This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Measuring Calorie and Nutrition From Food Image

Parisa Pouladzadeh, Shervin Shirmohammadi, *Senior Member, IEEE*, and Rana Al-Maghrabi

***Abstract*— As people across the globe are becoming more interested in watching their weight, eating more healthy, and avoiding obesity, a system that can measure calories and nutrition in every day meals can be very useful. In this paper, we propose a food calorie and nutrition measurement system that can help patients and dietitians to measure and manage daily food intake. Our system is built on food image processing and uses nutritional fact tables. Recently, there has been an increase in the usage of personal mobile technology such as smartphones or tablets, which users carry with them practically all the time. Via a special calibration technique, our system uses the built-in camera of such mobile devices and records a photo of the food before and after eating it to measure the consumption of calorie and nutrient components. Our results show that the accuracy of our system is acceptable and it will greatly improve and facilitate current manual calorie measurement techniques.**

***Index Terms*— Calorie measurement, food image processing, obesity management.**

I. Introduction

OBESITY in adults has become a serious problem. A person is considered obese when the body mass index is higher than or equal to 30 (kg/m^2) [1]. In 2008, more than one in ten of the world's adult population were obese [1], but in 2012 this figure has risen to one in six adults [2], an alarming growth rate. Recent studies have shown that obese people are more likely to have serious health conditions such as hypertension, heart attack, type II diabetes, high cholesterol, breast and colon cancers, and breathing disorders. The main cause of obesity is the imbalance between the amount of food intake and energy consumed by the individuals [3]. Therefore, to lose weight in a healthy way, as well as to maintain a healthy weight for normal people, the daily food intake must be measured [4]. In fact, all existing obesity treatment techniques require the patient to record all food intakes per day to compare the food intake with consumed energy. However, in most of the cases, unfortunately patients face difficulties in estimating and measuring the amount of food intake due to the self-denial of the problem, lack of nutritional information, the manual process of writing down this information (which is tiresome and can be forgotten), and other reasons. As such, a semiautomatic monitoring system to record and measure the amount of calories consumed in a meal would be of great help not only to patients and dietitians in the treatment of obesity, but also to the average calorie-conscious person. Indeed, a number of food intake measuring methods have been developed in the last few years. However, most of these systems have drawbacks such as usage difficulties or large calculation errors. Furthermore, many of these methods are for experimental practices and not for real life usage, as we shall see in the Section II.

In this paper, we propose a personal software instrument to measure calorie and nutrient intake using a smartphone or any other mobile device equipped with a camera. Our system uses image processing and segmentation to identify food portions (i.e., isolating portions such as chicken, rice, vegetables, and so on, from the overall food image), measures the volume of each food portion, and calculates nutritional facts of each portion by calculating the mass of each portion from its measured volume and matching it against existing nutritional fact tables. While a preliminary description of our work has been presented in [5], here we extend it by proposing a more accurate measurement method for estimating food portion volume, which also works for food portions with an irregular shape, and by evaluating our approach with more food items. More importantly, the segmentation features are enriched by involving texture as well as color, shape, and size of the objects. Our results show reasonable accuracy in the estimation of nutritional values of food types for which our system has been trained.

Color and texture are the fundamental characters of natural images, and play an important role in visual perception. Color has been used in identifying objects for many years. Texture is one of the most active topics in machine intelligence and pattern analysis since the 1950s which tries to discriminate different patterns of images by extracting the dependency of intensity between the pixels and their neighboring pixels [6], or by obtaining the variance of intensity across pixels [7]. Recently, different features of color and texture are combined together to measure food nutrition more accurately [8].

In our proposed system, we also aim at using smartphones as monitoring tools as they are widely accessible and easy to use. However, compared with existing work, our system has the following contributions.

1) Our system is currently the only one that not only explains and discusses uncertainties in image-based food calorie measurement, but also measures and presents actual uncertainty results using food images and its application scenario. This puts our system properly

Manuscript received September 24, 2013; revised December 3, 2013; accepted January 3, 2014. The Associate Editor coordinating the review process was Dr. Domenico Grimaldi.

P. Pouladzadeh and S. Shirmohammadi are with the Distributed and Collaborative Virtual Environment Research Laboratory, University of Ottawa, Ottawa, ON K1N 6N5, Canada, and also with the College of Engineering and Natural Sciences, Istanbul Şehir University, Istanbul 34469, Turkey (e-mail: ppouladzadeh@discover.uottawa.ca; shervin@discover.uottawa.ca).

R. Al-Maghrabi with the Distributed and Collaborative Virtual Environment Research Laboratory, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: ralmaghrabi@discover.uottawa.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2014.2303533

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

in the context of instrumentation and measurement research, and leads to more meaningful results for food recognition systems.

2) To the best of our knowledge, this is the first study of a food image segmentation, classification, identification, and calorie measurement system that not only uses 3000 images, but also under different conditions such as using different cameras, lighting, and angles. We also use a variety of food such as solid or liquid food, and mixed or nonmixed food. Other existing work uses much fewer images (typically hundreds) of mostly very specific food, and also do not consider the above condition variations. For example, Qin *et al.* [9] has used the shape and texture features with only 180 images of food with a very distinct shape and texture, [10] has used only fruits in fruit salad, and [11] has used 120 pizza images. From a measurement perspective, our study and results are more comprehensive, meaningful, and generalizable.
3) In our proposed system, we use more features than other systems, including color, texture, size, and shape, whereas most of the existing methods in this area, such as [9], use only color and shape features. As we have shown in Section VI and Table II, using four features significantly increases the accuracy of the system compared with using fewer features.
4) We design a method to apply Gabor filter for texture segmentation of food images. To do this, a bank of Gabor filters with different desired orientations and wavelength are applied to an image. The outcome of each of these Gabor filters is a 2-D array, with the same size of the input image. The sum of all elements in one such array is a number that represents the matching orientation and spatial frequency of the input image. In our method, six orientations are used as Gabor parameter.

The rest of this paper is organized as follows. Section II covers related work in this area, while Section III presents a brief background of calorie measurement requirements and available calorie tables. Section IV presents our system design, which is followed by Section V, where our food portion volume measurement technique is proposed. Section VI covers the performance evaluation of our proposed method, while Section VII analyzes the proposed work. Finally, Section VIII concludes this paper as well as providing a brief discussion of future works.

