Detecting Cooking State with Gas Sensors During Dry Cooking

Sen H. Hirano, Jed R. Brubaker, Donald J. Patterson, Gillian R. Hayes

Department of Informatics

University of California, Irvine

shirano@ics.uci.edu, jrbrubak@ics.uci.edu, djp3@ics.uci.edu, gillianrh@ics.uci.edu

ABSTRACT

Gas sensors have the potential to assist cooking by providing feedback on the cooking process and by further automating cooking. In this work, we explored the potential use of gas sensors to monitor food during the cooking process. Focusing on dry cooking, we collected gas emissions using 13 sensors during trials in which food was cooked to various degrees of doneness. Using decision tree classifiers, we were able to predict doneness for waffles and popcorn with 73% and 85% accuracy, respectively. We reflect on the potential reasons for this variation and the ways in which gas sensors might reliably be used in ubicomp applications to support cooking.

Author Keywords

Electronic nose; gas sensors; food; cooking state

ACM Classification Keywords

H.4.m. Information systems applications: Miscellaneous.

INTRODUCTION

Inferring the doneness of food as it is being cooked can be difficult. Novice cooks, in particular, struggle because they are not familiar with the characteristics of an item cooked to the appropriate doneness. This lack of confidence is negatively correlated with success, and novices are significantly less confident performing dry cooking techniques such as frying and microwaving [2] as opposed to cooking in water-based liquids. Likewise, those with low or no vision may be unable to visually inspect the food. Finally, in industrial kitchens, cooking at large scale requires automated determinations of food state. Thus, in this work, we explore how ubicomp technologies—gas sensors—might be used to support dry cooking.

Technology has long been appropriated to assist in cooking, such as thermometers for level of doneness and computer vision for level of browning [5]. However, not all cooking scenarios are amendable to using vision or probe

thermometers, because the food can be visually obstructed, and inserting thermometers can be inconvenient. Gas sensors can "smell" odors and other gases released while cooking, and have had some success in classifying smells when used in testing chambers [9]. We build on this work by expanding these determinations into new, less sterile and controlled environments.

In this paper, we present our findings from using an electric nose to infer information while performing dry cooking methods in a residential kitchen. We describe an experiment testing two dry-cooked products: microwave popcorn and waffles. We used an electronic nose of 13 sensors to collect data from three samples of each food, cooked for successively longer durations until the samples burned and analyzed these using several classifiers.

With the data collected from our electronic nose we found decision trees to have potential for detecting optimal cooking doneness in popcorn (84.8% accuracy), waffles (73.3% accuracy), and a combined dataset (75% accuracy). In addition, through a post-hoc analysis of our data we were able to detect the type of food being cooked. While our experiment was not optimally designed for this finding, detection of food type suggests an exciting direction for future research. A smart kitchen, for example, could first detect what is being cooked and then support the proper doneness classifier. We discuss issues with our sensing platform and analysis approach that could explain the difference in our results.

RELATED WORK

This research touches on two distinct but important areas of inquiry. First, from a domain standpoint, we consider how to teach people to cook in ubicomp-enabled kitchens, our primary motivating scenario. Second, in terms of underlying technology, we describe the ways in which gas sensors have been used in similar efforts.

Previous work has described cooking related systems that use a variety of technologies, including accelerometers and pressure plates [9], and vision [7]. However, these systems are primarily concerned with providing recipe suggestions, nutrition tracking, or assisting ingredient preparation and recipe navigation. One system, Panavi, is notable here for measuring the temperature of the pan to determine doneness [9]. However, even in this case, the system neglects the actual food itself. Thus, there is an open area of inquiry for

Food	Portion	Cooking medium	Heat	Evaluation criteria	Sec until burnt	# Trials
Waffle	$\frac{1}{3}$ cup of batter	Waffle Iron	415 F	Doneness rating of 0-20	400	60
Popcorn	2.75oz bag	Microwave	1000Watt	Popped % and burned %	220	33

Table 1. Units of analysis

measuring food state while cooking. An optimal situation would provide data from within the food, however gas sensors provide a more practical proxy.

