Smart-U: Smart Utensils Know What You Eat

Qianyi Huang*, Zhice Yang†* and Qian Zhang*
*Department of Computer Science and Engineering, Hong Kong University of Science and Technology
†SIST, ShanghaiTech University
Email:{qhuangaa, qianzh}@ust.hk*, yangzhc@shanghaitech.edu.cn†

***Abstract*—Mobile sensing enables unobtrusive monitoring of our daily activities, sleep quality, breathing and heart rate, revolutionizing the health-care system. Dietary information is also a critical dimension for health management but has no convenient solution yet. In this paper, we ask whether we can track meal composition unobtrusively. We introduce Smart-U, a new utensil design that can recognize meal composition during the intake process, without user intervention or on-body instruments. Smart-U makes use of the fact that light spectra reflected by foods are dependent on the food ingredients. By analyzing the reflected light spectra, Smart-U can recognize what food is on top of the utensil. We describe the prototype design of Smart-U and the food recognition algorithm. We demonstrate that Smart-U can recognize 20 types of foods with 93% accuracy. It can work robustly under different conditions. We envision that Smart-U can enable automatic food intake tracking, and provide personalized food suggestions based on nutrient recommendations and prior consumption. In the long run, Smart-U can contribute to the study of chronic diseases.**

I. Introduction

The past decade has seen a surge of interest in daily activity monitoring. Nowadays, smartphones and smart watches can monitor our exercise [1], [6], sleep quality [27] and even mental states [24]. With recent development in technologies, we envision that health-monitoring techniques can go further to monitor what we are eating. This information can be used to improve our health with self-awareness. With this dietary information, we can answer questions like "Does my child get enough nutrients today?" or "How much exercise is needed to burn the calories I consumed?" Furthermore, if we can monitor food intake over a long term, it could enable doctors and nutritionists to study how chronic diseases are related to dietary habits and provide food suggestions for the population.

In this paper, we ask whether it is possible to track what foods are consumed in an unobtrusive and detailed manner. Unfortunately, existing techniques for monitoring food intake are not suitable for this purpose. Computer vision [23], [33] and on-body [25], [31] devices has been applied for automatic diet monitoring. However, they have limited capabilities in recognizing meal composition. They mainly focus on coarse-grained food recognition, and they cannot provide detailed dietary information, such as distinguishing between whole milk or skimmed milk. Past research has also studied the feasibility of monitoring dietary behavior by instrumented objects and places, such as smart forks [8], [21] and a smart tablecloth [37]. They typically detect eating-related actions (*e.g.*, stirring and cutting), and distinguish general meal types, such as having the main dish, soup or salads, without concrete information on what food is consumed. Thus, there is no convenient solution for dietary tracking yet.

In this paper, we introduce Smart-U, a new method for tracking meal composition. Smart-U are utensils with food recognition capability. Smart-U does not require the users to wear any on-body devices or perform any extra actions. It recognizes what foods are consumed by the users during the intake process. Furthermore, Smart-U can work with many types of utensils, such as spoons, glasses, and dishes. It can recognize both solid and liquid foods.

Smart-U works by analyzing the light spectra reflected by the foods on the utensils. Specifically, Smart-U contains an array of LEDs, including both near-infrared (NIR) bands and visible light bands. Smart-U modulates the LEDs sequentially and captures the spectra reflected by the foods, which depends on the chemical properties of the foods [28]. By looking at the reflected light spectra, Smart-U can infer what foods are placed on the utensils.

However, there are three main challenges in achieving this. The first challenge is that various eating environments, especially the ambient light conditions, would interfere the light spectra we get. We address this challenge by investigating how ambient light would affect foods' light spectra. It turns out that ambient light and food reflected light are combined linearly at the receiver side, *i.e.*, the photodiode. Thus we can cancel out the ambient light interference by taking the ambient light off the total light intensities. The second challenge is to minimize disturbance of LEDs to users' eyes. We tackle this by first using NIR LEDs to detect whether there are foods placed on top of the utensils and only turn on visible light LEDs when foods are covering the LEDs. In this way, LED lights are blocked by the foods and will not penetrate directly into human eyes. The last challenge is how to recognize foods and nutritions. We handle this challenge by designing lighting patterns for LEDs and build machine learning models to predict food category and nutrients.

We build prototypes of Smart-U in the form-factors of a spoon and a glass. We test their performance under various conditions. Experiment results show that Smart-U can reliably detect food presence and recognize 20 types of food with 93% accuracy. Smart-U can work robustly under different temperatures, lighting conditions, and when in motion. Smart-U can also recognize 6 types of drinks with high accuracy. We also take the primarily attempt to predict nutrition information in milk and recognize mixed foods. We believe that Smart-U moves a significant step toward automatic dietary monitoring that enables people to track their meal composition and has a significant impact on our health-care system.

We summarize our contributions as follows:

- We propose a new method, Smart-U, for food recognition. Smart-U integrates an LED array and works by analyzing foods' reflectance profiles. It can recognize meal composition unobtrusively during the intake process and can work with many types of utensils.
- We present the design of Smart-U, including both prototype design and food recognition algorithm. To make Smart-U a user-friendly utensil, we handle ambient light interference and minimize the disturbance to human eyes.
- We build two prototypes of Smart-U, a spoon and a glass, and conduct extensive experiments to test the performance of Smart-U. It can recognize up to 20 types of food with 93% accuracy and can work robustly under various conditions.

The rest of the paper is organized as follows. In Section II, we present the theory of spectroscopy. We describe the prototype design in Section III and algorithm design in Section IV. Implementation details can be found in Section V and experiment results are shown in Section VI. We summarize existing literature in Section VII and discuss some remaining issues in Section VIII. Finally are the conclusion remarks.

II. Theory of Operations: Spectroscopy

In this section, we will briefly introduce spectroscopy, which is the working principle of Smart-U. To make it simple, we will discuss diatomic molecules here. For the more complicated cases of polyatomic molecules, readers may refer to [17].

