MimiCook: A Cooking Assistant System with Situated Guidance

Ayaka Sato
The University of Tokyo
Hongo, Bunkyo, Tokyo, Japan
ayakasato@acm.org

Keita Watanabe
Meiji University
Nakano-ku, Tokyo, Japan
watanabe@gmail.com

Jun Rekimoto
The University of Tokyo / Sony Computer Science Laboratory
Hongo, Bunkyo, Tokyo, Japan
rekimoto@acm.org

ABSTRACT
Referring to documents is common when making things, but there is a difficulty caused by the gap between a written description and the actual context of making. For example, when cooking following a recipe, people may lose their current position in the recipe, misunderstand the required amount of ingredients because of complicated measuring units, or skip steps by mistake. We address these problems by selecting cooking as our domain. Our proposed cooking support system, MimiCook, embodies a recipe in a real kitchen counter and directly navigates a user. The system consists of a computer, a depth camera, a projector, and a scaling device. It displays step-by-step instructions directly onto the utensils and ingredients, and controls the guidance display in accordance with the user's situations. The integrated scaling device also helps users to avoid mistakes with measuring units. Results of our user study shows participants found it easier to cook with the system and even subjects who had never cooked the assigned recipe did not make any mistakes.

Author Keywords
Kitchen; recipes; cooking; situated guidance; depth sensing;

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
We often refer to documents when making things, such as cooking following a recipe, building a plastic model in accordance with an instruction manual, or building an electric circuit in accordance with a handwritten schematic. However, we sometimes make mistakes by skipping steps, misunderstanding instructions, or ignoring context. This is because the document and the environment in the real world exist separately. Because of this gap, a reader of the document has to link them by comparing them, and if the reader makes a mistake in linking, a failure occurs. In this paper, we focused on a staple domestic activity: cooking.

TEI 2014, Feb 16–19, 2014, Munich, Germany.

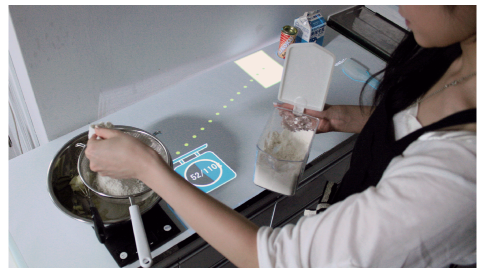

Figure 1. ***MimiCook*** **system: Cooking instructions are directly projected onto the kitchen counter. Guidance information automatically changes in accordance with the user's progress.**

Cooking following a recipe is not as trivial as it may seem. A chef has to be aware of which steps are done and which are not to know what the next step is. However, several steps are written in one sentence or one block of text, so the chef sometimes loses his/her place in the recipe. This leads to mistakes such as missing steps or using the wrong amount of an ingredient without noticing. As cooking is a real time task, excessive efforts to understand the recipe interrupt the cooking activity itself and might result in a bad dish.

Another problem is measurement. There are so many different units (grams, pounds, ounces, spoons, milliliters etc.) and measuring tools (kitchen scales, teaspoons, tablespoons, cups, etc.) that it is difficult for novice chefs to handle them. Differences in amounts are very problematic because even a 50ml difference may change the taste drastically. Units also differ between countries; for example, "a cup" is normally equivalent to 250ml in most of North America and Europe but is 200ml in some Asian countries. In addition, if the weight of available ingredients, such as meat, is different from the corresponding weight written on a recipe, all other amounts of ingredients must be recalculated accordingly, but this is very cumbersome and may cause mistakes.

In this paper, we propose a kitchen-embodied recipe system named *MimiCook* (Figure 1). This system offers an Augmented Reality (AR) cooking environment configured by a computer, a depth camera, a projector, and a scaling