There has also been substantial work using electric noses to detect differences in food in highly controlled experimental environments [1]. Researchers have applied gas sensors in "simple" applications, such as tracking toast as it burns [3], tracking the stages of baking [8], and determining freshness of ingredients [1]. Most projects from the aforementioned literature use laboratory equipment, a situation unrealistic for the dynamic kitchen environment. Thus, there remain research questions around how these sensors might be used in an uncontrolled open-air environment, such as a smart kitchen. Goschnick et al. placed an electronic nose above a toaster to identify cooking phases [3]. We extend this work by looking at two additional foods that undergo more complex changes while cooking.

Although gas sensing has not yet been used to detect cooking states, we are particularly inspired by Kobayashi et al.'s use of gas sensors and accelerometers worn around the neck to infer location and activity [6]. In this paper, we present a solution that uses sensors fixed in the environment to infer food and cooking states. In this way, this work contributes to the state of the art in ubicomp cooking support by detecting the internal food states without piercing or touching the food and by exploring the use of sensors in an open air uncontrolled environment.

METHODS
Data Collection
Individual gas sensors are designed to react to specific molecules, but other molecules impact readings, complicating the detection of specific gases. Due to this limitation, multiple gas sensors are typically used in an array, generally referenced as an "electronic nose", so that more unique patterns can be detected in sampling readings across the multiple sensors [9].

In this work, we created a custom electronic nose that includes: two Alcohol, two Methane (CH_4), two Carbon Monoxide (CO), two Liquefied Petroleum Gas (LPG), Hydrogen (H_2), Carbon Dioxide (CO_2), Hydrogen Sulfide (H_2S), and Ammonia (NH_3). These sensors were chosen based on previous testing and sensitivity to molecules released during cooking, such as H_2, H_2S, and NH_3 produced in the Maillard reaction. To evaluate the ability of an electronic nose to infer the state of foods cooked using dry cooking methods in a naturalistic environment, we performed an experiment cooking popcorn and waffles in a home kitchen with no forced-air flow while recording.

The sensors were preheated at the beginning of each day's trials to adhere to manufacturer specified preparation

requirements. Before each trial, the cooking device was cleaned and a vent was turned on to clear the air in the kitchen until the sensor readings stabilized (~4 minutes) then the vent was turned off. Then the sensors were placed two feet above the cooking equipment. This distance was chosen because it was approximately the distance to the vent in local apartments.

We cooked three samples of each food at 20 second intervals up to the point of burning. Thus, if a food burned at 60 seconds, there would be 9 samples cooked (*i.e.*, three cooked for 20 seconds, three for 40 seconds, and three for 60 seconds). For each sample, we recorded the start and stop times of cooking, corresponding to the application and removal of the heat source. Each of the foods types were cooked using the preparation methods described on the package, except for the timing as previously noted (see Table 1). In total, 93 trials were performed.

We rated doneness of waffles on a 20-point scale across two researchers, who disagreed by no more than one point and the doneness of popcorn by counting the number of kernels popped, unpopped, and burnt. We then transformed the popcorn data to a 1-20 scale and translated both food items from the 1-20 scale to one of three nominal values: undercooked (0-7), cooked (8-11), and overcooked (13-20). These ranges were determined based notes taken about human-taste tests performed for each trial.

Feature Extraction
In preparation for the classifiers, we processed the data to highlight the effect of time and distance on the gas sensors' ability to react to the gas molecules released in an odor. First, we segmented the time series data into five-second frames. We used this frame to detect potency and diffusion changes in near real time, while avoiding the delayed feedback that would be caused by a larger window. Then, each sample, values were shifted so the lowest value was represented as zero. We then extracted eight statistical features from each sensor' values collected from the 13 gas sensors over each five-second frame. These features include: min, max, mean, standard deviation, sum, variance, slope, and y-intercept.

After computing these features, we used the last five-second frame from each sample and the associated doneness of the sample, as the target class, to train the classifiers. We used only the last frame, because that frame corresponds to the classification task that would be required in a live application of this technology (*i.e.,* disregarding temporal changes).

Modeling
For analysis, we selected seven different algorithms: SMO, Random Tree, J48, Random Forest-10, Random Forest-100,

Simple Logistic, and Logistic. We selected the best performing algorithm based on the greatest ROC Area. These ended up being SMO, Simple Logistic, and Random Forest-100. An overview is provided in Table 2.

