Paper (Consumer Services)

Food Recognition via Monitoring Power Leakage from a Microwave Oven

Wei Wei[1,a]   Akihiro Nakamata[1]   Yoshihiro Kawahara[1,2]   Tohru Asami[1]

Abstract: In this paper, we demonstrate a food recognition method by monitoring power leakage from a domestic microwave oven. Universal Software Radio Peripheral (USRP) is applied as a low-cost spectrum analyzer to measure the microwave oven leakage as received signal strength indication (RSSI). We aim to recognize 18 categories of food that are commonly cooked in a microwave oven. By analyzing 180 features that contain the information of heating-time difference, we attain an average recognition accuracy of 82.3%. Using 138 features excluding the heating-time difference information, the average recognition accuracy is 56.2%. The recognition accuracy under different conditions is also investigated, for instance, utilizing different microwave ovens, different distances between the microwave oven and the USRP as well as different data down-sampling rates. Finally, a food recognition application is implemented to demonstrate our method.

Keywords: food recognition, power leakage, microwave oven, USRP, Wi-Fi access point

1. Introduction

Food recognition has been an important topic that researchers in related fields focus on. According to the statistic published by the World Health Organization (WHO), in 2008, more than 1.4 billion adults, in their 20 s and older, were overweight. Over 200 million men and nearly 300 million women were obese [1]. Thus food recognition plays a part in dietary monitoring and logging, helping to control the obesity problem. Methods have been proposed for food recognition by other researchers, such as recognizing food images and monitoring the chewing sound. These methods call for users to submit a picture of the foods or wear a device when having meals. The extra efforts asked from users might introduce usage burden or complexity though we should admit the ubiquity of previous methods. We propose to exploit the features of food-cooking facilities to recognize the foods automatically as a complementary method co-operating with previous methods. The solution we propose is to monitor the feature changing of the power leakage from the microwave oven when different categories of food are cooked within.

The conceptual block diagram is indicated in **Fig. 1**. As we can see from Fig. 1, the source of food recognition in the proposed system is the microwave oven leakage. There are advantages in recognizing foods via the power leakage from the microwave oven. First, most households have a microwave oven, which is usually used to cook foods [2]. We illustrate the usage frequency of different cooking facilities in **Fig. 2**. Second, there is small amount of power leaks from a microwave oven at the frequency
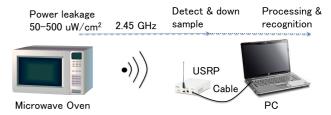

Fig. 1   Conceptual block diagram of proposed food recognition system.
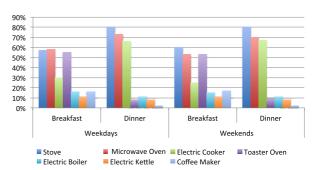

Fig. 2   Usage frequentness of different cooking facilities published by Tokyo Electric Power Company (Tepco).

of about 2.45 GHz when heating foods. The center frequency of leaking power shifts slightly when heating different categories of food [3]. Third, the power leakage is different when different categories of food are being heated inside the microwave oven. We show the absorbed power by food in Eq. (1) [4], showing that material dielectric property determines power absorption efficiency. We denote the frequency with f and the electric field intensity with E. ε_r is the relative permittivity and $tan\delta$ is the loss tangent, two of which change depending on the food materials and cooking stage. To sum up, based on the three advantages that we list, foods can be recognized via detecting the power leakage from the microwave oven as RSSI that varies according to the material

1   Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113–8656, Japan
2   JST PRESTO
a)   weiwei@akg.t.u-tokyo.ac.jp

Table 1   Comparison between food recognition methods. ○ stands for good and × stands for not as good as others.

METHOD	Usage Ubiquity	Information	Automatic	Accuracy	Deployment Cost
Image Processing	○	○	×	○	○
Wearable Device	○	×	×	○	×
Microwave Oven Leakage	×	×	○	○	×

dielectric features of different categories of food.

$$P = 0.556 \times 10^{-12} \times \varepsilon_r \times tan\delta \times f \times E^2 \qquad (1)$$

We clarify the contribution of this paper as follows: In this paper we demonstrate a food recognition system that recognizes foods via monitoring the power leakage from the microwave oven using USRP. This method is a complementary method to the existing food recognition methods using image processing and acoustic sound monitoring. Nevertheless, our system requires less user effort than the former two methods. We make use of the fact that for different kinds of food with different material characteristics (such as moisture content, configuration, and so forth), the food states are different during the heating, which causes the power leakage of the microwave oven to vary. We extract features from measured RSSI to distinguish the power leakage of the microwave oven when heating different kinds of food. In addition, we also illustrate the validity of the proposed system by investigating the recognition accuracy under different recognition conditions, such as different recognition distances, different data down sampling frequencies, using different ovens and foods that are made by different manufacturers.

2.   Related Works

The food recognition for dietary logging and monitoring has been focused on by researchers and two main categories of recognition methods based on two different schemes have been proposed. Solving recognition problems as image categorization or classification problems is the most popular method. The "Foodlog" system based on cell phone camera function was proposed by Kitamura, et al., and according to Refs. [6], [7], [8], the system extracts the features of food color, circle edge and SIFT features from food images taken by a user via a cell phone and uploaded to an online system, attaining an accuracy of 91.8% for food-nonfood recognition and an accuracy of 38.2% of food balance estimator of 5 food categories. In Refs. [9], [10], the authors selected color, texture, gradient, and SIFT features to do training with a separate classifier for each feature. Finally, all the classifiers are weighted combined with the multiple kernel learning method, and the recognition accuracy of 61.3% and 62.5% is achieved for 50 and 85 categories of Japanese foods using 9 and 17 features. And in Ref. [11], the authors utilized the pairwise statistics between local features computed over pixel-level segmentations into eight ingredient types and acquired the recognition accuracy of 28.2% with 61 food categories and 78.0% with 7 food categories. Other food recognition methods using wearable devices to recognize and record the food intake have also been proposed. P. Sebastian et al. proposed a food intake recognition method via investigating acoustics of chewing different kinds of food [12]. Actually, the research on the power leakage of the microwave oven has been conducted on energy harvesting [3] and WLAN network communication quality [13]. In this paper, we explore the usage of microwave oven leakage for food recognition.