A diatomic molecule is vibrating as two masses on a spring, as shown in Figure 1. Its potential energy is defined as

$$V = \frac{1}{2}kx^2,$$

where k is the force constant of the bond and x is the displacement from the equilibrium internuclear distance. Its natural frequency is

$$v = \frac{1}{2\pi}\sqrt{\frac{k}{m_r}}, \text{ where } m_r = \frac{m_1 m_2}{m_1 + m_2}.$$

In the example of Figure 1, m_1 and m_2 are the masses of the carbon and hydrogen atoms, respectively. The corresponding energy levels are

$$E_n = (n + \frac{1}{2})\hbar v, n = 0, 1, 2, 3, \cdots,$$

where $\hbar$ is Planck's constant.

Light can be thought of as a stream of photons. The energy of a photon is

$$E_p = \frac{\hbar c}{\lambda},$$

where λ is its wavelength. When a molecule absorbs a photon of light, its energy will escalate to a higher level. Absorption must obey the law of conservation of energy, that is, the increased amount of bond energy must be equal to the energy of the absorbed photon, *i.e.*,

$$E_p = \Delta E_{n_1 \to n_2}.$$

It indicates that a bond will absorb photons of specific wavelengths.

Figure 2 shows the possible energy transitions. The dominant one is the fundamental transition, *i.e.*, $n = 0 \to n = 1$. Transitions from $n = 0 \to n = 2, 3, 4, \cdots$ are called overtones. Other allowed bands such as $n = 1 \to n = 2$, $n = 2 \to n = 3$ are called "hot bands" [17]. For the fundamental transition, light of wavelength $\lambda = \frac{\hbar c}{\Delta E_{0 \to 1}}$ is absorbed. For other transitions, the light of corresponding wavelengths is absorbed. This results in many absorption peaks in the light spectrum. As different foods have different ingredients and thus different chemical bonds, foods can be distinguished by their light spectra.

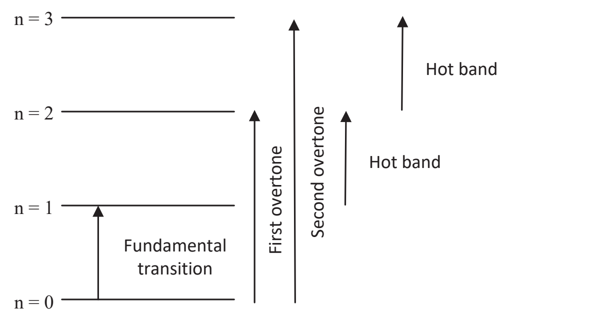


Fig. 2. Vibrational energy levels and possible transitions for a diatomic molecule.

The conventional method to obtain a light spectrum is shown in Figure 3. Light from a light source passes through a focusing and collimating lens, reaches the food samples. We can either observe the transmitted or reflected light. Here we show the latter case. A diffraction grating separates the reflected light into different beams, such that we can obtain the light intensity for each wavelength.

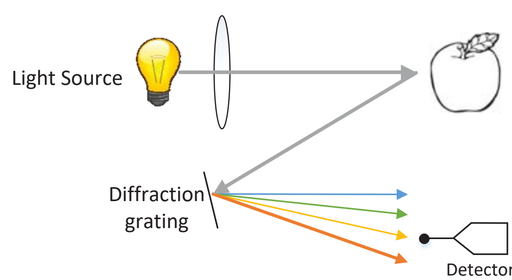


Fig. 3. Conventional spectroscopy system.

The whole system is very bulky. The best case is a handheld food scanner, such as NIRQuest512 from Ocean Optics [10] and TellSpec [13]. In this study, our goal is to miniaturize each component such that it can be integrated into a utensil and consider the various usage scenarios. Details can be found in Section III and Section IV.

Fig. 1. A diatomic molecule is vibrating as two masses on a spring.

III. Smart-U: Prototype Design

In this section, we briefly introduce how we build the Smart-U prototype from two aspects: the light source and the receiver circuit.

A. Light source

As Figure 3 shows, conventionally, we need a light source, which generates light covering a broad spectrum, and a diffraction grating to separate the light into beams. This makes the whole system hard to miniaturize. We notice that we can replace the light source with an LED array. LEDs have several advantages. First, we can control the LEDs such that only one LED is on each time. In this way, we no longer need a diffraction grating, which saves us a lot of space. Second, LEDs are available in a wide range of wavelengths, including the UV, visible and infrared regions. We can select the wavelength that can provide us useful information. Third, LEDs have tiny footprints. A PCB of 2 cm^2 can mount tens of LEDs.

As our goal is to miniaturize the whole system, we need to select LEDs of wavelengths that can provide us critical information. As many commercial NIR spectrometers [10], [13] cover the wavelength ranging from 900 to 1700 nm, we select 8 off-the-shelf LEDs spreading out in this range. In addition, we add 4 LEDs in the visible light bands into the LED array, as we believe that visible bands can also help to distinguish foods. They are red, blue, green, amber, respectively.

B. Receiver

To detect the intensities of reflected light, we use photodiodes as receivers. Photodiodes absorb photons and convert light into electrical current. As photodiodes usually produce a tiny amount of current, in the order of nanoamperes, we need an amplifier circuit to amplify the current so that we can get useful readings. The circuit is shown in Figure 4. Here we use the photodiode in the unbiased mode, as it is more sensitive in the unbiased mode than in the reverse biased mode. The amplifier keeps the voltage across the photodiode zero, *i.e.*, in the unbiased mode. The feedback resistor converts the small photocurrent to a voltage that we can measure at the amplifier's output.

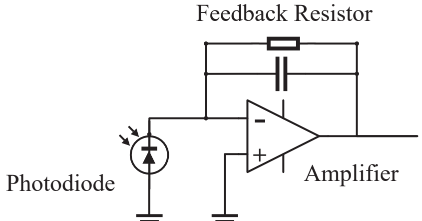


Fig. 4. The amplifier circuit for photodiodes.

As a photodiode covering the full range from 400nm to 1700nm is very expensive (about $120), we use two separate ones, one from 300nm to 1100nm to cover the visible light bands and the other from 800nm to 1700nm to cover the NIR bands.