II. Related Work

There have been a number of proposed methods for measuring daily food's dietary information. One example, which is typical of current clinical approaches, is the 24-h dietary recall [12]. The idea of this method is the listing of the daily food intake using a special format for a period of 24 h. This method requires a trained interviewer, such as a dietician, to ask the respondent to remember in details all the food and drinks she/he has consumed during a period of time in the recent past (often the previous 24 h). The 24 h requires only short-term memory, and if the recall is unannounced, the diet is not changed. In addition, the interview is relatively brief (20–30 min), and the subject burden is less in comparison with other food recording methods [13]. However, it is not always easy for a person to remember the actual contents as well as the amount of the food intake. In addition, to see an expert every 24 h is difficult and in many cases not feasible. The great majorities of existing clinical methods are similar to this, and typically require food records to be obtained for three to seven days, with seven days being the gold standard [5]. The problem with this manual approach is obvious: people not remembering exactly what they ate, forgetting to take note, and needing to see an expert dietician on a very frequent basis so the dietician can guess how much calories and nutrient the person has taken.

To alleviate the shortcomings of these clinical methods, researchers have been trying to come up with improved techniques. Some of these techniques require the person to take a picture of the food before eating it, so that the picture can be processed offline, either manually or automatically, to measure the amount of calorie. For example, the work in [14] proposes a method that uses a calibration card as a reference; this card should be placed next to the food when capturing the image, so that the dimensions of the food are known. However, this card must always be present in the photo when the user wants to use the system. The drawback is that the system will not work without this card, which means that in the case of misplacement or absence of the card, the system will not work. Another method uses the photo of the food and feeds that to a neural network developed by researchers in [15]. But the user must capture the photo in a special tray (for calibration purposes), which might not be always possible and so the method might be difficult to follow for the average user. A personal digital assistive (PDA) system has also been proposed for food calorie measurement in [16], where patients use the PDA to record their daily food intake information on a mobile phone. But it has been shown that the result of the portion estimation has significant error and also it takes a long time for the user to record the information [17]. Yet another approach appears in [18], where the picture of the food taken with a smartphone is compared with the photos of predefined foods with known nutritional values, which are stored in a database, and the values are estimated based on picture similarity. The main disadvantage of this system is that it does not consider the size of the food, which is extremely important.

Compared with the above methods, our proposed system has fewer of their shortcomings. Our measurement system also uses a photo of the food, taken with the built-in camera of a smartphone, but uses the patient's thumb for calibration, which solves the problem of carrying cards or special trays. More specifically, an image of the thumb is captured and stored with its measurements in the first usage time (first time calibration). This unique method will lead to relatively accurate results without the difficulties of other methods. Food images will then be taken with the user's thumb placed next to the dish, making it easy to measure the real-life size of the portions. We then apply image processing and classification techniques to find the food portions, their volume, and their nutritional facts. But before discussing the details of our system, let us first review some background about calorie measurement and its requirements.

III. Background

A. Required Accuracy of the Measurement System

Before discussing any technical issues, it is important to understand what level of accuracy is expected from our system. To answer this question, we must first see what level of accuracy existing clinical methods have in their measurement of food's nutritional facts. There are two things to consider. First, if we put a plate of food in front of an expert dietician, she/he cannot give an accurate measurement of its nutritional facts by simply looking at it or even examining it manually, because it is impossible to know the exact contents of the dish, such as if this dish contains salt, and if so how much, or contains oil, and if so what type (olive, corn, animal-based, etc.), and how much, and so on. In addition, some food portions can be obstructed, for example, a piece of meat could be deep inside a soup, making it invisible to the dietician. Therefore, we can see already that high accuracy of calorie measurement is not possible in real life. Second, when we add this to what happens in existing clinical methods such as [4], in which the dietician goes over a list of food items recorded by the patient without necessarily even seeing the actual food or its picture, and without knowing size of portions, it becomes clear that accuracy is decreased even more.

This is very important, because it directly affects the objectives of our system. The goal of our measurement system is therefore to design an automated measurement tool running on a smartphone or other mobile devices with built-in camera that facilitates (i.e., makes it easier) to record food intake, measure the size of food portions, and measure nutritional facts, compared with the existing clinical methods. Our goal is not to necessarily have high accuracy, because as explained above such accuracy is not possible in practice. Of course, the more accurate the system is the better the end results, and this is why in this paper we have tried to measure the size of food portions as accurately as possible. However, it is very important to understand that high accuracy is not possible when dealing with food pictures only.

B. Measurement Unit: Calorie Definition and Nutritional Tables

Calorie is a typical measuring unit, which is defined as the amount of heat energy needed to raise the temperature of one gram of water by 1° [19]. This unit is commonly used to measure the overall amount of energy in any food portion that consists of the main food components of carbohydrate, protein, and fat. In addition to gram units, calorie units are also adopted in developing nutritional facts tables. Each person should take a certain amount of calories daily. If this amount is increased, it will lead to gain weight.

Table I shows a small sample of a typical nutritional facts table, this specific one from Health Canada [20]. Such tables are readily available from international or national health organizations around the world. Our proposed system relies on such tables as a reference to measure nutritional facts from any selected food photo.

TABLE I
SAMPLE OF A TYPICAL NUTRITIONAL TABLE

Food Name	*Measure*	*Weight (grams)*	*Energy*
Apple with skin	1	140	80
Potato, boil, no skin	1	135	116
Orange	1	110	62
tomatoes, raw	1	123	30
Bread white, commercial	1	100	17
Cake	1	100	250
Egg	1	150	17
Cucumber	1	100	30
Banana	1	100	105
Orange	1	110	62

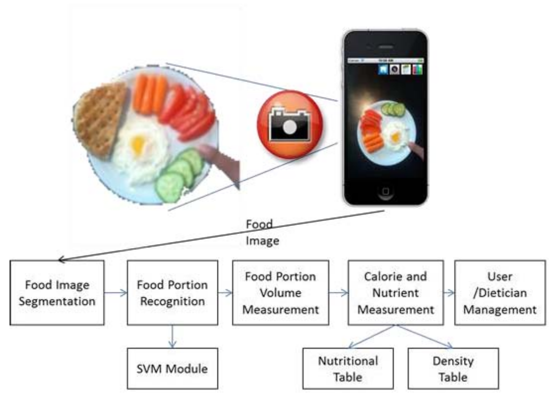

Fig. 1. Overall system design.