We intend to compare our proposed method with the two primary categories of methods above, so as to elucidate the contribution of our proposal. We conduct the comparison through the following parameters:

- Usage Ubiquity. "Usage Ubiquity" means whether the usage of the method is pervasive in daily life or not. Compared to the food recognition methods using image processing and a wearable device, our proposed method can only be utilized in the household environment with a microwave oven and a commodity 2.4 GHz band RF Receiver such as a Wi-Fi access point (router).
- Information. "Information" means that the recognition scheme (image, acoustic sound and microwave oven leakage) contains more or less information corresponding to the food itself. Among three categories of methods, the image processing scheme contains the most information of the food itself.
- Automatic. "Automatic" means that the method can recognize and record food categories automatically, or the method needs much or less user effort. Among all the three methods, the methods using image processing and a wearable device both demand the users to do extra work in order to conduct recognition. Our proposed method can measure the power leakage and conduct the recognition automatically. The users do not need to take pictures or wear extra devices.
- Accuracy. "Accuracy" stands for the recognition accuracy. For all the three categories, different methods proposed by researchers can achieve different accuracy. However, all methods could achieve an accuracy level of about 35% for a large food category number and about 80% for a small food category number.
- Deployment Cost. "Deployment Cost" means the expense or cost of the recognition system or device. Due to the fact that most people have a cell phone equipped with a camera currently, the recognition method using image processing will not impose extra system deployment cost to the users. The recognition method using a wearable device and our proposed method both demand the users to equip themselves with extra devices. However, our proposed method demands the usage of a microwave oven and a Wi-Fi access point (router), which are also commonly used by normal households now.

We summarize the comparison of the results among all the three categories of methods in **Table 1**. We should note that the proposed food recognition method based on the microwave oven leakage is appealing in the aspect of auto-log (low users effort demand) compared to the other two methods. However, due to the limitation of other comparison parameters, we place our proposed

method as a complementary method co-working with the image processing and wearable device based food recognition methods.

3.   Recognition Scheme

In this section, we illustrate the recognition scheme of the proposed method. We first describe the system configuration. Then we list the detailed information of the food categories in our recognition experiment. Finally we will go into data measurement and down sampling before feature extraction. The workflow block diagram of the proposed recognition system is illustrated in **Fig. 3**. In this section, we mainly concentrate on the first three of all five steps.

3.1   System Configuration

In order to investigate how the distance between the microwave oven and the USRP (which we call "recognition distance") affects the recognition result, we take the recognition distance into consideration as one of the parameters. We investigate three recognition distances that are 0.3 m, 5 m and 10 m respectively. We conduct recognition with the distance of 0.3 m because with this distance there is almost no interference introduced by people moving or other electrical devices on the path between the microwave oven and the USRP. For the recognition distances of 5 m and 10 m, we reckon that the recognition device is usually deployed indoors while the normal size of a room is about 5 m and the normal size of a house is about 10 m.

We set up three measurement points at all the three recognition distances simultaneously as shown in **Fig. 4**. And **Fig. 5** shows that one USRP (set 1 in Fig. 4) is deployed 0.3 m from the front door of the microwave oven. In each measurement set, the USRP is connected to the laptop via an ethernet cable. Together they perform data measuring & down sampling, feature extraction and

recognition, as exhibited in **Fig. 6**.

In our recognizing system, the microwave oven we utilize is the NE-EZ2 manufactured by National, a turning-plate microwave oven that is ordinarily available on the market. When the NE-EZ2 microwave oven is heating the food, a leakage signal around the frequency of 2.45 GHz can be observed by a spectrum analyzer. For different kinds of food, the center frequency of the leakage signal will slightly shift. We show the spectrogram of the microwave oven leakage signal measured with the spectrum analyzer RSA3308B-R3 by Tektronix when heating water and French fries in **Fig. 7**.

As for the USRP utilized in our system, we adopt the USRP2 manufactured by Ettus Research with the antenna VERT2450 by the same manufacturer [14]. The software defined radio (SDR) tool GNU Radio is utilized to control the USRP [15]. We briefly describe the working scheme of USRP. After the RF signal is received by the antenna, raw signal (data) is first sampled by the internal A/D converter with a sampling frequency of 100 MHz. Then the signal (data) goes through processing such as downsampling with FPGA and filtering. The processed data finally is transmitted to the PC via an ethernet cable as I/Q signal. In our system, the downsampling rate of the FPGA is set up to 312 kHz.

3.2   Food Category

We select 18 categories of food that is usually sold at grocery stores. The detail information about the 18 categories of food is listed in **Table 2**. The "Time" column in Table 2 stands for the heating-time of each kind of food. We should note that all food categories we select are off-the-shelf products from food

Fig. 6   USRP and laptop connected via an ethernet cable in each measurement set.

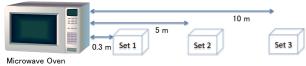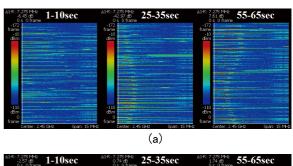**Fig. 3**   The workflow block diagram of the proposed recognition system.

