient intensity will define the boundaries of the objects, and the pixels with gradient converging to a local intensity minimum are pixels from the same segment. This method was rejected due to the bad results obtained by our initial tests as shown in figure 1, where we can note how the edge and contour definition is not proper, and the final result is producing erroneous detection. The methods evaluated allowed us to understand the general characteristics of the images obtained with the mobile devices. The application needs to deal with illumination problem, distance between the camera and the food, and the quality and resolution provided by the different cameras present in the mobile devices. We therefore selected a combination of methods, in which a color rasterization is performed with a 4^{th} level pyramid, this algorithm allow us to increase the differences between the objects present in the image, reducing the poor illumination effect over the entire scene. The pyramid simplifies the characteristics and physical attributes of the objects present inside the images, the colors are defined as one per object, and the textures of the food that can produce wrong effects over the final result are removed. With this condition applied, then we move to edge accentuation, to reproduce as accurate as possible the edges present in each object, and define with these edges, the corresponding contours. The edge detection is obtained by the algorithm of the nearest neighbor evaluation with an 8x8 matrix, this process was developed based on known schemes such as the one defined Irwin Sobel, but for our set of test images the analysis performed, generated better results when we focus our edge detection using the 8x8 matrix. Once the contours are obtained, the areas of interest are considered, to extract one by one the different portions present inside the images. Figure 2 shows a set of images, that exemplify the deficient results obtained by the initial techniques evaluated.

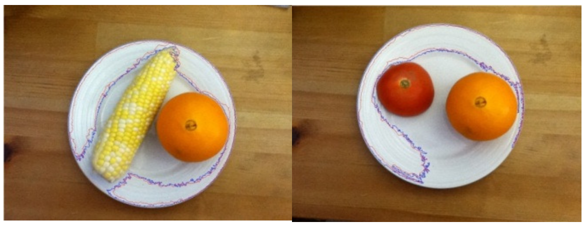

Figure 2: Deficient Results from Watershed Transformation

Figure 3 shows two initial images captured by our approach; note that inside the image the thumb of the patient is present because, as explained above, we need it to find the real life size of the food. Figure 4 shows the images after it has been processed by the application. Rasterization of the average color per object is done to locate the contour of the food portions in a better way. Figure 5 shows the images after the initial process of contour definition; the red and blue lines are defining the contours in each portion and the green lines are showing the regions of interest for each portion located in the dish.

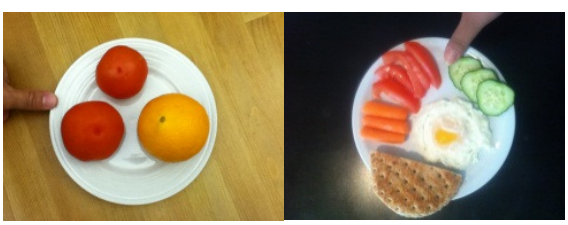

Figure 3: Original Images taken by the user

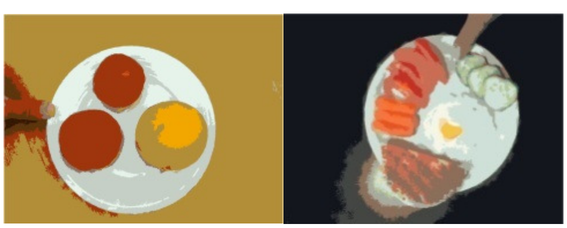

Figure 4: The images after color rasterization

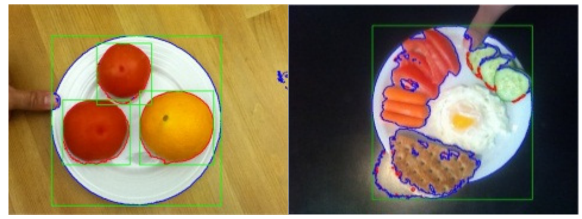

Figure 5: The Images with contours defined

Figure 6 shows images of the extracted portions (in this case one of the oranges) with the average color applied. This will help us in the next steps to define the specific characteristics we need to apply the SVM and make the recognition of the food present in the portion.

First we use these images to measure the size of the portions in pixels, and then we analyze the RGB channels corresponding to each image. We move from RGB to HSV to extract the specific color of each portion and use it as part of the information needed to feed the procedure of prediction. In the process of image analysis and section extraction, one of the key features of our application is the extraction and proper measurement of the thumb present in each picture. For this extraction we move the images from RGB to YCbCr color space, this transformation enable us to successfully locate the sections of the picture containing the skin color, with this the finger was correctly located, and extracted, and the size of the finger established within the picture measurements. Due to the initial configurations of the applications, where the real size of the thumb is introduced, the application is now able to calculate with a good approximation the real size of each portion, based on the measurement pattern obtained with these procedures. Figure 7 shows the image of the thumb of the user, extracted from the food image and its corresponding measurements. With the thumb's size in pixels and its corresponding real life size measurements stored in the application database, we are able to translate the size of each food portion from pixel size to real life size.

Figure 8 shows the simple shapes used to compare the portions and define the basic geometrical shape of the food. For this we implement the Hough lines and circles to enhance the basic characteristics of each portion. The orange of the image shows a Hough circle drawn over the image, based on the contour shape of the portion. Finally we apply the Hu moments analysis to establish what shape is more related to the portion and obtain more details to feed the SVM procedure.