IV. Proposed System

The overall design of our system and its blocks are shown in Fig. 1. As the figure shows, at the early stage, images are taken by the user with a mobile device followed by a preprocessing step. Then, at the segmentation step, each image will be analyzed to extract various segments of the food portion. It is known that without having a good image segmentation mechanism, it is not possible to process the image appropriately. That is why we have jointly used color and texture segmentation tools. We will show how these steps lead to an accurate food separation scheme. For each detected food portion, a feature extraction process has to be performed. In this step, various food features including size, shape, color, and texture will be extracted. The extracted features will be sent to the classification step where, using the support vector machine (SVM) scheme, the food portion will be identified. Finally, by estimating the area of the food portion and using some nutritional tables, the calorie value of the food will be extracted. The thumb of the user and its placement on the plate are also shown in Fig. 1. There is a one-time calibration process for the thumb, which is used as a size reference to measure the real-life size of food portions in the picture. We reported the concept of using the thumb for calibration, as well as its implementation and evaluation in [21] and [22], respectively, and so we do not repeat them here. An example of food picture capturing and thumb isolation and measurement are shown in Fig. 2.

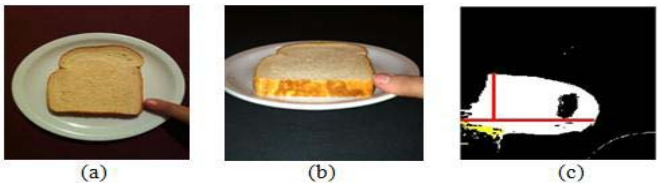

Fig. 2. (a) and (b) Test images with food and thumb. (c) Calculation of the thumb dimensions.

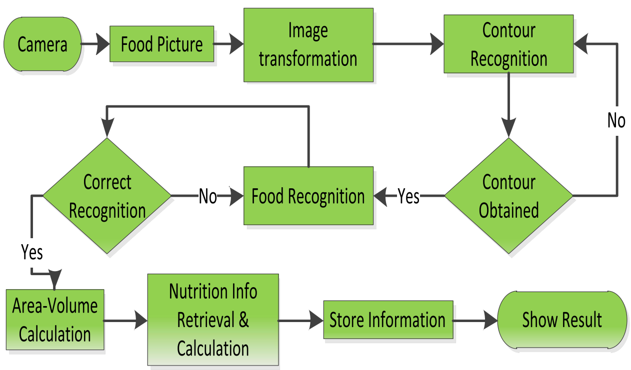

Fig. 3. System's flowchart.

Compared with the calibration method of similar systems, using the thumb is more flexible, controllable, and reliable. For users with thumb disability or amputated thumbs, another finger or a coin can be used instead, the latter still more ubiquitous than special plates or cards used in other systems.

Fig. 3 shows the overall sequence of steps in our system. The user captures two photos of the food: one from above and one from the side; the side photo is needed to measure depth, to have a more accurate volume measurement, as will be explained in Section VI.

The system uses image segmentation on the photo taken from the top and uses contours to isolate various food portions.

The detailed design, implementation, and evaluation of this image processing and segmentation component were described in [22]. For texture features, we used Gabor filters to measure local texture properties in the frequency domain.

We used a Gabor filter-bank proposed in [23]. It is highly suitable for our purpose where the texture features are obtained by subjecting each image to a Gabor filtering operation in a window around each pixel. We can then estimate the mean and the standard deviation of the energy of the filtered image. The size of the block is proportional to the size of the segment. A Gabor impulse response in the spatial domain consists of a sinusoidal plane wave of some orientation and frequency, modulated by a 2-D Gaussian envelope. It is given by

$$h\,(x, y) = -\exp\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right\}\cos\,(2\pi U_x + \varphi) \qquad (1)$$

where U_x and φ are the frequency and phase of the sinusoidal plane wave along the z-axis (i.e., the 0° orientation), and σ_x and σ_y are the space constants of the Gaussian envelope along the z- and y-axis, respectively.

A Gabor filter-bank consists of Gabor filters with Gaussian kernel function of several sizes modulated by sinusoidal plane waves of different orientations from the same Gabor-root filter, as defined in (1), it can be represented as

TABLE II

DIFFERENT TEXTURE

Label	*Class*	*Samples*
1	Soft	400
2	rough	450
3	smooth	180
4	porous	320
5	wavy	200

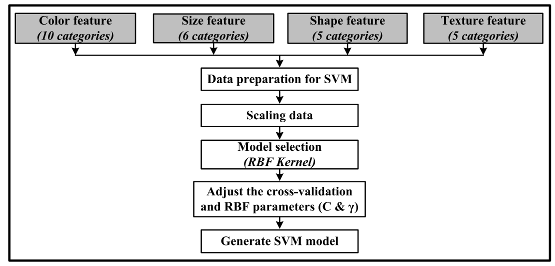

Fig. 4. SVM algorithm.

$$g_{m,n}(x, y) = a^{-m}h\big(x', y'\big) \qquad a > 1 \qquad (2)$$

where

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \\ \theta &= \frac{n\pi}{k} \quad (k = \text{total orientation},\ n = 0, 1, \ldots, k-1, \\ &\qquad\qquad\qquad\qquad m = 0, 1, \ldots, s-1). \end{aligned}$$

Give an image $I_E\,(r, c)$ of size HxW, the discrete Gabor filtered output is given by a 2-D convolution

$$I_g\,(r, c) = \sum_{s,t} I_E\,(r-s, c-t)g_{m,n}(st). \qquad (3)$$

As a result of this convolution, the energy of the filtered image is obtained and then the mean and standard deviation are estimated and used as features. We used the following parameters: five scales ($S = 5$) and six orientations ($K = 6$). In our model, we used Gabor filter for texture segmentation. In the implementation phase, each image is divided into 4×4 blocks, and each block is convolved with Gabor filter. Six orientations and five scales Gabor filters are used, and the mean and variance of the Gabor sizes are calculated for each block. In our project, using Gabor filter, we can identify five different textures and their identities as soft, rough, smooth, porous, and wavy, as shown in Table II. In this table, for each texture, the number of used image samples for training phase is reported as well.