Fig. 4   System deployment of the three measurement sets at different recognition distances.

Fig. 5   USRP deployed 0.3 m away from the front door of microwave oven.

Fig. 7   Spectrogram of the microwave oven leakage when heating (a) water and (b) French fries.

Table 2   Detail information of 18 categories of food.

Food	Brand	Weight (g)	Time (s)
Corn dog	LAWSON	60	40
Cream stew	LAWSON	250	100
Curry sauce	LAWSON	250	100
Dumpling	LAWSON	80	80
French fries	OreIda	100	90
Fried rice	LAWSON	230	200
Gratin	Meji	200	270
Fried chicken	AjiNoMoto	100	80
Okonomiyaki	TableMark	294	240
Spaghetti	Nissin	300	310
Pizza	AQLI	100	90
Porridge	Home-made	250	80
Rice	LAWSON	250	170
Rice ball	Nissui	80	110
Siumai	Nissui	85	90
Taiyaki	LAWSON	92	90
Octopus dumplings	TableMark	100	140
Water	Home-made	100	120

Fig. 8   Raw data measured with the recognition distance of 0.3 m and the down sampling frequency of 2 kHz. Red: pizza. Green: French fries.

manufacturers, which are normally heated in their packet as they are. Because different kinds of food are packed with different net weights, the weights of different food are different in Table 2. For each category of food, we heat ten packages with the same weight and manufacturer. In other words, raw data measurement is repeated 10 times for each kind of food. Thus we utilize 180 data sets to conduct recognition.

3.3   Data Down Sampling

As we show in Fig. 3, after measuring raw data (signal) via the USRP, we proceed data down sampling before we extract features from raw data and conduct recognition.

The measured RSSI data has been down sampled within the USRP with the sampling frequency of 312 kHz before being transmitted to the PC. However, for the RSSI data that is measured for the time length within the range of 40–310 s (according to Table 2), although the raw data have already been down sampled to the frequency of 312 kHz with the USRP, it is still too large for a normal PC to process and recognize. Thus, the signal is down sampled to 312 kHz on USRP and further to the following four frequencies on PC, 500 Hz, 1 kHz, 2 kHz and 5 kHz respectively. We selected these four frequencies in order to investigate how down sampling frequency affects recognition accuracy.

We show the raw data with the recognition distance of 0.3 m and the down sampling frequency of 2 kHz in **Fig. 8**. Three main feature aspects are marked with a number in Fig. 8, which we illustrate in the following list.

(1) Average RSSI level. The average RSSI level of French fries is higher than that of pizza according to Fig. 8. There is a similar average level gap between other different categories of food. We can extract features such as mean, max, min, median, etc. to evaluate such differences between food categories.

(2) Fluctuation. The fluctuation feature such as the amplitude of French fries is higher than that of pizza according to Fig. 8. We can extract other features such as range, standard deviation, etc. to evaluate the difference between food categories.

(3) Turning cycle. We note that the raw data for all 18 categories of food is varied with a time cycle of approximately 12 s. Besides, the 12-second time cycle is the turning cycle
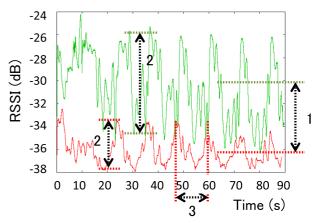
of the turning-plate inside the microwave oven.

To sum up, these three aspects are the main root from which we can draw out more specific features for recognition.

4.   Feature Extraction and Optimization

In this section, we introduce the features we extract to conduct recognition. We have summarized three main aspects of features from the raw data of different kinds of food in the previous section. We first extract specific features from the three aspects above. Then we conduct feature optimization via evaluating the importance of each feature and the relationship between recognition accuracy and the amount of adopted features. The content of this section includes the last two steps in Fig. 3.

4.1   Feature Extraction

In order to make use of the first two feature aspects, which are average RSSI level and fluctuation, we select 46 features as we demonstrate in **Table 3**. The $x_1, x_2...x_n$ stands for the value of raw data at each sampling point. And the $y_2, y_3...y_n$ stands for the step difference of the x array. For instance, y_2 equals to $x_2 - x_1$. Furthermore, we exploit the third feature aspect, which is a 12-second cycle of raw data. As we can see from Fig. 7, the spectrogram of microwave leakage varies for different foods along with the heating time. In Fig. 7, we show the spectrogram of water and French fries during different heating-time slots. Considering the 12-second turning cycle (for all 18 kinds of food) of the turning plate in the microwave oven, we make use of this common turning cycle of all kinds of food (12 s) to divide the time-varying raw data into data frames with the time length of 12 s. Considering the heating-time length in Table 2 (the heating-time of corn dog is the shortest, which is 40 s), we utilize the first three data frames (time: 1–12 s, 13–24 s, 25–36 s) for all 18 kinds of food as shown in **Fig. 9**. We extract features in Table 3 from the all-time-length data, the first frame raw data (1–12 s), the second frame raw data (13–24 s) and the third frame raw data (25–36 s) (thus we extract totally 184 features = 46 features × 4) and conduct recognition.

We utilize all features in Table 3 to the raw data under all recognition conditions (recognition distances and down sampling frequencies). Machine learning software WEKA (Waikato Environment for Knowledge Analysis) is applied to conduct recogni-

Table 3   46 features for recognition.