IV. Smart-U: Recognizing Foods

In this section, we will present how we recognize foods using Smart-U. We first present how Smart-U cancels ambient light interference, and then describe how Smart-U detects food presence. After that, we show how Smart-U recognizes foods and nutrition.

A. Ambient light interference cancellation

When there is food on the utensil, the light intensity received by the photodiodes, denoted by n_T, indeed comes from four sources. First is the ambient light that penetrate through the food and reaches the receiver side, denoted as n_A; second is the direct path between the LEDs and the photodiode, denoted as n_D; third is the LED light reflected by the circuit board, surface of the utensils, and other instances (excluding food), denoted as n_R; last is the LED light reflected by the food, denoted as n_F, which is the information that we intend to get. As the distance between the transmitter (LEDs) and the receiver is fixed, n_D is constant. Similarly, when the geometry of the circuit board and the utensils is fixed, n_R is also constant. However, ambient light could be changing and how it affects foods' light spectra is unknown.

To study how ambient light affects foods' light spectra, we use a Yeelight Bulb from XiaoMi [14] to mimic various lighting conditions. We change the light intensity of bulb from 0 to 100%. We put a slice of beef on the utensil and obtain readings from the receiver. Results are shown in Figure 5. In Figure 5(a) we can see that when there is no LED on, due to the increasing ambient light intensity, we get linearly increasing readings. When the LED is on, the readings increase by an offset. In Figure 5(b), we plot the offsets for three LEDs. We can see that the offsets are almost constant. It indicates that n_F will not change with the ambient light.

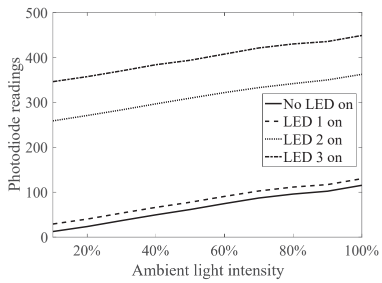
Smart-U cancels out the interference of ambient light by taking n_A off n_T. We can obtain n_A when there is no LED on, and the only light source is the ambient light. As it's hard to obtain n_F directly, we can get

$$n'_F = n_T - n_A, \tag{1}$$

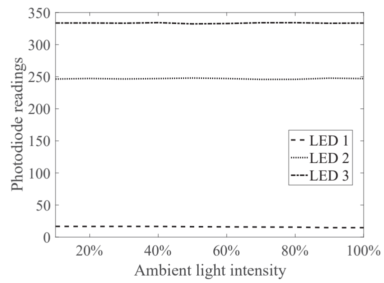
which is a combination of n_F, n_D and n_R, where the last two are constant. In this way, Smart-U cancels out the interference of ambient light.

B. Food detection

The second step is to detect whether there is food on top of the utensil. This serves two purposes. First, the spoon only needs to carry out food recognition procedure when there is food on the utensil so that it can save battery power; Second, as we have 4 visible-light LEDs on the LED array, which may cause disturbance to users' eye. Thus, we only turn on visible-light LEDs when there is food on top of the utensil so that light is blocked by the food and will not go directly into users' eye, causing any unpleasant user experience. We



(a) Raw photodiode readings.



(b) Offset of Photodiode readings.

Fig. 5. We show that ambient light and food reflected light combine linearly at the receiver side.

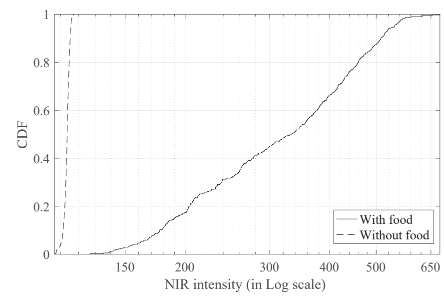


Fig. 6. We compare the NIR intensities with/without food.

use two thresholds for food detection, NIR intensity and time thresholds.

NIR intensity thresholding

We rely on NIR LEDs and photodiode for food detection. We use l to denote the light intensity received by the NIR photodiode. Smart-U turns on a NIR LED and continuously tracks the readings of l. When there is no food on the utensil, little amount of emitted NIR light is reflected back to the receiver. Thus, l has a small value. When there is food on the utensil, a large portion of emitted light is reflected. Thus, l has a large value. We test under three ambient light conditions: indoor environment with no light, indoor environment with fluorescent lights on, and under sunlight. We keep the utensil empty and repeat 200 times under each condition. We also put 20 types of foods on the utensil and conduct the test under the three environments. Figure 6 shows the comparison between with/without food cases.

From Figure 6, we can see that, as we have already canceled out the influence of ambient light, l is rather stable when there is no food. The values fall within a small range. However, when there is food on top of Smart-U, the values of l vary a lot. The value depends on the food ingredients. We further perform t-test on these two sets of data. We get $P < 0.001$, indicating statistically highly significant difference between the two data sets. Thus, we can set a threshold Γ_l to discriminate these two cases.

Time thresholding

Using only a NIR intensity threshold can lead to false positives. When there is no food, according to our observation, there could be short instants that l rises above Γ_l due to hardware imperfection. Thus, we further add another threshold in the time dimension, Γ_t. Only when l is above Γ_l for at least a duration of Γ_t, Smart-U confirms the presence of foods.

C. Recognizing Foods and Nutrition

1) LED Lighting Pattern: On Smart-U, we have 12 LEDs on the LED array. Smart-U turns on the LEDs one by one. For each LED, Smart-U changes both its duty cycle and frequency. Specifically, Smart-U first sets the flicker frequency of an LED and changes its duty cycle (0, 0.25, 0.5, 0.75, 1); then it fixes the duty cycle and changes the flicker frequency (3 frequencies). This lighting pattern is repeated for all the 12 LEDs.

Each combination of flicker frequency and duty cycle lasts for 15 milliseconds. Thus it takes about 1.5 seconds to complete the lighting sequence.

2) Feature extraction: Before performing food classification, we first extract related features.