RESULTS

Using classifiers on gas emission data while cooking was more accurate for popcorn (84.8% accuracy) than waffles (73.3% accuracy). The classifiers were able to correctly classify 28 out of 33 instances of popcorn data (see Table 3). Although each were 84.8% accurate, Random Forest-100 provided the greatest ROC at .981. Upon closer inspection, misclassifications were more common for the cooked category. For waffles, the classifier was only able to classify 44 of the 60 instances for 73.3% accuracy (see Table 4). In this case, simple Logistic provided the best ROC of the algorithms with the highest accuracy (.902). The classifier was not able to correctly identify any of the cooked waffle trials.

To simulate the cooking of the food in a kitchen that might not be aware of what food is being cooked, we combined the data from both popcorn and waffles, resulting in 75.3% accuracy from the Random Forest classifier (see Table 5). In practice, a smart kitchen could improve its accuracy by first teasing apart the types of food being cooked [1, 9]. In other words, classifiers will likely be able to differentiate between the foods being cooked and then follow up with the proper algorithm to determine how cooked the food is.

DISCUSSION

The decrease in accuracy when classifying waffles can be explained by the more complex patterns exhibited while cooking. Future studies should explore the benefit from temporally informed features or models that allow memory of prior state changes, such as Hidden Markov Models, to help account for the repeating feature patterns in the data.

After the classifiers are trained, a gas sensor based cooking system should be able to determine how close to the "ideal" state the food currently is. In this way, gas sensors could usefully support cooks, particularly those who cannot successfully visually inspect food for ideal cooking times. Beyond our primary motivating example of residential cooks, these kinds of predictions could be helpful in industrial kitchens as well, where tight control of product consistency and quality is essential. Thus, future work also demands consideration for how such a system might provide appropriate feedback to cooks, including novices, those with visual impairments, and even other machines.

Although promising, the small sample size, particularly of the "done" samples, limits the power of this work. A study with more samples at each time period might determine whether the lower accuracy of cooked samples was caused by the small sample size or misclassification. Furthermore, evaluation by human raters was used to produce our target class; however, a more objective analysis of the samples may create cleaner data that machine learning could exploit. In particular, the difficulty of detecting ideal doneness

Algorithm	popcorn		waffles	
	Avg TP	**ROC**	**Avg TP**	**ROC**
SMO	0.733	0.827	0.727	0.838
J48	0.55	0.631	0.788	0.765
Random Forest-10	0.65	0.84	0.848	0.927
Random Forest-100	0.733	0.856	0.848	0.981
Simple Logistic	0.733	0.902	0.848	0.946
Logistic	0.633	0.803	0.727	0.839

Table 2. Overview of algorithms and effectiveness

	under-	cooked	over-	recall
undercooked	15	0	0	1
cooked	1	6	2	0.66667
overcooked	0	2	7	0.77778
precision	0.9375	0.75	0.77778	

Table 3. Confusion matrix for popcorn classification[1]

	under-	cooked	over-	recall
undercooked	18	1	2	0.85714
cooked	1	7	3	0.63636
overcooked	5	4	19	0.67857
precision	0.75	0.58333	0.79167	

Table 4. Confusion matrix for waffle classification[1]

	under-	cooked	over-	recall
undercooked	30	1	5	0.83333
cooked	4	9	7	0.45
overcooked	3	3	31	0.83784
precision	0.81081	0.69231	0.72093	

Table 5. Confusion matrix for the combined dataset[1]

compared to "undercooked" suggests that initial applications in this area might choose to focus on a binary classifier of "undercooked" which, while not enabling autonomous kitchens, might be beneficial for the prevention of food borne illnesses associated with undercooked food. However, a user study to determine what kinds of interactions would most benefit novices should be conducted first.

Automatic determination of the state of a food regardless of its type is a necessary step for ubicomp technologies to be able to operate autonomously in a smart kitchen. As expected, multiple food types create problems when trying to identify specific states for one type of food; however if systems start with food type, selection of the proper model to use can be facilitated.

The discrepancy between the two cooking methods might be explained by the difference in the intensity of the cooking. Microwaves heat the popcorn more efficiently, thus burning the popcorn after 220 seconds instead of the 400 seconds required to burn waffles. Meanwhile, a visual analysis of the both popcorn and waffle data suggests the existence of cooking phases. These cooking phases are state

---

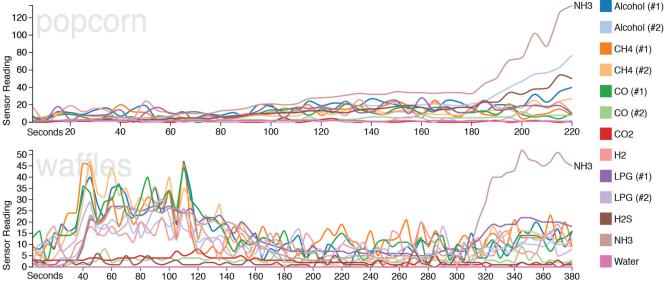[1] Classified using Random Forest-100.