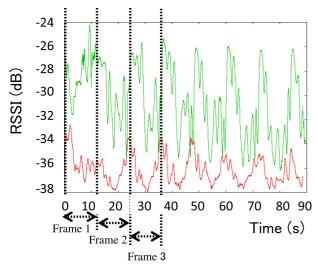| no.	Feature Name	Detail	no.	Feature Name	Detail
1	average	$\dfrac{x_1 + x_2 + x_3 + ... + x_n}{n}$	2	max most frequent value	max value among the most frequent values
3	range	$\text{Max}(x_1...x_n)\text{-Min}(x_1...x_n)$	4	min most frequent value	min value among the most frequent values
5	skewness	skewness of $x_1, x_2...x_n$	6	kurtosis	kurtosis of $x_1, x_2...x_n$
7	mean deviation	mean deviation of $x_1, x_2...x_n$	8	standard deviation	standard deviation of $x_1, x_2...x_n$
9	maximum	Max value among $x_1, x_2...x_n$	10	minimum	Min value among $x_1, x_2...x_n$
11	median	Median value among $x_1, x_2...x_n$	12	root mean square	$\sqrt{\dfrac{\sum x^2}{n}}$ (x: $x_1, x_2...x_n$)
13	coefficient of variation	coefficient of variation of $x_1, x_2...x_n$	14–18	auto-covariance	0.05 s, 0.1 s, 0.5 s, 1.0 s, 2.0 s shift auto-covariance
19–23	auto-correlation	0.05 s, 0.1 s, 0.5 s, 1.0 s, 2.0 s shift auto-correlation	24–46	all 1–23 features for step difference	change x_n to y_n ($= x_n - x_{n-1}$) in all calculation

Fig. 9   First three frames of raw data with the recognition distance of 0.3 m and the down sampling frequency of 2 kHz. Red: pizza. Green: French fries.

Table 4   Recognition accuracy of 18 categories of food using the all-time-length data and the first three frames of raw data (totally 184 features) with different recognition distances and down sampling frequencies.

Distance vs Frequency	500 Hz	1 kHz	2 kHz	5 kHz
0.3 m	80.6%	81.7%	80.0%	83.9%
5 m	80.6%	80.6%	84.4%	84.4%
10 m	79.4%	81.7%	85.6%	84.4%

lower down sampling frequency, more information is lost from the original raw data during down sampling. Thus the recognition accuracy is lower than that of high down sampling frequency. Second, for some features extracted from all-time-length raw data, the heating time length information is contained within such features. We take Feature 24 in Table 3 for example, which is the average of step difference. For all-time-length raw data, the Feature 24 is calculated as shown in Eq. (2).

$$\frac{(x_2 - x_1) + (x_3 - x_2) + ...(x_n - x_{n-1})}{n} = \frac{x_n - x_1}{n} \qquad (2)$$

As we can see from Eq. (2), the heating-time length (proportional to n with certain down sampling frequency) determines the Feature 24 because the difference of the $x_n - x_1$ among all 18 kinds of food is negligible compared to the difference of n among different varieties of food. Because of the features such as the average of step difference which are evoked from all-time-length raw data (the heating-time duration of different foods is mostly different according to Table 2), the recognition accuracy remains while we increase the recognition distance. Such features enhance the robustness of the proposed recognition scheme against the effect of recognition distance.

However, heating-time length information might not be suitable to be used as features in some other application scenarios. For instance, users do not care about the detail of heating time length recommended by the manufacturer of the frozen food product, or users heat home-made foods instead of the off-the-shelf products by food manufacturers. In such cases, the heating-time length is not fixed for certain kinds of food. As well, using the analog turning-button to set up heating-time instead of the using digital press-button also makes the heating-time for the same food vary. We should note that even for different kinds of food, the heating-time can still be the same not only because of the user's personal intention but also the manufacturers' recommendation. In all the cases we have listed above, heating-time length information is not suitable to be used as recognition feature. It

tion [16], [17]. We select Attribute Selected Classifier combined with Simple Logistic to conduct recognition. We also utilize Rank Search as the search method to acquire the importance ranking of all features. 10-fold cross-validation is used to evaluate the feature data. The recognition accuracy is specified as the percent of correctly classified sample numbers out of all 180 samples (data sets) utilized for recognition. We present the recognition accuracy result using the feature extraction above (184 features) in **Table 4**. We show the confusion matrix with 5-meter recognition distance and 2-kHz down sampling frequency (recognition accuracy of 84.4%) in **Table 5**. We should mention the following findings from Table 4:

- The recognition accuracy shows an increasing trend with the same recognition distance as we increase the data down sampling frequency.
- With the same down sampling frequency, the recognition accuracy does not decrease while the recognition distance increases, but maintains at the same level.
- The average recognition accuracy under all recognition conditions (recognition distance & down sampling frequency) is 82.3%, which is comparable with other related work. We show the comparison results between our proposed method and other existing works in **Table 6**.

For the findings above, we clarify the facts that: first, with

Table 5   Confusion matrix with all 184 features, 5-meter recognition distance and 2-kHz down sampling frequency.

All values in %	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
a: Corn dog	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b: Cream stew	0	60	30	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0
c: Curry sauce	0	30	60	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
d: Dumpling	0	0	0	80	10	0	0	10	0	0	0	0	0	0	0	0	0	0
e: French fries	0	0	0	0	80	0	0	0	0	0	0	0	0	0	10	10	0	0
f: Fried rice	0	0	0	0	0	90	0	0	0	0	0	0	10	0	0	0	0	0
g: Gratin	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
h: Fried chicken	0	0	0	10	0	0	0	50	0	0	0	0	30	0	0	10	0	0
i: Okonomiyaki	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
j: Spaghetti	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
k: Pizza	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
l: Porridge	0	0	0	0	0	0	0	20	0	0	0	80	0	0	0	0	0	0
m: Rice	0	0	0	0	0	10	0	0	0	0	0	0	90	0	0	0	0	0
n: Rice ball	0	0	0	0	0	0	0	0	0	0	0	0	0	90	0	0	10	0
o: Siumai	0	0	0	10	0	0	0	10	0	0	0	0	0	0	80	0	0	0
p: Taiyaki	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	90	0	0
q: Octopus dumplings	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	90	0
r: Water	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	10	80

Table 6   Recognition accuracies of different methods.