For each lighting pattern (a combination of flicker frequency and duty cycle), Smart-U computes the average and standard deviation of $[n'_{F,t=1,2,\dots}]$. For one LED, we have 16 features. Thus, there are 192 features in total.

3) Recognition: After extracting food features, we perform food classification.

Recognizing single food

Among many popular classification algorithms, Smart-U builds a Random Forest food recognition model. This is because Random Forest has a lower risk of overfitting by averaging over many trees and is therefore more accurate. The model takes 192 features as input and outputs the food label with the highest probability.

Recognizing nutrition information

After we obtain the label for the food, we proceed to recognize the nutrition information in the food. In this paper, we only consider predicting nutrition information in milk. It would be our future work to generalize to other foods.

We consider three nutrients: protein, fat, and carbohydrate. Given that they are all milk, their light spectra are very similar with only slight differences. We need to emphasize on the variation in spectra. Thus, we use Principal Component Analysis to find the components with large variance. Then we build a regression model for each nutrient, with the principal components as predictors and nutrient content as the response value. Thus, we obtain three regression models, each corresponding to one nutrient.

Recognizing mixed foods

We further consider that the food on the utensil may be a mixture, for example, a mixture of milk and cereal. When doing food classification, we have no idea about how many types of foods there are on the utensil. There may be N types of foods, where N is an unknown number. This is a typical problem of multi-label classification. We select the state-of-the-art multi-label classification algorithm, Random Forest of Predictive Clustering Tree (RF-PCT) [22], as previous studies point out that RF-PCT has high prediction accuracy and

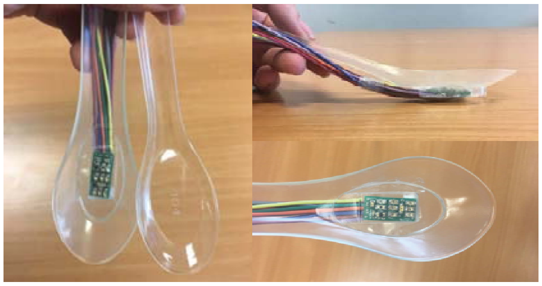


Fig. 7. A spoon example of Smart-U. In the left figure, we compare the size of Smart-U with a normal spoon; right top is the side view and right bottom is the top view.

low complexity [26]. The RF-PCT model will output all the possible labels for the foods.

V. Implementation

In this section, we introduce the implementation details.

3D printing Smart-U prototype: We design the 3D utensil model using the online tool, Tinkercad. The model is printed using transparent photopolymer resin. We place an opening underneath the utensil for holding the circuit board. The opening is 23mm (Long) $\times$12mm (Width) $\times$5mm (Height) in size. The total weight of the spoon model is 8g. Light comes from underneath the utensil and reaches the surface of the food, and then reflected back to the photodiode.

Circuit Design: We fabric the LED array and the receiver circuit on a two-layer print circuit board. The LEDs and photodiodes are on the top layer. The wavelengths for the LEDs are 470nm (blue), 525nm (green), 592nm (amber), 625nm (red), 860nm, 940nm, 1050nm, 1200nm, 1300nm, 1450nm, 1500nm and 1650nm. For the components in Figure 4, we use an AD8505ARJZ-R7 amplifier from Analog Device, a $820k\Omega$ resistor and a $30pF$ capacitor. The two photodiodes are from Luna Optoelectronics. Their part numbers are SD012-151-001 and SD040-101-411CT-ND, respectively.

Control Unit: The LED array is controlled by an Arduino Mega [2]. The default flicker frequency and duty cycle are 122Hz and 50%. We first fix the flicker frequency and change the duty cycle among $0\%, 25\%, 50\%, 75\%$ and 100%. Then we fix the duty cycle and change the flicker frequencies among three values. The default frequency on Arduino Mega is 31250Hz, and we divide the frequency by 1, 64 and 256.

Figure 7 gives photos of our prototype, in the form-factor of a spoon. We also build a glass prototype by attaching the circuit to a transparent Polypropylene glass. Other utensils, such as dish and bowl, can be designed similarly.

VI. Evaluation

In this section, we show the evaluation results. We first show the results for food detection and food recognition, and then evaluate the effects of many factors, including ambient light, temperature, and movement. Last, we show the results for recognizing drinks, nutrients in milk and mixed food. The majority of tests are performed using the spoon prototype.

A. Food Detection Accuracy

We implement the food detection algorithm on an Arduino Mega and test with four types of foods: pork slice, black pepper beef, carrot, and almonds. We put the food on the spoon and see whether Smart-U detects the presence of food. We do the experiments under four lighting conditions: indoor environment with lights off on a rainy day, indoor environment with a fluorescent light on, indoor environment with an LED lamp on, natural sunlight on a sunny day. We repeat 10 times for each type of foods. Results are shown in Figure 8.

We can see that the detection rates are very high. Miss detections only happen when the food is placed at the corner of the spoon. As we can see in Figure 7, some parts of the spoon surface cannot be covered by the receiver on the PCB. When the food is small in size, *e.g.*, almonds, Smart-U may miss detecting its presence. This problem can be solved by adding more receivers on the spoon so that the whole utensil surface can be covered.

We also want to see whether Smart-U would give false positives. We put the empty spoon in each environment for about half an hour. Although there are some instants that l raise above Γ_l, these moments are marked as false alarms by the time threshold. Thus, during a total duration of two hours, there is no false positive. We also swing our hands over Smart-U to see whether it would trigger false alarms. We notice that only when hands are about 1 cm above the spoon, it will give the false alarm. When hands rise to $2-3$ cm above the spoon, Smart-U will not trigger false alarms.

B. Recognizing Foods

In this subsection, we show the accuracy for Smart-U to recognize foods. According to the *Australian Guide to Healthy Eating* [3], we prepare 20 types of foods, covering the five categories of foods in the food selection guide. There are:

1) Rice, muesli, bread, wheat biscuit.
2) Dried pork slice, black pepper sliced beef, satay pork, satay beef.
3) Mango, apple, grapefruit, banana, dragon fruit, carrot.
4) Almond, peanuts, cashew, jujube.
5) Scrambled egg, yogurt.