As the figure below shows, we have used these features as our classification inputs and the results will be the input of the SVM phase. For each feature, several categories are engaged, as shown in Fig. 4.

Some examples of various food types and their segmented portions are shown in Fig. 5.

Once the food items are segmented and their features are extracted, the next step is to identify the food items using

Fig. 5. Segmentation of dishes into food portions.

Fig. 6. SVM module verifies with the user the type of foods it has determined [18].

statistical pattern recognition techniques. Afterward, the food item has to be classified, using SVM mechanism [24], [25].

SVM is one of the popular techniques used for data classification. A classification task usually involves training and testing data, which consist of some data instances. Each instance in the training set contains one class label and several features. The goal of SVM is to produce a model, which predicts target value of data instances in the testing set, which are given only by their attributes.

In our model, we use the radial basis function (RBF) kernel, which maps samples into a higher dimensional space in a nonlinear manner. Unlike the linear kernels, the RBF kernel is well suited for the cases in which the relation between the class labels and attributes is nonlinear.

In our proposed method, the feature vectors of SVM contain five texture, five color, three shape, and five size features. The feature vectors of each food item, extracted during the segmentation phase, will be used as the training vectors of SVM.

For increasing the accuracy, after the SVM module has determined each food portion type, the system can optionally interact with the user to verify the kind of food portions. For instance, it can show a picture of the food to the user, annotated with what it believes are the portion types, such as chicken, meet, vegetable, and so on, as described in [21], and shown in Fig. 6. The user can then confirm or change the food type. This changes the system from an automatic one into a semiautomatic one; however, it will increase the accuracy of the system.

The system then measures the volume of each food portion and converts it to mass, using available density tables, and finally uses the mass and nutritional tables to measure the overall calorie and nutrients in the food. These two latter components; i.e., food portion volume measurement and calories measurement, are the focus of this paper and will be explained in the following section.

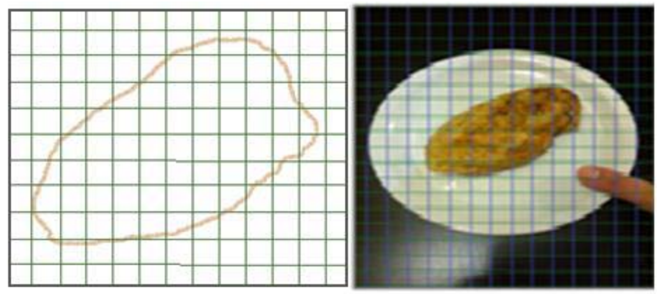

Fig. 7. Methodology for food portion area measurement.

The system also has a module that allows the user or the dietician to use the measurement results and manage the user's eating habits or clinical program. This module provides useful graphs such as daily intake, weekly intake, comparison between various dates, and percentage change in calorie consumption, as discussed in [21].

V. Proposed Measurement Method

A. Food Portion Volume Measurement

As explained before, to measure the size of the food inside the dish, two pictures must be taken: one from the top and one from the side, with the user's thumb placed beside the dish when taking the picture from the top. The picture from the side can be used to see how deep the food goes, and is needed for measuring the food portions' volumes. The system, which already has the dimensions of the user's thumb, can then use this information to measure the actual area of each food portion from the top picture, and can multiply this area by the depth (from the side picture) to estimate the volume of food. Let us see this in more details in the following paragraphs.

To calculate the surface area for a food portion, we propose to superimpose a grid of squares onto the image segment so that each square contains an equal number of pixels and, therefore, equal area. Fig. 7 shows an example with an actual food portion. The reason for using a grid is twofold. First, compared with other methods, the grid will more easily match with irregular shapes, which is important for food images because most of the food portions will be irregular. Naturally, there will be some estimation error, but this error can be reduced by making the grid finer. Second, depending on the processing capabilities of the user's mobile device and the expected system response time from the user's perspective, we can adjust the granularity of the grid to balance between the two factors. If the grid is made finer, measurements become more accurate but will take longer time, and if the grid is made coarser, measurements become less accurate but the response time will be faster.

The total area (TA) of the food portion is calculated as the sum of the sub areas (Ti) for each square (i) in the grid, as

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Fig. 8. Calculating area and volume of regular shapes in food images [5].

shown in equation (4)

$$\mathrm{TA} = \sum_{i=1}^{n} Ti \tag{4}$$

where n is the total number of squares in the food portion's area. After that, and using the photo from the side view, the system will extract the depth of the food, d, to calculate the food portion's volume, V, using the following:

$$V = \mathrm{TA} \times d. \tag{5}$$

For better accuracy, if some food portions happen to be regular shapes such as square, circle, triangle, and so on, we can use geometric formulas to calculate their area, instead of using a grid. This, however, requires an additional module that can recognize regular shapes. Fig. 8 shows some example calculations for regular shapes in a set of different food images.

B. Calorie and Nutrition Measurement

The volume measurement method described above is really just an interim step to measure the mass of the food portion. Mass is what we really need since all the nutritional tables are based on food mass. Once we have the mass, we can use these tables to calculate the amount of calories and other nutrition, as described next.

It is known that the nutritional facts database is an important component for a useful and successful food recognition system [26]. The data of nutritional values of foods are stored in these tables and are available from national and international health organizations. These tables, similar to the one shown in Table I, help us to calculate the amount of calories quickly and without reference to the Internet or an expert.

At this point, we have the measurement for the volume of each food portion, and we can use the following general mathematical equation to calculate their mass:

$$M = \rho V \tag{6}$$

where M is the mass of the food portion and ρ is its density. Food density can also be obtained from readily available tables. For example, aquacalc provides a volume to mass conversion for 3199 food items and ingredients [27].

To extract the density of each food portion, the system needs to know the type of the food, which is done by our SVM-based food recognition module. An example of the information that is fed into the SVM module is shown in Fig. 9 right column. The SVM module uses this information and recognizes the type of food for each portion [28]. In addition, as mentioned earlier, at this stage, the system can ask the user to verify whether the food type recognized by the SVM module is correct. If not, the user can then enter the correct type, as shown in Fig. 6.