Methods	Food Category	Accuracy
Proposed	18	82.3%
[9]	50	61.3%
[10]	85	62.5%
[11]	7	78.0%

Table 7   Recognition accuracy of 18 categories of food using the first three frames of raw data (totally 138 features) with different recognition distances and down sampling frequencies.

Distance vs Frequency	500 Hz	1 kHz	2 kHz	5 kHz
0.3 m	62.2%	57.8%	52.2%	58.3%
5 m	52.8%	51.1%	60.0%	54.4%
10 m	52.8%	55.6%	55.6%	61.1%

has been shown in Table 5 that cream stew and curry sauce is partially mixed because they are heated with the same heating-time length (100 s) according to Table 2.

In order to exclude the impact of the different heating-time duration of different kinds of food, we conduct recognition using 46 features in Table 3 extracted from only the first three frames as pictured in Fig. 9. For each of the three frames of any food category in Table 2, the heating time length is the same, which is 12 s. Therefore we utilize in total 138 features extracted from the three frames of raw data. The recognition accuracy with different recognition distances and down sampling frequencies is shown in **Table 7**. The average recognition accuracy with different recognition distances and down sampling frequencies is about 56.2%. As we have excluded the effect of different heating-time length during feature selection, the results in Table 7 also show that the distance increasing from 0.3 m to 10 m does not impose a negative impact on the recognition accuracy of the proposed scheme. We read the confusion matrix with 5-meter recognition distance and 2-kHz down sampling frequency (recognition accuracy of 60.0%) in **Table 8**.

4.2   Feature Optimization

Utilizing the features that include (when using 184 features) or exclude (when using 138 features) the difference of the heating-time duration of different kinds of food, we acquire the recognition accuracy as presented in Table 4 and Table 7. We now con-

centrate on the importance of each feature and the relationship between the feature amount and recognition accuracy.

As for the case utilizing 184 features, we show the five most important features with different recognition distances and down sampling frequencies in **Table 9**. The feature number corresponds to the number in Table 3. The feature number with the suffix [0-12] stands for the feature extracted from the first raw data frame while the suffix [12-24] and [24-36] stand for the features of the second and third frames respectively. The feature number with no suffix stands for the feature of all-time-length raw data. The importance rank is according to the Rank Search method of WEKA.

As shown in Table 9, we can observe that all top 5 features for different recognition distances and down sampling frequencies are 1) the features extracted from all-time-length raw data and 2) step difference features that include heating-time length information. We can deduce that the features related to the heating-time length difference of different kinds of food are the most robust features for recognition while all 184 features are utilized. The more directly the feature is determined by heating-time length difference, the more important the feature is for recognition.

We select the recognition condition of 5-meter recognition distance and 2-kHz down sampling frequency to investigate the relationship between the feature amount and recognition accuracy as demonstrated in **Fig. 10**. The reason why we select 5-meter distance is that this distance is the most similar to the real size of a normal room of people's homes among all three recognition distances (0.3 m, 5 m and 10 m). And with the 2-kHz down sampling frequency we acquired the highest recognition accuracy at 5-meter distance with a smaller amount of data, compared to 5-kHz down sampling frequency according to Table 4. As shown in Fig. 10, with the top ten features, which contains the information of the heating-time length difference among different foods, the recognition accuracy increases from 84.44% to 88.30%. This result shows that the total heating-time difference among different kinds of food is decisive if we can use this difference as the feature for recognition.

We also investigate the top 5 features while 138 features are utilized for recognition with different recognition distances and

Table 8 Confusion matrix with all 138 features, 5-meter recognition distance and 2-kHz down sampling frequency.

All values in %	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
a: Corn dog	80	0	0	0	0	0	0	0	0	0	0	0	0	10	0	10	0	0
b: Cream stew	0	0	30	0	0	0	10	0	0	0	0	40	0	0	0	0	10	10
c: Curry sauce	0	0	30	0	0	10	0	0	10	0	10	10	10	0	10	0	0	10
d: Dumpling	0	0	0	80	10	0	0	0	0	0	0	0	0	0	0	0	10	0
e: French fries	0	10	0	0	70	0	0	0	0	0	0	0	10	0	0	0	0	10
f: Fried rice	0	0	0	0	10	60	10	0	0	10	10	0	0	0	0	0	0	0
g: Gratin	0	0	0	0	0	0	90	0	0	0	0	10	0	0	0	0	0	0
h: Fried chicken	10	0	0	0	0	0	0	50	0	0	0	10	10	0	10	0	0	10
i: Okonomiyaki	0	0	0	0	0	0	0	0	80	20	0	0	0	0	0	0	0	0
j: Spaghetti	0	0	10	0	0	10	10	0	20	50	0	0	0	0	0	0	0	0
k: Pizza	0	0	0	0	0	0	0	0	0	20	70	10	0	0	0	0	0	0
l: Porridge	0	30	0	0	0	10	0	0	0	0	0	40	20	0	0	0	0	0
m: Rice	0	0	0	0	10	0	0	20	0	0	0	30	30	0	10	0	0	0
n: Rice ball	10	0	0	0	0	0	0	0	0	0	0	0	0	90	0	0	0	0
o: Siumai	0	10	0	0	0	0	10	20	0	0	0	10	0	0	50	0	0	0
p: Taiyaki	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	80	10	0
q: Octopus dumplings	0	0	10	0	10	0	0	0	0	0	0	10	10	0	0	0	50	10
r: Water	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	10	80

Table 9 Top 5 features among 184 features with different recognition distances and down sampling frequencies.