For each type of food, we prepare 10 food samples. We compare 5 commonly used classification algorithm: Decision tree, Naive Bayes, SVM, k-NN and Random Forest. We vary the percentage of training/testing data and compare their accuracy. The results are shown in Figure 9. From Figure 9 we can see that Random Forest outperforms the other 4 classifiers, indicating that Random Forest is a good choice for food recognition.

We use 50% of data for training and the remaining 50% for testing. We show the confusion matrix in Figure 10. The overall accuracy is 93%. We notice that the majority of errors occur among the same category of foods. Take satay pork and satay beef for example. They are both meat and have been processed using the same method. Thus they are hard to distinguish. The same is also true for peanuts and cashews.

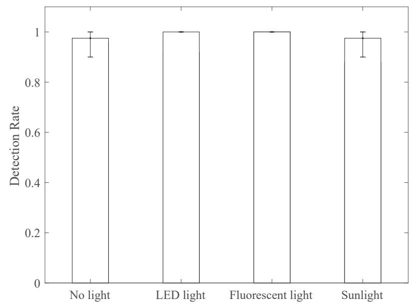


Fig. 8. Food detection rate under different light conditions.

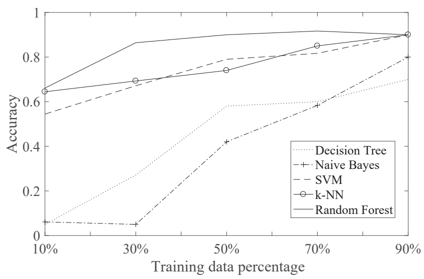


Fig. 9. We vary the percentage of training/testing data and compare the accuracy of 5 classification algorithms. Results show that Random Forest outperforms the other 4 classifiers in all cases.

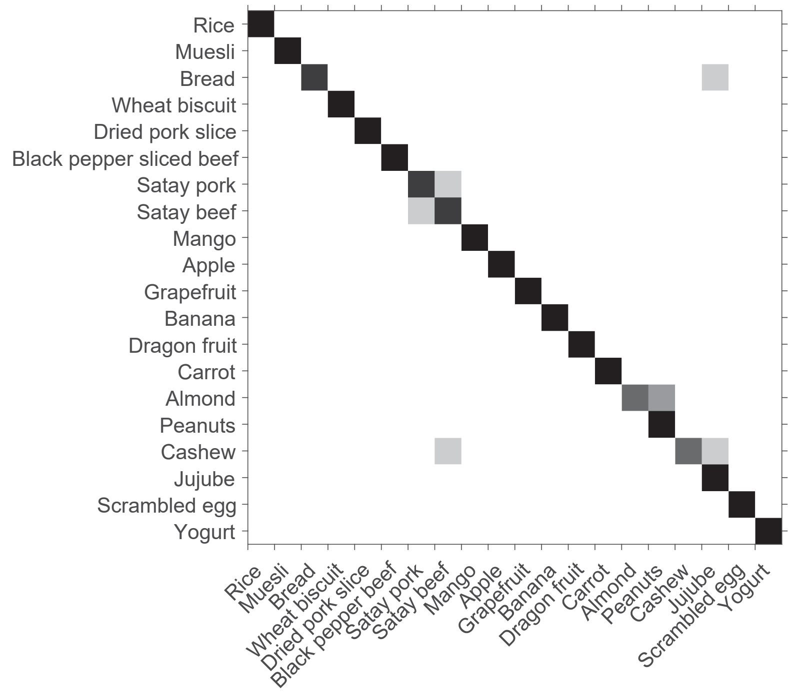


Fig. 10. Confusion matrix for recognizing 20 types of foods. The overall accuracy is 93%.

We show their pictures in Figure 14. From the figure, we can see that they are very similar in appearance.

If Smart-U does not include the four LEDs in visible light bands, the accuracy drops to 67.5%, indicating that visible light bands are beneficial for food classification.

C. Effects of Ambient Light

Here we test the performance of Smart-U under different lighting conditions. The four lighting conditions are the same as in Section VI-A. We use the Random Forest model trained in Section VI-B to classify five types of foods: rice, pork, banana, carrot, and almond. The results are shown in Figure 11. The accuracy under sunlight (0.88) is a little bit lower than the other three cases. This is because, under intense sunlight, some readings from the photodiodes are saturated. This would be solved by using an adjustable feedback resistor in Figure 4. Smart-U could dynamically adjust the amplifier gain according to the intensity of ambient light, *i.e.*, use a large gain to achieve better sensitivity in a dark environment and use a small gain to avoid saturation in a bright environment. The accuracies in the other three conditions are all above 0.92.

D. Effects of Temperature

In this subsection, we show how temperature would affect the performance of Smart-U. We cool foods by putting them in the refrigerator and heat foods using a microwave. According to common practice, we test yogurt, mango and apple by putting them in the fridge and heat rice, sliced pork and sliced beef using a microwave. The results are shown in the first two bars in Figure 12. In these two cases, the accuracies are 0.93 and 0.9, respectively.

The above experiments are conducted in an environment with stable room temperature, ranging from $25°C$ to $29°C$ with the average at $27°C$. We want to see whether Smart-U can work robustly under extreme weather condition. So we put the spoon in the refrigerator to cool it down, mimicking an environment with low room temperature. The results are shown as the third bar in Figure 12. The errors come from the confusion between pork slices and satay pork. They are both pork just with different processing methods, and thus it is difficult to distinguish them. This two sets of experiments show that Smart-U can work robustly under different food and environmental temperatures.

E. Effects of Movement

As Smart-U takes about 1.5 seconds to obtain the light spectra of foods, we want to see whether motion during this period would affect the performance. We hold the spoon and mimic the hand motion of bringing food to the mouth. We test five types of foods: rice, sliced pork, banana, almond, and yogurt. The average accuracy is 0.92. One sample of banana is confused with apple, mainly for that they are both fruits thus share some similarity in ingredients. Another sample of banana is confused with scrambled egg, maybe due to their similarity in color (light yellow).