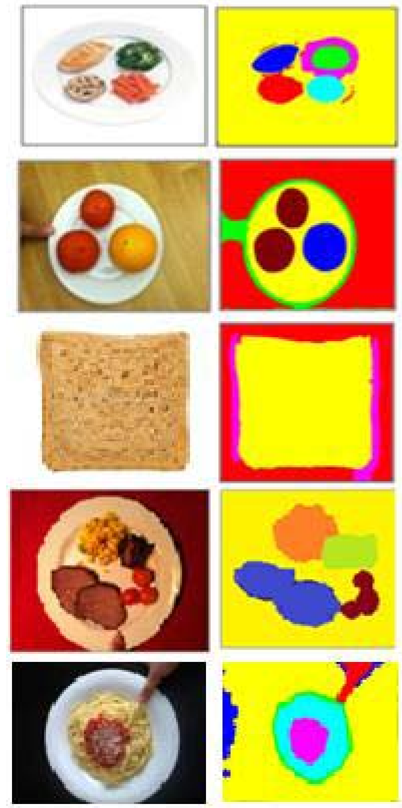

Fig. 9. Before (left) and after (right) color analysis and contour detection. Right column is fed into SVM.

Now, the system can calculate the mass by having the type of food. Therefore, the amount of calorie and nutrition of each food portion can be derived using nutritional tables, such as Table I, and based on the following:

$$\text{Calorie in the photo} = \frac{\text{Calorie from table} \times \text{Mass in the photo}}{\text{Mass from table}}. \tag{7}$$

C. Partially Eaten Food

It is possible that a user does not finish the entire food captured in the first picture that was taken before eating the food. If so, we propose a simple technique to increase measurement accuracy in such cases. If a user does not finish a meal, she/he should take another top picture of what is left of the meal. All of the above process can then be repeated on this new picture to calculate the amount of calorie and nutrient in the remaining food. The actual value of in-take is then adjusted by deducting the values of the remaining food.

VI. Performance Evaluation

A. Evaluation Strategy

We have implemented our system as a software prototype, where we successfully segmented the food images

TABLE III
RESULTS OF FOOD AND FRUIT RECOGNITION SYSTEM

No.	Food items	Recognition Rate (%)					
		Using Color Features	*Using Texture Features*	*Using Size Features*	*Using Shape Features*	*Using All Features*	*Using All Features (10 fold cross-validation)*
1	Apple	60.33	85.25	31.22	22.55	97.64	91.41
2	Orange	65.38	79.24	41.04	71.33	95.59	90.19
3	Corn	52.00	81.93	71.33	34.61	94.85	97.00
4	Tomato	71.29	69.81	48.09	45.01	89.56	79.82
5	Carrot	74.61	79.67	69.30	65.19	99.79	92.34
6	Bread	56.11	61.56	35.55	35.20	98.39	93.50
7	Pasta	71.22	81.57	52.09	48.30	94.75	96.10
8	Sauce	72.45	78.45	40.56	55.00	88.78	85.00
9	Chicken	69.81	71.45	28.02	34.27	86.55	84.52
10	egg	45.12	75.71	31.00	48.37	77.53	92.53
11	Cheese	61.67	83.62	42.67	33.65	97.47	93.43
12	Meat	75.38	71.67	55.00	44.61	95.73	97.73
13	Onion	45.81	79.98	31.78	22.59	89.99	84.48
14	Bean	76.80	79.55	76.71	65.11	98.68	96.73
15	Fish	58.55	64.81	18.96	62.73	77.70	81.50
Total Average		**63.76**	**76.28**	**44.88**	**45.90**	**92.21**	**90.41**

and identified food portions using their contour inside of the dish [22]. We then extracted one-by-one each portion and analyzed them using the methods described in this paper. For the SVM part, we used around 3000 different images for our method, which means a set of more than 300 images for each food portions, Approximately 150 for training set and then another 150 images as a testing set. In the experiment, the color, texture, size, and shape properties of the food images were extracted after preprocessing, as shown in the examples of Fig. 9. We then checked the recognition result with features separately, which were color, texture, size, and shape, respectively. In addition, we have evaluated the performance of the system when all of the features are involved in the recognition phase. Furthermore, to test the accuracy of the SVM method, we have applied tenfold cross-validation on different food portions. In cross-validation, the original sample is randomly partitioned into k equal size subsamples. In our model, we have 10 different rotation of our sample, a single subsample is retained as the validation data for testing the model, and the remaining $k-1$ subsamples are used as training data. The cross-validation process is then repeated k times (the folds), with each of the k subsamples used exactly once as the validation data. The k results from the folds then can be averaged to produce a single estimation. The advantage of this method over repeated random subsampling is that all observations are used for both training and validation, and each observation is used for validation exactly once.

B. Evaluation of the Recognition Systems

The results of the above-mentioned evaluations are shown in Table III. As the table shows, we have low accuracy results for each separate feature, whereas, involving joint combination of all features works well with an accuracy of approximately 92.21%. Finally, as shown in the last column of Table III, we have examined the system performance using tenfold cross-validation technique, and we can see that the accuracy of results are acceptable as well. Since in tenfold cross-validation,

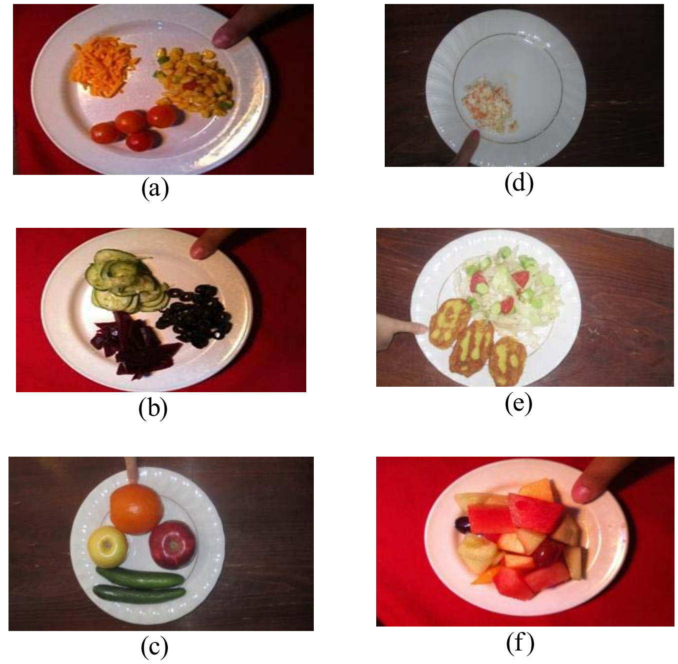

Fig. 10. Nonmixed food (left) and mixed food (right).