Figure 1. Sensor readings across time from popcorn (top) and waffle (bottom) trials. With popcorn, sensor levels consistently rise as the popcorn begins to burn at 180 seconds. Waffles exhibit a bimodal shape, when food begins to burn at 340 seconds, particularly NH_3.

changes that food goes through while cooking. In waffles, these phases could be: heating, setting, drying, and crisping. In this case, doneness cannot be detectable based solely on the statistics gathered in a five-second window.

As popcorn cooks, sensor values steadily grow until the popcorn begins to burn, at which point, the slopes increase considerably -- particularly NH_3 (see Figure 1, top). The consistently positive relationship is what allows a particular five-second frame to be used to reliably infer the current state of the food being cooked. Waffles, on the other hand, show a bimodal shape, in which the end of the first hump coincides with the time when waffles first met our "cooked" condition. Afterwards, readings rise slightly and the sensors produce noisy data until the waffles burn a few minutes later (see Figure 1, bottom). Because the current classification replies on distinct differences between five-second timeframes, this noisy area is likely a strong contributor to the misclassification of the waffle data. But samples taken from various points in time during the noisy area are indistinguishable. Additionally, ML techniques that are able to detect and remember distinctive phases from an array of sensor streams should support tracking complex cooking processes. Ultimately, real world cooking scenarios require a phased approach in order to detect both transitions in food during the cooking process, as well as transitions in recipes that involve multiple ingredients (*e.g.*, adding meat after onions have browned).

CONCLUSION & FUTURE WORK
In this paper, we conducted an analysis of gas sensor data from cooking popcorn and waffles collected in a residential kitchen. We found that using Random Forest to create a model from these data is promising for popcorn (84.8% accuracy) and somewhat promising for waffles (73.3%); however, the combination dataset is only 75% accurate. Thus, if food types can be reliably differentiated, classification would be more accurate for popcorn.

Using gas sensors is a step to having a fuller picture of what is happening in a kitchen. In addition to the suite of other ubicomp sensor technologies (*e.g.*, audio or video), gas sensors have the added benefit of being able to provide information previously invisible to many other kitchen tracking systems, providing the ability to not only classify the type of food, but also to help with nutritional tracking.

ACKNOWLEDGEMENTS

REFERENCES
1. Balaban, M. and Korel, F. Electronic Nose Technology in Food Analysis. Handbook of Food Analysis Instruments. CRC Press, 2008.
2. Caraher, M. *et al.* The state of cooking in England: the relationship of cooking skills to food choice. *British Food Journal*, 8 (1999), 590–609.
3. Goschnick, J. An electronic nose for intelligent consumer products based on a gas analytical gradient microarray. *Microelectronic Engineering* 57-58, (2001), 693–704.
4. Hall, M. et al. The weka data mining software: an update. SIGKDD Explor. Newsl, 11, 10–18, November 2009.
5. Hernández, J.A., Heyd, B., and Trystram, G. On-line assessment of brightness and surface kinetics during coffee roasting. *Food Engineering* 87, 3 (2008), 314–322.
6. Kobayashi, Y *et al.* A Context Aware System Based on Scent. In *Proc. ISWC 2011*, 47–50.
7. Panger, G. Kinect in the kitchen: testing depth camera interactions in practical home environments. In *Proc. CHI 2012*, 1985–1990.
8. Romani, S., *et al.* Physico-Chemical And Electronic Nose Measurements On The Study Of Biscuit Baking Kinetics. *Italian Journal of Food Science* 24, 1 (2012), 32–40.
9. Uriu, D., *et al.* Panavi: Recipe Medium With a Sensors-Embedded Pan For Domestic Users to Master Professional Culinary Arts. In *Proc. CHI 2012.* 129-138.
10. Wilson, A.D. and Baietto, M. Applications and Advances in Electronic-Nose Technologies. *Sensors* 9, 7 (2009), 5099–51.