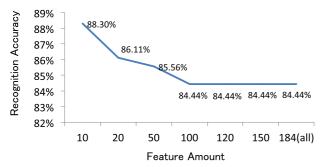| D vs F	500 Hz	1 kHz	2 kHz	5 kHz
0.3 m	24, 30, 34 39, 36	24, 30, 34 36, 11	24, 30, 34 29, 28	24, 30, 34 29, 28
5 m	24, 30, 36 34, 29	24, 30, 36 34, 29	24, 30, 28 29, 36	24, 30, 29 28, 36
10 m	24, 30, 36 35, 31	24, 30, 29 28, 36	24, 30, 28 29, 36	24, 30, 29 28, 36

Fig. 10 The relationship between the top feature amount and recognition accuracy for the case of 184 features with 5 m recognition distance and 2 kHz down sampling frequency.

Table 10 Top 5 features among 138 features with different recognition distances and down sampling frequencies.

D vs F	500 Hz	1 kHz	2 kHz	5 kHz
0.3 m	1 [0-12] 11 [0-12] 12 [0-12] 12 [24-36] 11 [24-36]	9 [24-36] 11 [24-36] 1 [0-12] 12 [0-12] 11 [0-12]	9 [24-36] 3 [24-36] 11 [24-36] 1 [0-12] 24 [24-36]	1 [0-12] 11 [24-36] 12 [0-12] 4 [0-12] 11 [0-12]
5 m	1 [24-36] 12 [24-36] 11 [0-12] 17 [0-12] 12 [0-12]	34 [24-36] 11 [0-12] 1 [24-36] 34 [0-12] 12 [24-36]	34 [24-36] 18 [24-36] 11 [24-36] 34 [0-12] 1 [24-36]	18 [24-36] 11 [24-36] 1 [24-36] 34 [0-12] 34 [24-36]
10 m	1 [24-36] 19 [0-12] 30 [0-12] 34 [12-24] 12 [12-24]	1 [24-36] 34 [12-24] 12 [12-24] 1 [12-24] 8 [24-36]	1 [24-36] 7 [12-24] 17 [12-24] 11 [12-24] 18 [24-36]	1 [24-36] 18 [24-36] 17 [12-24] 12 [12-24] 16 [24-36]

Fig. 11 The relationship between the top feature amount and recognition accuracy for the case of 138 features with 5-m recognition distance and 2-kHz down sampling frequency.

5 important features are among Features 1–23, which is different with the results in Table 9. This result shows that without the all-time-length raw data, Features 24–46 do not contain the difference of the heating-time length among different kinds of food anymore, which makes them not as important as in Table 9. The features as average (Feature 1), median (Feature 11), etc. become the most important features according to Table 10.

Again we look into the relationship between the top feature amount and recognition accuracy for the condition with in total 138 features, 5-meter recognition distance and 2-kHz down sampling frequency. The solution is depicted in **Fig. 11**.

5. Discussion

In this section, we discuss the ubiquity of the proposed food recognition scheme. We expand our previous experiment to the situation that we utilize other microwave ovens or foods that are produced by other manufacturers. In addition, we also conduct discussion on recognizing foods with different weights and replacing USRP in our scheme with Wi-Fi access point so as to expand the utility range of our system to daily household usage.

5.1 Using Other Microwave Oven

As we mentioned, we utilized the microwave oven, National NE-EZ2, to conduct previous data measurement and food recognition. We also utilized another microwave oven (Panasonic NE-

down sampling frequencies as indicated in **Table 10**. The feature number is the same with Table 3 and the suffix is also the same as we illustrated. As we can see from Table 10, most of the top
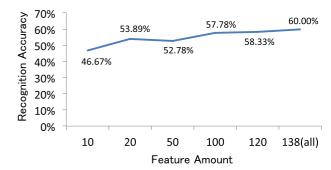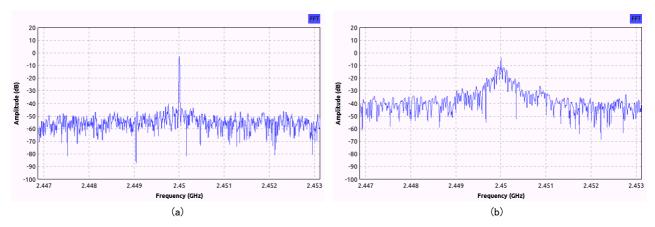
Fig. 12   Spectrum of two microwave ovens. (a) National NE-EZ2. (b) Panasonic NE-EH225.

EH225) to conduct data measurement, processing and recognition with a recognition distance of 5 m and down sampling frequency of 5 kHz. As for the difference between National NE-EZ2 and Panasonic NE-EH225, we show the spectrogram of two microwave ovens working in **Fig. 12**, which is measured by the spectrum-analyzing function of USRP at a distance of 30 cm from the door of the microwave oven. As we can see from Fig. 12, the spectrum range of the leakage from the NE-EH225 is wider than that of the NE-EZ2 while the microwave oven is working.