TABLE IV
AREA MEASUREMENT EXPERIMENT RESULTS

Food type	*Error percentage*
Bread	0.63%
Cake	2.30%
Spaghetti	-3.07%
Cookies	0.50%
Omelet	10.5%

we divided input data into 10 different groups, in each iteration we have to test the method on the group of images, meaning that the results are for a group of images, not only for one single image. Compared with the tenfold cross method with the previous model in which we have tested the system using only one image in each step and the result is the accuracy of finding one food portion, we may reach lower accuracy in some food portions, which is why the last column of Table III is generally lower than its second last column, with the exception of fish and egg.

Furthermore, we have evaluated the tenfold cross-validation technique with a nonmixed and mixed plate shown in Fig. 10 and reported the results in Table III. We can observe that in nonmixed and mixed plate food, we could not obtain accuracy as high as that of single food portions shown in Table II. This is expected as we have discussed before in this paper.

C. Evaluation of the Area Measurement Technique

We have evaluated our proposed area measurement technique on a variety of simple food (not liquid like soup, curry, and so on). We measured the area of each food portion twice: once by hand from the image, and once using our proposed method. Our experimental results, some of which are shown in Table IV, show that our area measurement method achieves a reasonable error of about 10% in the worst case, and less than 1% in the best case.

TABLE V
ACCURACY OF PROPOSED METHOD IN COMPARISON WITH REAL VALUES

Food Portions	*Weight (grams)*	*Calculated Calorie*	*Real Calorie*	*Absolute Accuracy (%)*
Cake	100	275	250	90
Egg	150	15	17	88
Apple	200	100	114	87
Tomato	150	23	30	76
Cucumber	100	27.5	30	91
Bread	100	21	17	76
Banana	150	140	157	89
Orange	160	98	90	91
Average Accuracy				**86**

D. System Accuracy

To evaluate the accuracy of the proposed method, we have performed two different simulation scenarios. In the first one, our proposed method is applied on several food portions, and their type and volume are extracted. Using the type and volume of each food portion, its mass is extracted using a density table [25]. Using the extracted mass, the calorie of each food portion is derived using Table I. In the second scenario, the real food portion is actually weighted and its real calorie is extracted using the tables. Finally, we have compared the extracted calories from these two scenarios. Some of the results are shown in Table V.

As the table shows, the accuracy of the proposed method in nonmixed food is approximately around 86%. The results are lower than the recognition rate shown in Table III, though not significantly inaccurate.

E. Uncertainty Measurements

One way to increase the confidence in experimental data is to repeat the same measurement many times and to better estimate uncertainties [1] by checking how reproducible the measurements are. When dealing with repeated measurements, there are three important statistical quantities: average (or mean), standard deviation, and standard error. These are summarized in Table VI.

In our system, the following parameters may have effects on the results: illumination, camera angle, and the camera itself.

Illumination is one of the important parameters, which affect the system outcome because illumination directly affects the segmentation algorithm, which in turn affects the rest of the algorithms. To consider, we put the same plate in three different locations with different illuminations and we took pictures. This strategy was repeated for all of the images in our database.

The second effective parameter is the angle of photography; we have chosen three different angles, which are approximately 30°, 90°, and 150° from the plate of food for all pictures. This means that for each plate in three different locations we have also gotten three more pictures from different angles.

Finally, the camera itself will have an effect on the results in terms of its lens, hardware, and software. As such, we used three different cameras for our experiments, consisting of Canon SD1400, iphone 4, and Canon SD1300.

We discussed above that we have selected three different illuminations for our plates, each illumination combined with three different angles, and each angle taken with three different cameras. This means that we have 27 images for each plate of food in various conditions. This gives a good opportunity to measure uncertainties. Since we cannot show the values for each food's 27 different images, in Table VIII we show for each parameter the average values combined with the other two parameters. For example, the column that corresponds to angle at 30° represents the average for all images in all three illuminations and taken with all three cameras when the angle was 30°. As we can observe from the table, the results show that different illuminations with different angles and also different cameras did not change the final results and they are approximately in the same range. Because of this, the standard error is in an acceptable range in each food potion and the overall error percentage is small compared with real calories. All in all this can tell us the method can work well with passable uncertainty in nonmixed plate of food.

TABLE VI
DEFINITION OF STATISTICAL QUANTITIES

Statistic	**What it is**	**Statistical interpretation**	**Symbol**
Average	estimate of the "true" value of the measurement	the central value	x_{ave}
Standard deviation	a measure of the "spread" in the data	You can be reasonably sure that if you repeat the same measurement one more time, that next measurement will be less than one standard deviation away from the average.	s
Standard error	estimate in the uncertainty in the average of the measurements	You can be reasonably sure that if you do the entire experiment again with the same number of repetitions, the average value from the new experiment will be less than one standard error away from the average value from this experiment.	SE

VII. ANALYSIS

We applied our method to three different categories of food: single, nonmixed, and mixed foods, and from the results which are shown in Tables III and VII, we saw that the SVMs accuracy is approximately 92.21%, 85%, and 35%–65%, respectively.

While the above results are encouraging, there are still some limitations with our system, as follows.