F. Recognizing Drinks

In this subsection, we show that Smart-U can also recognize drinks. Here we prepare 6 types of drinks: water, tea, coffee,

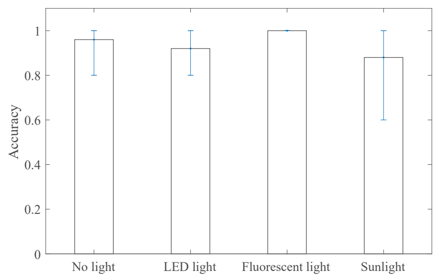


Fig. 11. Accuracy under different lighting conditions.

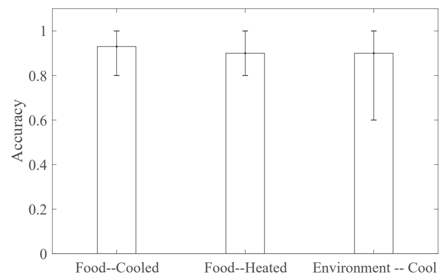


Fig. 12. Accuracy under different food/environment temperatures.

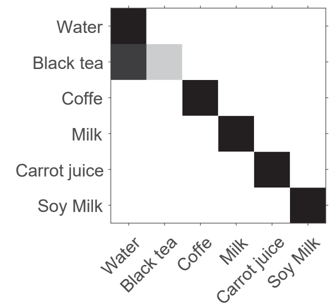


Fig. 13. Confusion matrix for 6 types of drinks.

(a) Satay pork. (b) Satay beef. (c) Peanuts. (d) Cashews.

Fig. 14. Some food samples. Satay pork and satay beef look very similar. The same is true for peanuts and cashews.

milk, carrot juice and soy milk. We perform the test with the glass prototype. We prepare 10 samples (about 25ml) for each drink, 5 samples for training and 5 samples for testing. The results are shown in Figure 13.

4 samples of tea are confused with water. This is because the tea samples are made using a tea bag. The first few samples are strong and the following samples become weak gradually. Thus, the last 4 samples are confused with water. We verify this by preparing 5 new samples of tea, made with a new tea bag. All the 5 new samples are correctly classified as tea. Results show that Smart-U can recognize 6 types of drinks with high accuracy.

G. Predicting Nutrition Information

In this subsection, we show the results for predicting nutrition information in milk. We buy five types of milk from a local supermarket and read their nutrition information from food package. Their information are shown in Table I.

TABLE I
NUTRITION INFORMATION IN 5 TYPES OF MILK. THE NUMBER IS IN GRAMS FOR EVERY 100ML SERVING.

Milk	Protein	Fat	Carbohydrate
Whole milk (Brand 1)	3.2	3.9	5.2
Hi-Calcium Whole Milk (Brand 1)	3.2	3.7	5.1
Hi-Calcium Low Fat Milk (Brand 1)	4.3	1.4	7
Hi-Calcium Chocolate Milk (Brand 1)	2.4	1.4	10.8
Whole milk (Brand 2)	3.26	3.64	4.62

We obtain 10 samples for each type of milk and use 5 samples from the first four types to build the regression model and test on the remaining samples. The predicted results are shown in Figure 15. We can see that the predicted values have a high correlation with the actual values. The deviations from the real values are 2.7%, 8.2% and 1.6% for protein, fat, and carbohydrate, respectively. For the unseen samples (Type 5), the deviations are 9.7%, 19.3%, 3.5%. In this study, we only consider these three nutrients, and it is our future work to examine the concentration of minerals.

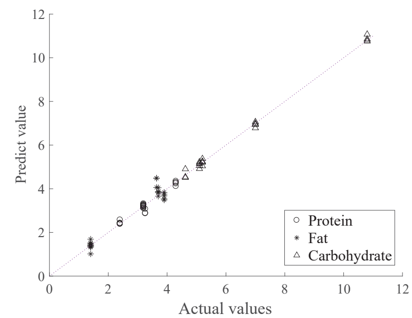


Fig. 15. Predicted results for the three nutrients in milk.

H. Recognizing Mixed Foods

In this subsection, we show the results for identifying mixed foods. We prepare five types of foods: pork slice, black pepper beef, almond, peanut, and cashew. In total we get 9 cases: 5 pure cases, one for each type, and 4 mixed cases: pork & beef mixed, almonds & peanuts mixed, almonds & cashews mixed, and peanuts & cashews mixed. For each combination, we prepare 20 samples. We train an RF-PCT multi-label classifier using CLUS [5] with 50% of data and test on the remaining 50% of data. Table II shows the both the true positive rates and true negative rates. As we show in Figure 14, peanuts and cashews are very similar and thus difficult to distinguish, they are confused in a number of testing cases, leading to low true positive rates and true negative rates for both foods. The

TABLE II
RECOGNIZING MIXED FOODS.

Food	True Positive Rate	False Positive Rate
Pork	0.88	1
Beef	1	1
Almond	0.85	0.95
Peanut	0.77	0.87
Cashew	0.7	0.89
Average	0.84	0.94

accuracy of the multi-label classifier is generally lower than single-label classifier because multi-label classification problem is much more challenging and difficult than single-label classification. The accuracy of Smart-U would be improved by enhancing the hardware design. We can add more receivers on Smart-U, and thus multiple tx-rx pairs can provide more information on the food ingredients.

VII. Related Works

In this section, we review related works.

On-body dietary monitoring devices. A number of sensors are used for dietary monitoring, including acoustic sensors, electromyography (EMG) sensors, force sensors and motion sensors. They come in various form-factors. Acoustic sensors are integrated in earpiece [15], [25] and neckpiece [31], [35]. It works by monitoring the sounds of chewing and swallowing. EMG sensors are built into eyeglasses to measure the muscle activity of temporal muscle during chewing [20], [36]. The activity of temporal muscle can also be measured using force sensor [19]. Motion sensor can detect the eating related motions, including both head motions [25] and arm/wrist motions [16]. Other sensors such as piezoelectric and strain sensors are also used in dietary monitoring. A comprehensive summary of on-body approaches can be found in [34]. Although on-body methods can provide full day monitoring, they usually have low social acceptance. Furthermore, these methods are undesirable among some groups of people. For example, the elderly often feel encumbered by wearable devices.