We select five kinds of food from Table 2 (curry sauce, dumpling, pizza, rice and water) and use NE-EH225 to heat them with 5 m recognition distance and 2 kHz down sampling frequency. For each food, we collect five sets of new raw data (by NE-EH225) and replace five sets of original raw data (by NE-EZ2) out of the total ten sets. We conduct recognition with the same classifier of WEKA and 184 features, acquiring the recognition accuracy of 83.3%. With 138 features we achieve a recognition accuracy of 60.6%. As for utilizing the proposed recognition scheme with other microwave ovens, we suggest users provide the learning data after using the new oven for a certain period of time (a few months or so). With the new learning data, the proposed scheme can work with the new oven.

5.2   Using Foods by Other Manufacturers

As well as microwave ovens of another pattern, we also recognize food that is produced by other manufacturers. We chose 5 categories of food within the original 18 categories and make the substitution. We list the five categories of food in **Table 11**. Similar to recognition with another microwave oven, we replace five sets of original raw data of these five kinds of food (by manufacturers in Table 2) with the new data (by manufacturers in Table 11). The recognition distance is 5 m and down sampling frequency is 2 kHz. We first utilize 184 features that contain different heating-time length information of different kinds of food and acquire a recognition accuracy of 72.8%. When we use 138 features that exclude the difference of heating-time length, we acquire a recognition accuracy of 56.7%.

5.3   Using Foods of Different Weights

For a certain category of food, people may heat it with different weights or quantities. We can expect two main differences

Table 11   Five categories of food that are produced by other manufacturers.

no.	Food	Brand
1	French fries	LAWSON
2	Fried rice	LAWSON
3	Siumai	AjiNoMoto
4	Octopus dumpling	LAWSON
5	Fried chicken	LAWSON

from our previous experiment. First, the heating-time length will change for a certain kind of food. Second, the general RSSI level will vary because the power absorbed by the food is proportional to the food quantity. At this time, the heating-time difference should not be utilized as a feature to recognize food categories but to recognize the food quantity with the food category already known. Also the features containing the information of the general RSSI level should be utilized in the same way. Thus we can recognize the food weight in two steps, which will be one of the main focuses in our future work.

(1)   Utilize the features that do not contain the information of the heating-time difference or the general RSSI level to recognize the food category.

(2)   Utilize the features that contain the information of the heating-time difference or the general RSSI level to recognize the food weight.

Although our system cannot recognize the nutrition intake currently, we should note that our system can help analyzing the users' dietary habit or tendency (favorite foods and meal time), which not only can aid people to defeat the obesity problem as an auto food logging or monitoring system but also can provide customers' preference information for the food manufacturers.

5.4   Replacing USRP with Wi-Fi Access Point

In our current food recognition scheme, USRP is utilized as a low-cost spectrum analyzer to monitor the time-varying power leakage from the microwave oven. According to the RSSI measurement result, we can conclude that the leakage strength level differs between different kinds of food. In addition, the power leakage from the microwave oven @ 2.45 GHz interferes with the WLAN communication at the same frequency, taking Wi-Fi as an example. Thus we can infer that different leakage strength by heating different kinds of food causes varying degrees of interference to the Wi-Fi communication, which can be detected

by a Wi-Fi access point. As a result, the USRP in the recognition system can be replaced by a normal Wi-Fi access point. The interference features detected by the Wi-Fi access point can be utilized to recognize food categories. This will reduce the deployment cost of our system because almost all households have a Wi-Fi access point such as a Wi-Fi router at the present time. And this replacement will also be one of the focuses in our future work.

6. Implementation

We now illustrate the implemented demo of our proposed recognition scheme. We have taken the video of the demo (uploaded as a supporting document) showing how the demo works and successfully recognizes three kinds of food in Table 2 as examples, which are water, French fries and pizza. In the demo system, the USRP is deployed 5 m away from the microwave oven. When the user switches on the recognition system, the system starts working and keeps monitoring the RSSI @ 2.45 GHz. When the microwave oven starts to heat the food, the system detects the RSSI increase caused by the power leakage of the mi-
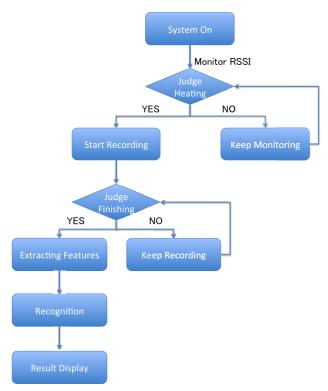
crowave oven and records the time-varying RSSI data until the RSSI level returns to normal value (after the heating stops). Then the system extracts 184 features from recorded RSSI data and utilizes the features as test data for recognition. The training data that we use are the features of 18 categories of food (ten sets data for each food), which has been stored in the laptop beforehand. We show the working flow block diagram in **Fig. 13**. We have also designed the application UI for users to control the recognition system and showing the recognition results as shown in **Fig. 14**.

7. Conclusion

In this paper, we proposed a food recognizing system via monitoring the power leakage from the microwave oven using the universal software radio peripheral. This system exploits the difference of the power leakage from the microwave oven caused by heating different kinds of food to conduct recognition. 18 categories of food have been recognized with an average recognition accuracy of 82.3% using 184 features that contain the information of heating-time difference of different kinds of food, while the average recognition accuracy is 56.2% using 138 features excluding the information of the heating-time difference among food categories. The parameters such as recognition distance (between the USRP and the microwave oven) and data down sampling frequency have also been investigated by conducting recognition with the combination of three recognition distances (0.3 m, 5 m, 10 m) and four down sampling frequencies (500 Hz, 1 kHz, 2 kHz and 5 kHz). In order to expand the ubiquity of the proposed system, we also illustrated the performance of the system using different patterns of microwave oven and the foods within the 18 categories but produced by other manufacturers. Finally, we implemented a demo system, including a control program and user interface to demonstrate our work.