1) Our method still has problems in detecting some mixed foods. In the current version of our proposed method, the segmentation step often fails to properly detect various food portions in mixed foods. In addition, illumination of food portions in a mixed food may be changed as they get mixed, making it harder to extract different food portions. Furthermore, the size of food portions in different mixed food are not similar, hence the method fails to segment food portions properly. To solve this problem, we are working on improving the segmentation mechanism to better support mixed food as well, with the following plan for our future work.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE VII
RESULTS OF TENFOLD CROSS-VALIDATION TECHNIQUES ON NONMIXED AND MIXED FOOD

10 fold cross validation		**Accuracy (%)**					
		Non-mixed			**Mixed**		
		a)	**b)**	**c)**	**d)**	**e)**	**f)**
Train classifier on folds: 2 3 4 5 6 7 8 9 10;	Test against fold: 1	85.34	82.25	91.05	65	44.29	35.62
Train classifier on folds: 1 3 4 5 6 7 8 9 10;	Test against fold: 2	79.36	78.24	100.21	65.25	45	33
Train classifier on folds: 1 2 4 5 6 7 8 9 10;	Test against fold: 3	81.66	77.68	95.3	61.49	45	34.82
Train classifier on folds: 1 2 3 5 6 7 8 9 10;	Test against fold: 4	73.92	89.98	75.41	64.5	43.25	32.38
Train classifier on folds: 1 2 3 4 6 7 8 9 10;	Test against fold: 5	89.22	79.81	100.5	66.81	41.75	34
Train classifier on folds: 1 2 3 4 5 7 8 9 10;	Test against fold: 6	81.3	89.89	95.18	60.15	45	34.3
Train classifier on folds: 1 2 3 4 5 6 8 9 10;	Test against fold: 7	89.28	81.56	94.75	65.63	42.8	35.28
Train classifier on folds: 1 2 3 4 5 6 7 9 10;	Test against fold: 8	91.26	91.57	70.19	64.5	44.19	33.19
Train classifier on folds: 1 2 3 4 5 6 7 8 10;	Test against fold:9	85.1	78.45	87.13	65.5	45.21	35.12
Train classifier on folds: 1 2 3 4 5 6 7 8 9;	Test against fold:10	89	81.45	69.01	64.25	45	35.01
Average		**84.54**	**85.34**	**87.9**	**64.30**	**44.14**	**34.27**

TABLE VIII
REPEATED UNCERTAINTY OF MEASUREMENT

Food items	**Real Calories**	**Calories Measured by App**										
		Illumination			**Angle**			**Camera**			**Average**	**Standard Error**
		Location 1	**Location 2**	**Location 3**	**30^o**	**90^o**	**150^o**	**Canon SD1400**	**iphone 4**	**Canon SD1300**		
Red Apple	80	77.39	79.24	79.99	76.81	80.01	79.40	77.46	81.31	78.46	**78.89**	**0.49**
Orange	71	71.23	71.60	70.39	71.31	70.92	71.02	70.92	71.40	71.61	**71.15**	**0.12**
Tomato	30	21.49	22.51	22.30	25.12	28.01	22.93	23.35	23.71	24.66	**23.78**	**0.65**
Carrot	30	29.61	29.01	29.50	30.21	30.39	30.29	29.77	29.41	29.10	**29.69**	**0.16**
Bread	68	66.81	67.12	67.81	68.29	68.99	69.16	70.31	67.52	71.72	**68.63**	**0.53**
Pasta	280	270.14	268.00	259.91	281.56	285.01	279.48	269.10	271.88	259.93	**271.66**	**2.97**
Egg	17	15.63	16.00	15.99	17.32	16.89	16.93	14.59	15.12	15.52	**15.99**	**0.30**
Banana	10	8.50	8.29	8.31	8.45	8.45	8.00	7.90	7.91	7.23	**8.11**	**0.13**
Cucumber	30	27.34	28.01	28.00	28.21	28.00	28.49	27.37	27.61	27.99	**27.89**	**0.12**
Green Pepper	16	18.27	18.21	18.44	18.5	18.5	18.92	18.27	18.5	18.30	**18.43**	**0.07**
Strawberry	53	45.5	46.53	46.12	46.10	45.17	46.13	46.00	47.02	46.38	**46.10**	**0.18**

a) We are going to apply and test other methods such as graph cut segmentation to improve our segmentation steps. Having a more accurate segmentation method helps us to extract more reliable features for recognition phase.
b) We are going to train the system with more mixed foods, to expand the operation range of the system.
c) To increase the accuracy of segmentation, also we are going to increase the range of each feature; e.g., expanding the range of color or texture features.

2) Following plans can be used to improve the measurement of the mass of the food to achieve higher accuracy.
a) Better estimation of the area of each food portion, which can be improved using more accurate segmentation methods, as described in item 1) above.
b) Coming up with an approach to measure the depth of the food more accurately, instead of assuming that the depth is uniform throughout the food portion's area, which is what we assume now.

3) All of our simulations are performed on white plates with a smooth texture. We need to expand our work to various plates with different shapes, textures, and colors as well.

VIII. CONCLUSION

In this paper, we proposed a measurement method that estimates the amount of calories from a food's image by measuring the volume of the food portions from the image and using nutritional facts tables to measure the amount of calorie and nutrition in the food. As we argued, our system is designed to aid dieticians for the treatment of obese or overweight people, although normal people can also benefit from our system by controlling more closely their daily eating without worrying about overeating and weight gain. We focused on

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

identifying food items in an image using image processing and segmentation, food classification using SVM, food portion volume measurement, and calorie measurement based on food portion mass and nutritional tables. Our results indicated reasonable accuracy of our method in area measurement, and subsequently volume and calorie measurement.

An obvious avenue for future work is to cover more food types from a variety of cuisines around the world. In addition, more work is needed for supporting mixed or even liquid food, if possible.

REFERENCES

[1] World Health Organization, Geneva, Switzerland. (2011, Oct.). *Obesity Study* [Online]. Available: http://www.who.int/mediacentre/factsheets/fs311/en/index.html

[2] World Health Organization, Geneva, Switzerland. (2012). *World Health Statistics* [Online]. Available: http://www.who.int/gho/publications/world_health_statistics/2012/en/index.html

[3] G. A. Bray and C. Bouchard, *Handbook of Obesity*, 2nd ed. Baton Rouge, LA, USA: Pennington Biomedical Research Center, 2004.

[4] J. Wenyan, Z. Ruizhen, Y. Ning, J. D. Fernstrom, M. H. Fernstrom, R. J. Sclabassi, *et al.*, "A food portion size measurement system for image-based dietary assessment," in *Proc. IEEE 35th Bioeng. Conf.*, Apr. 2009, pp. 3–5.