Computer-vision based approaches. Cameras have been used in the domain of dietary tracking. DietCam [23] can not only recognize food on the plate but also estimate the food volume. It works by taking photos or a short video of the meal before and after eating. In this way, it can estimate how much calories are consumed by the user. eButton [33] also conducts dietary assessment using life-logging camera. There are also some image recognition tools available online, *e.g.*, Foodai [7] and CloudSight [4], which can be used for food recognition. Computer vision based approaches are subject to ambient light conditions and cannot distinguish similar foods, such as whole milk and skimmed milk.

Smart environment and smart objects. We can instrument the environments and daily objects to make them smart. Smart kitchen [18] can track how the food is processed and estimate calorie in the meal. Smart dining table [37] can monitor the actions performed during mealtime (*e.g.*, stir, scoop, cut) and infer how the meal is consumed. There are also some smart utensils. Spün [12] are coming in the form-factor of forks and spoons. When combined with a smartphone camera, it enables calorie and nutrition tracking. HAPIfork [8] vibrates when the user is eating too fast. Sensing Fork [21] uses persuasive technology to improve the eating behavior of children by enabling the interaction between eating and a smartphone game. We note that Smart-U is different from existing smart utensils, as it can provide detailed information on food category, without the help from another device (smartphone).

Near-infrared in food analysis. Near-infrared spectroscopy has been used in food industry for quality assessment [30] and food analysis [29]. Traditional methods require bulky and expensive equipment to do the food analysis, which is not portable and infeasible for daily dietary tracking. Recently, there are several miniaturized handhold food scanners available on the market, *e.g.*, Tellspec [13] and SciO [11]. Although their working principle is similar to Smart-U, they require the user actively scan the food and can only provide information for the parts that are scanned. Nutrilyzer [32] is a portable device for characterizing food with photoacoustic effect, but it works only on liquid food. Smart-U does not require users to perform specific actions during meal time. As it can be integrated into any utensil, *e.g.*, a spoon, it can provide information on every spoon serving.

VIII. Discussion

In this section, we discuss some remaining issues.

Food amount estimation. This paper mainly discuss how to recognize meal composition. Estimating food amount is not in the scope of this study. We note that there are two possible ways that Smart-U can be enhanced to estimate food amount. First, as Smart-U can detect whether there is food on top of the utensil, it will detect interleaving positive and negative results during the meal consumption. In the case of our spoon prototype, Smart-U can infer the number of spoon servings, which is directly related to the amount of food consumed. If Smart-U wants to be more accurate, the second approach it can take is to integrate some additional sensors. For example, it can include 3-axis accelerometer or gyroscope to track the food serving motion [16]. It can also include electrodes to detect when the user put the spoon in his/her mouth [21].

Complex mixture recognition. In our experiments, we only consider cases when two types of foods are mixed. We did not consider the more complicated cases with complex mixtures. To recognize complex mixtures, we can add more receivers, such that multiple tx-rx pairs can provide us more information.

Our evaluation results show that the accuracy of recognizing mixed foods is lower than that of the non-mixed cases. The situation can be improved by better hardware design. The current LED technology enables integrating multiple LEDs in a small chipset. It means that the LED array in Smart-U can be shrunken into a small chip. In this way, we can deploy multiple such chips in Smart-U and each chip can treat the food on top as single food. By combining the outputs from all chips, we can get to know what foods are in the mixture.

Food recognition speed. The total amount of time takes to recognize food includes the time to obtain food light spectra and to perform food classification. Smart-U takes 1.5 seconds to obtain food light spectra. For food recognition, as the algorithm is not implemented on a microcontroller, we do not have the concrete number. But we believe it could be rather short, as testing a sample using a well-trained model is supposed to be fast. As it usually takes about $2-3$ seconds to bring food to mouth, Smart-U has sufficient time to perform food recognition.

Safety and life time of Smart-U. As for utensils, we may be concerned about whether Smart-U is safe. We require the surface of the utensils to be transparent such that light can pass through the surface. Some materials such as Polypropylene, what we used in the glass prototype in Section VI-F, is transparent and is often used for food packaging, even for microwave-proof packaging. Smart-U can be manufactured using these safe materials.

Another concern about Smart-U is that it should be waterproof. The circuit board and the control unit can be sealed inside the utensil. If Smart-U further adopts wireless charging, it does not need to have any open charging port. As some smart utensils can be washed under water, *e.g.*, Liftware Steady [9], HAPIfork [8], Smart-U can adopt the similar techniques to make it waterproof and durable.

IX. Conclusion

In this paper, we have presented the design, implementation, and evaluation of Smart-U, a new method for recognizing meal composition. Smart-U uses an LED array, photodiodes and a control unit to capture the reflected light spectra of foods on the utensil. The unique light spectra enable Smart-U to recognize meal composition. We have implemented a spoon prototype and a glass prototype. We have demonstrated that Smart-U can identify 20 types of foods with 93% accuracy. We conduct tests under different lighting conditions, temperatures and with user movement. Results show that Smart-U can work robustly under different conditions. It can also recognize 6 types of drinks with high accuracy. We also take the primarily attempt to predict nutrients in milk and identify mixed foods. We believe that Smart-U moves a significant step toward automatic dietary monitoring that can provide personalized food suggestions based on nutrient recommendations and prior consumption. In the long run, Smart-U can contribute to the study of chronic diseases.

Acknowledgment

This work was supported in part by the RGC under Contract CERG 16212714, 16203215, Contract ITS/143/16FP-A, and Contract R8015, in part by the Huawei-HKUST Joint Laboratory (by the IoT WiFi Key Technologies). We would like to thank anonymous reviewers for their constructive comments. We also thank Sijie Pan for helping with hardware design.