Fig. 13 Working flow block diagram of the food recognition demo system.

Fig. 14 User interface of demo recognition system. (a) switching on/off the system. (b) recognition result that shows the user just heated pizza including the food name and time.

References

[1] available from ⟨http://www.who.int/mediacentre/factsheets/fs311/en/⟩.
[2] available from ⟨http://www.tepco.co.jp/cc/press/⟩.
[3] Kawahara, Y., Bian, X., Shigeta, R., Narusue, Y., Rushi V., Tentzeris, M. and Asami, T.: Power Harvesting from Microwave Oven Electromagnetic Leakage, *Ubicomp'13*, Zurich, Switzerland (2013).
[4] available from ⟨http://www.microdenshi.co.jp/microwave/⟩.
[5] available from ⟨https://www.ettus.com/product/details/UN210-KIT⟩.
[6] Kitamura, K., Yamasaki, T. and Aizawa, K.: Food log by analyzing food images, *ACM MM*, pp.999–1000, ACM (2008).
[7] Kitamura, K., De Silva, C., Yamasaki, T. and Aizawa, K.: Image processing based approach to food balance analysis for personal food logging, *IEEE ICME*, pp.625–630 (2010).
[8] Aizawa, K., De Silva, C., Ogawa, M. and Sato, Y.: Food balance estimation by using personal dietary tendencies in a multimedia food log, *IEEE Trans. Multimedia*, Vol.15, Issue 8, pp.2176–2185 (2013).
[9] Joutou, T. and Yanai, K.: A food image recognition system with multiple kernel learning, *IEEE ICIP*, pp.285–288 (2009).
[10] Hoashi, H., Joutou, T. and Yanai, K.: Image recognition of 85 food categories by feature fusion, *IEEE ISM*, pp.296–301 (2010).
[11] Yang, S., Chen, M., Pomerleau, D. and Sukthankar, R.: Food recognition using statistics of pairwise local features, *IEEE CVPR*, pp.2249–2256 (2010).
[12] Sebastian, P., Matthias, W. and Wolf, F.: Food intake recognition conception for wearable devices, *MobileHealth '11 Proc. First ACM MobiHoc Workshop on Pervasive Wireless Healthcare*, No.7 (2011).
[13] Rondeau, T.W., D'Souza, M.F. and Sweeney, D.G.: Residential microwave oven interference on Bluetooth data performance, *IEEE Trans. Consumer Electronics*, Vol.50, Issue 3, pp.856–863 (2004).
[14] available from ⟨https://www.ettus.com/product⟩.

[15] Redmine. Wikistart—gnu radio—gnuradio.org (Retrieved Feb. 5, 2014), available from 〈http://gnuradio.org.〉.
[16] Witten, I., Frank, E. and Hall, M.: *Data mining: practical machine learning tools and techniques*, Third Edition (2011).
[17] Holmes, G., Donkin, A. and Witten, I. Weka: A machine learning workbench, *Proc. 1994 Second Australian and New Zealand Conference on Intelligent Information Systems*, pp.357–361, IEEE (1994).

Wei Wei received his B.E. degree in the department of Electronic Engineering, Tsinghua University (Beijing) in 2008. After that he received the M.E. degree in the department of Information and Communication Engineering, The University of Tokyo in 2012. He is currently a Ph.D. student in the department of Information and Communication Engineering, The University of Tokyo. His research interests are mainly in the aspects of Wireless Power Transmission (WPT) via Resonance Coupling and Ubiquitous Computing. He is a student member of IEICE, IPSJ and IEEE. He was doing research on WPT at Georgia Insitute of Technology as a visiting scholar from Oct. to Dec. in 2011.

Akihiro Nakamata is a student in the department of Information and Communication Engineering, The University of Tokyo (Master's Degree). His research interests are in the area of Ubiquitous Computing and Wireless Power Transfer. He is currently working on developing wireless power transfer platform for implantable medical devices. He is a student member of IEICE.

Yoshihiro Kawahara is an associate professor in the department of Information and Communication Engineering, The University of Tokyo. His research interests are in the areas of Computer Networks, Ubiquitous and Mobile Computing. He is currently interested in developing energetically autonomous information communication devices. He received his Ph.D. in the department of Information and Communication Engineering in 2005, as well as M.E. in 2002, and B.E. in 2000, in the same department of the University of Tokyo. He joined the faculty in 2005. He is a member of IEICE, IPSJ, and IEEE. He's a committee member of IEEE MTT TC-24 (RFID Technologies.) From 2011 to 2013, he was a visiting scholar at Georgia Insitute of Technology. He was a visiting assistant professor at Massachusetts Institute of Technology in 2013.

Tohru Asami received his B.E. degree and M.E. degree in electrical engineering from Kyoto University in 1974 and 1976 respectively, and Ph.D. from the department of Information and Communication Engineering from The University of Tokyo in 2005. In 1976, he joined KDD (KDDI) in Tokyo. Since that time, he has been working in several research areas such as UNIX-based data communication systems, network management systems, etc. After C.E.O. of KDDI R&D Labs. Inc., in 2006, he moved to The University of Tokyo as a professor of Dept. of Information and Communication Engineering, Graduate School of Information Science and Technology. He is a member of the IEEE. From 2003 to 2005, he was a vice chairman of the board of directors of Information System Society in The Institute of Electronics, Information and Communication Engineers, Japan (IEICE-ISS).