[5] R. Almaghrabi, G. Villalobos, P. Pouladzadeh, and S. Shirmohammadi, "A novel method for measuring nutrition intake based on food image," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, Graz, Austria, May 2012, pp. 366–370.

[6] B. Kartikeyan and A. Sarkar, "An identification approach for 2-D autoregressive models in describing textures," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 2, pp. 121–131, 1993.

[7] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, no. 6, pp. 610–621, Nov. 1973.

[8] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 124–128, Jan. 1998.

[9] S. Qin, Y. Wu, and Y. Deng, "An automatic food recognition algorithm with both," *Proc. SPIE*, vol. 7489, pp. 1–8, Jul. 2009.

[10] S. A. Madival and B. C. Vishwanath, "Recognition of fruits in fruits salad based on color and texture features," *Int. J. Eng. Res. Technol.*, vol. 1, no. 7, Sep. 2012.

[11] C.-J. Du and D.-W. Sun, "Pizza sauce spread classification using colour vision and support vector machines," *J. Food Eng.*, vol. 66, no. 2, pp. 137–145, 2005.

[12] M. Livingstone, P. Robson, and J. Wallace, "Issues in dietary intake assessment of children and adolescents," *Brit. J. Nutrition*, vol. 92, pp. S213–S222, Oct. 2004.

[13] (2011, Oct.). *Obesity: A Research Journal* [Online]. Available: http://www.nature.com/oby/journal/v10/n11s/full/oby2002192a.html

[14] S. Mingui, L. Qiang, K. Schmidt, Y. Lei, Y. Ning, J. D. Fernstrom, *et al.*, "Determination of food portion size by image processing," in *Proc. 30th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2008, pp. 871–874.

[15] B. Schölkopf, A. Smola, R. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, May 2000.

[16] L. E. Burke, M. Warziski, T. Starrett, J. Choo, E. Music, S. Sereika, *et al.*, "Self-monitoring dietary intake: Current andfuture practices," *J. Renal Nutrition Off. J. Council Renal Nutrition Nat. Kidney Found.*, vol. 15, no. 3, pp. 281–290, 2005.

[17] J. Beasley, "The pros and cons of using PDAs for dietary self-monitoring," *J. Amer. Dietetic Assoc.*, vol. 107, no. 5, pp. 739–740, 2007.

[18] C. Gao, F. Kong, and J. Tan, "Healthaware: Tackling obesity with health aware smart phone systems," in *Proc. IEEE Int. Conf. Robot. Biometics*, Dec. 2009, pp. 1549–1554.

[19] K. E. Barrett, S. Boitano, S. M. Barman, and H. L. Brooks, "Digestion, absorption & nutritional principles," in *Ganong's Review of Medical Physiology*. New York, NY, USA: McGraw-Hill, 2009, ch. 27.

[20] Health Canada. Province, ON, Canada. (2011, Nov.). *Health Canada Nutrient Values* [Online]. Available: http://www.hc-sc.gc.ca/fn-an/nutrition/fiche-nutri-data/nutrient_value-valeurs_nutritives-tc-tm-eng.php

[21] G. Villalobos, R. Almaghrabi, B. Hariri, and S. Shirmohammadi, "A personal assistive system for nutrient intake monitoring," in *Proc. ACM Workshop Ubiquitous Meta User Inter.*, Scottsdale, AZ, USA, 2011, pp. 17–22.

[22] G. Villalobos, R. Almaghrabi, P. Pouladzadeh, and S. Shirmohammadi, "An image processing approach for calorie intake measurement," in *Proc. IEEE Symp. Med. Meas. Appl.*, Budapest, Hungary, May 2012, pp. 1–5.

[23] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognit.*, vol. 24, no. 12, pp. 1167–1186, 1991.

[24] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl.*, vol. 2, no. 2, pp. 121–167, 1998.

[25] K. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 181–201, Mar. 2001.

[26] W. Wu and J. Yang, "Fast food recognition from videos of eating for calorie estimation," in *Proc. Int. Conf. Multimedia Expo*, Jul. 2009, pp. 1210–1213.

[27] AquaCalc. *Food Volume to Weight Conversions* [Online]. Available: http://www.aqua-calc.com/page/density-table

[28] P. Pouladzadeh, G. Villalobos, R. Almaghrabi, and S. Shirmohammadi, "A novel SVM based food recognition method for calorie measurement applications," in *Proc. IEEE Int. Workshop Interact. Ambient Intell. Multimedia Environ.*, Melbourne, Australia, Jul. 2012, pp. 495–498.

Parisa Pouladzadeh received the M.Sc. degree from the University of Ottawa, Ottawa, ON, Canada, in 2012, where she is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Computer Science.

She is involved in food recognition systems. Her current research interests include image processing, artificial intelligence, and classification.

Mrs. Pouladzadeh was nominated for the Best Thesis Award.

Shervin Shirmohammadi (SM'04) received the Ph.D. degree in electrical engineering from the University of Ottawa, Ottawa, ON, Canada.

He is currently a Full Professor with the School of Electrical Engineering and Computer Science, University of Ottawa. He is the Co-Director of the Distributed and Collaborative Virtual Environment Research Laboratory, and Multimedia Communications Research Laboratory, conducting research in multimedia systems and networking, specifically in gaming systems and virtual environments, video systems, and multimedia-assisted biomedical engineering. He has published more than 200 publications, over 12 patents and technology transfers to the private sector, and a number of awards and prizes.

Dr. Shirmohammadi is an Associate Editor-in-Chief for the IEEE INSTRUMENTATION AND MEASUREMENT MAGAZINE, Associate Editor for the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, Senior Associate Editor for the *ACM Transactions on Multimedia Computing, Communications, and Applications*, and was an Associate Editor for the *Springer's Journal of Multimedia Tools and Applications* from 2004 to 2012. He is a University of Ottawa Gold Medalist, licensed Professional Engineer in Ontario, and Lifetime Professional Member of the ACM.

Rana Al-Maghrabi received the B.Sc. degree in family science from Taibah University, Madinah, Saudi Arabia, and the M.Sc. degree in systems science from the University of Ottawa, Ottawa, ON, Canada, in 2013.

She was a Nutritionist and Systems Analysis. Her current research interests include health measurement systems and applications in health care and the adoption of technology in health and health management.