References

[1] "Apple Health," https://www.apple.com/ios/health/.
[2] "Arduino Mega," https://store.arduino.cc/usa/arduino-mega-2560-rev3.
[3] "Australian Guide to Healthy Eating," https://www.eatforhealth.gov.au/guidelines/australian-guide-healthy-eating.
[4] "Cloudsight," https://cloudsight.ai/.
[5] "Clus," http://clus.sourceforge.net/doku.php?id=home.
[6] "Fitbit," https://www.fitbit.com/home.
[7] "Foodai," http://www.foodai.org/.
[8] "HAPIfork," https://www.hapi.com/product/hapifork.
[9] "Liftware Steady," https://www.liftware.com/intl/us/steady/.
[10] "NIRQuest512," https://oceanoptics.com/product/nirquest512/.
[11] "Scio," https://www.consumerphysics.com/scio-for-consumers/.
[12] "Spün," http://spunutensils.com/.
[13] "Tellspec," http://tellspec.com/.
[14] "Yeelight Smart LED Bulb," https://xiaomi-mi.com/smart-lighting/xiaomi-yeelight-smart-led-bulb-e27/.
[15] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster, "Analysis of chewing sounds for dietary monitoring," in *UbiComp*, Tokyo, Japan, 2005, pp. 56–72.
[16] O. Amft and G. Tröster, "Recognition of dietary activity events using on-body sensors," *Artificial intelligence in medicine*, vol. 42, no. 2, pp. 121–136, 2008.
[17] L. Bokobza, "Near infrared spectroscopy," *Journal of Near Infrared Spectroscopy*, vol. 6, no. 1, pp. 3–17, 1998.
[18] P.-Y. P. Chi, J.-H. Chen, H.-H. Chu, and J.-L. Lo, "Enabling calorie-aware cooking in a smart kitchen," in *Persuasive*, Oulu, Finland, 2008, pp. 116–127.
[19] J. Chung, J. Chung, W. Oh, Y. Yoo, W. G. Lee, and H. Bang, "A glasses-type wearable device for monitoring the patterns of food intake and facial activity," *Scientific Reports*, vol. 7, p. 41690, 2017.
[20] Q. Huang, W. Wang, and Q. Zhang, "Your glasses know your diet: Dietary monitoring using electromyography sensors," *IEEE Internet of Things Journal*, 2017.
[21] A. Kadomura, C.-Y. Li, K. Tsukada, H.-H. Chu, and I. Siio, "Persuasive technology to improve eating behavior using a sensor-embedded fork," in *UbiComp*, Seattle, Washington, 2014, pp. 319–329.
[22] D. Kocev, C. Vens, J. Struyf, and S. Džeroski, "Ensembles of multi-objective decision trees," *Machine Learning: ECML 2007*, pp. 624–631, 2007.
[23] F. Kong and J. Tan, "Dietcam: Automatic dietary assessment with mobile camera phones," *Pervasive and Mobile Computing*, vol. 8, no. 1, pp. 147–163, 2012.
[24] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong, "Moodscope: Building a mood sensor from smartphone usage patterns," in *MobiSys*, Taipei, Taiwan, 2013, pp. 389–402.
[25] M. Mirtchouk, C. Merck, and S. Kleinberg, "Automated estimation of food type and amount consumed from body-worn audio and motion sensors," in *UbiComp*, Heidelberg, Germany, 2016, pp. 451–462.
[26] N.-Y. Nair-Benrekia, P. Kuntz, and F. Meyer, "Learning from multi-label data with interactivity constraints: an extensive experimental study," *Expert Systems with Applications*, vol. 42, no. 13, pp. 5723–5736, 2015.
[27] R. Nandakumar, S. Gollakota, and N. Watson, "Contactless sleep apnea detection on smartphones," in *MobiSys*, Florence, Italy, May 2015, pp. 45–57.
[28] B. G. Osborne, *Near-Infrared Spectroscopy in Food Analysis*. John Wiley & Sons, Ltd, 2006. [Online]. Available: http://dx.doi.org/10.1002/9780470027318.a1018
[29] Y. Ozaki, W. F. McClure, and A. A. Christy, *Near-infrared spectroscopy in food science and technology*. John Wiley & Sons, 2006.
[30] N. Prieto, R. Roehe, P. Lavin, G. Batten, and S. Andres, "Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review," *Meat Science*, vol. 83, no. 2, pp. 175–186, 2009.
[31] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury, "Bodybeat: A mobile system for sensing non-speech body sounds," in *MobiSys*, Bretton Woods, NH, 2014, pp. 2–13.
[32] T. Rahman, A. T. Adams, P. Schein, A. Jain, D. Erickson, and T. Choudhury, "Nutrilyzer: A mobile system for characterizing liquid food with photoacoustic effect." in *SenSys*, 2016, pp. 123–136.
[33] M. Sun, L. E. Burke, Z.-H. Mao, Y. Chen, H.-C. Chen, Y. Bai, Y. Li, C. Li, and W. Jia, "ebutton: a wearable computer for health monitoring and personal assistance," in *Design Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE*. IEEE, 2014, pp. 1–6.
[34] T. Vu, F. Lin, N. Alshurafa, and W. Xu, "Wearable food intake monitoring technologies: A comprehensive review," *Computers*, vol. 6, no. 1, p. 4, 2017.
[35] K. Yatani and K. N. Truong, "Bodyscope: a wearable acoustic sensor for activity recognition," in *UbiComp*, Pittsburgh, PA, 2012, pp. 341–350.
[36] R. Zhang and O. Amft, "Bite glasses: measuring chewing using emg and bone vibration in smart eyeglasses," in *ISWC*. ACM, 2016, pp. 50–52.
[37] B. Zhou, J. Cheng, M. Sundholm, A. Reiss, W. Huang, O. Amft, and P. Lukowicz, "Smart table surface: A novel approach to pervasive dining monitoring," in *PerCom*. IEEE, 2015, pp. 155–162.