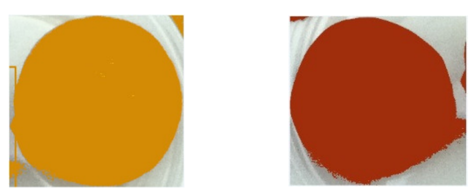

Figure 6: Images of portions and average color

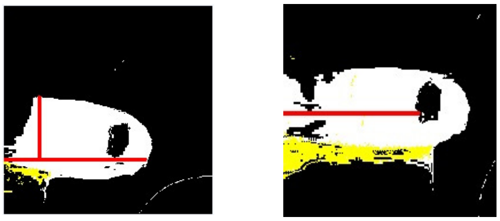

Figure 7: Images of the extracted pattern and its measurement

Figure 8: Images of the shape comparison and definition

In the classification phase, the features of each image are feed into a look up table and its most probable food type will be generated as output of this process. In this paper we have engaged the SVM method as our classifier. It is clear that each classifier needs to be trained before adopted in the food recognition system. After training phase the classifier will be as simple as a look up table, which receives the features and gives the food type. Please note that the training phase has a significant role in categorizing the foods, and its accuracy is related to the number of features and training data. Unlike many existing methods, we have trained our classifier engine using a vast number of features, including shape, color and size. In addition, the classifier is enriched by using wide threshold range for each feature.

While the SVM part of our system is not the focus of this paper (as it requires much more research to be able to work with the majority of food types), we explain here briefly how it works. We select a set of images I_{tr} as a training set and a second set of images I_{ts} as a testing set, which is used to validate the reliability of the model. In the first level, every image is segmented and portions are identified. These portions will then be fed into the SVM which will recognize the portions. In our experiment, the color, size and shape properties of the fruit and food images are extracted after pre-processing. The SVM uses color properties or shape properties for the food recognition.

After recognizing the foods, the nutrient information will be extracted easily. At first, the size of recognized food in centimeter is extracted using the thumb and recognized food size in terms of pixels. Then, using the nutrition table, the appropriate calorie of each food will be reported to user.

VII. CONCLUSIONS

In this paper we proposed a method for measuring food nutritional value; this method can be used as a personal assistive application on smartphones. The system analyzes food and fruit images using a set of image processing and analysis procedures, and makes use of a database of images to extract the information needed like color, shape and size properties from the images. It then uses an SVM to suggest the food portion types. Experimental results show that our image processing technique works better than others in segmenting and extracting the food portions.

REFERENCES

[1] W. H. Organization, "Obesity Fact Sheet," [Online]. Available: http://www.who.int/mediacentre/factsheets /fs311/en/index.html. [Accessed 10 01 2012].

[2] M. Livingstone, P. Robson and a. J.Wallace, "Issues in dietary intake assessment of children and adolescents," *Br.J.Nutr,* vol. 92, p. S213–S222, 2004.

[3] L. Harnack, L. Steffen, D. Arnett, S. Gao and R. Luepker, "Accuracy of estimation of large food portions," *J.Am.Diet.Assoc.,* vol. 104, p. 804–806, 2004.

[4] L. E. a. J. R. R. Klesges, "Who underreports dietary intake in a dietary recall? Evidence from the Second National Health and Nutrition Examination Survey," *Consult.Clin .Psychol,* vol. 63, p. 438–444, 1995.

[5] R. Johnson, R. Soultanakis and D. Matthews, "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," *J.Am.Diet.Assoc,* vol. 98, p. 1136–1140, 1998.

[6] J. Tooze, A. Subar, F. Thompson, R. Troiano, A. Schatzkin and V. Kipnis, "Psychosocial predictors of energy underreporting in a large doubly labeled water study," *Am.J.Clin.Nutr,* vol. 79, p. 795–804, 2004.

[7] L. Bandini, A. Must, H. Cyr, S. Anderson, J. Spadano and W. Dietz, "Longitudinal changes in the accuracy of reported energy intake in girls 10-15 y of age," *Am.J.Clin.Nutr.,* vol. 78, p. 480–484, 2003.

[8] W. Luo, H. Morrison, M. d. Groh, C. Waters, M. DesMeules, E. Jones-McLean, A.-M. Ugnat, S. Desjardins and M. L. a. Y. Ma, "The burden of adult obesity in Canada," *Chronic Diseases in Canada,* vol. 27, no. 4, pp. 135-144, 2007.

[9] S. Frank, N. Laharnar, S. Kullmann, R. Veit, C. Canova, Y. L. Hegner, A. Fritsche and H. Preissl, "Processing of food pictures: influence of hunger, gender and calorie content," *Brain Research ,* vol. 1350, pp. 159-166, 2010.

[10] M. Sun, Q. Liu, K. Schmidt, J. Yang, N. Yao, J. D. Fernstrom, M. H. Fernstrom, J. P. DeLany and R. J. Sclabassi, "Determination of Food Portion Size by Image Processing," *30th Annual International IEEE EMBS Conference,* pp. 871 - 874, August 2008.

[11] Y. Saeki and F. Takeda, "Proposal of Food Intake Measuring System in Medical Use and Its Discussion of Practical Capability," *Springer-Verlag Berlin Heidelberg,* vol. 3683, p. 1266–1273, 2005.

[12] C. K. Martin, S. Kaya and B. K. Gunturk, "Quantification of food intake using food image analysis," *Conference Proceedings of the International Conference of IEEE Engineering in Medicine and Biology Society,* vol. 2009, pp. 6869-6872, 2009.

[13] P.-Y. Chil, J.-H. Chen, H.-H. Chu and J.-L. Lo, "Enabling Calorie-Aware Cooking in a Smart Kitchen," *Computing 2008 Springer-Verlag,* vol. 5033, pp. 116-127, 2008.

[14] M. S. Westerterp-Plantenga, "Eating behavior in humans, characterized by cumulative food intake curves-a review," *Neuroscience and Biobehavioral Reviews,* vol. 24, p. 239–248, March 2000.