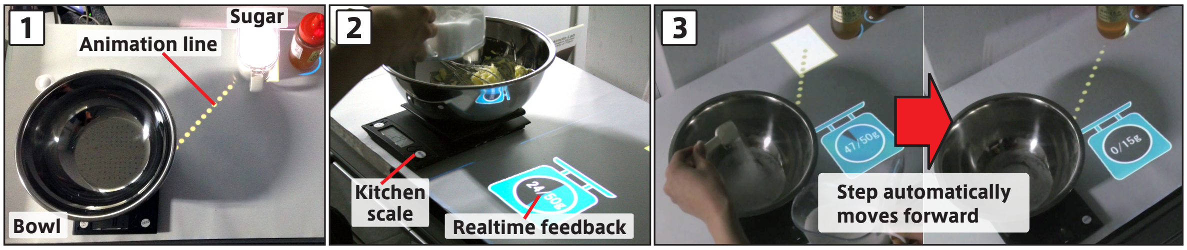

Figure 2. Main ideas. (1) Instructions are projected directly onto the kitchen counter and objects. (2) Scaling device with Bluetooth provides real-time feedback. (3) Step automatically moves forward in accordance with weights and objects' movements.

device. By using this setup, recipes are embodied in a kitchen counter. Furthermore, *MimiCook* also displays guidance directly onto the object of interest and controls the guidance display in accordance with the user's situation.

RELATED WORK

Several systems use methods for integrating recipes and kitchens [5,7,9,11,12,14]. Counteractive [7] uses projections of a recipe onto a kitchen counter that users can operate by touching the counter. Panavi [14] is a sensor-embedded frying pan that recognizes a user's situation by managing temperature. However, with these systems, users still have to read the instructions and be aware of where they are in the recipe. Kitchen of the Future [10] is a kitchen installed with several monitors, cameras, and foot switches so that users can always read the recipe no matter where they are in the kitchen. However, it does not recognize the context of cooking activities. Cooking Navi [3] and Video cooking [2] synthesize multimedia such as videos and pictures to make text-based recipes more understandable by presenting detailed information. Some systems specialize in recognizing user's activities and objects in the kitchen [1,4,6,8,13,15]. Closely related to our approach, [8] uses depth images added to RGB images to recognize fine-grained kitchen activities such as mixing and how many spoonfulls are poured. In contrast to these systems, our system focused more on supporting the whole cooking process in accordance with the context of cooking, rather than just recognizing cooking activities.

MIMICOOK SYSTEM

MimiCook is a system in which recipes are embodied in a kitchen counter. We first conducted a pilot study to observe what mistakes are likely to occur when cooking following a recipe. We recruited 10 participants (three males; seven females) for this study. We instructed all participants to cook the same muffin recipe. During the study, we observed that some participants were confused as to which measuring tool or utensil to use. Additionally, one participant misread the amount of ingredients, and another participant missed a step in which several ingredients were to be measured. Our observations and questionnaire results from this study confirmed the problems above.

In accordance with these observations, we designed the system on the basis of three main ideas as described below.

1. Direct and Spatial Navigation

To link a recipe with the real environment, the system projects navigation directions directly onto objects such as a kitchen counter, utensils, and ingredients. This enables users to act intuitively and avoid wrong usages of ingredients and utensils. Figure 2-1 shows an example scene of a step: "Put 50 grams of sugar into a bowl". The yellow line at the center animates the task directive to put the sugar from the container in to the bowl. Similarly, when the instruction is "Mix well with a whisk", a line animates a connection from a whisk to the bowl and shows "Mix well" at the center.

2. Integration of a Scaling Device

The system is integrated with a scaling device, which visualizes information about scaling in accordance with the recipe's scaling information and weight on the scales (Figure 2-2).

Numbers in recipes are more likely to be mistaken than words, especially when there are more than two numbers. For example when the instruction is "Add 30 grams of butter, an egg, and 80 grams of flour", a user first places a bowl (52 grams for example) on a kitchen scale, then adds butter (82 grams in total), then adds an egg (55 grams, 137 grams in total), and then adds the flour until the weight is 217 grams (Figure 3). However, once the user forgets the weight along the way, s/he has to scale again which is difficult after combining several different ingredients. Commercially available digital kitchen scales have a "TARE" button, which resets the scale to zero. When a user presses the button, the display shows "0" no matter what is on the scale. However, if the user forgets to press the button, the above problem may occur. Our system automatically resets to zero when a step moves forward. Thus, users do not need to calculate manually, which in turn helps users to avoid mistakes with weight units. Optionally, the user can adjust the quantity of the recipe in accordance with a specific ingredient's weight. To adjust, the user picks up the ingredient from the counter and places it on the scale, and the rest of the ingredient amounts will be adjusted automatically.

3. Situated Guidance

The system guides the next action step-by-step by recognizing the user's current progress in the recipe (Figure

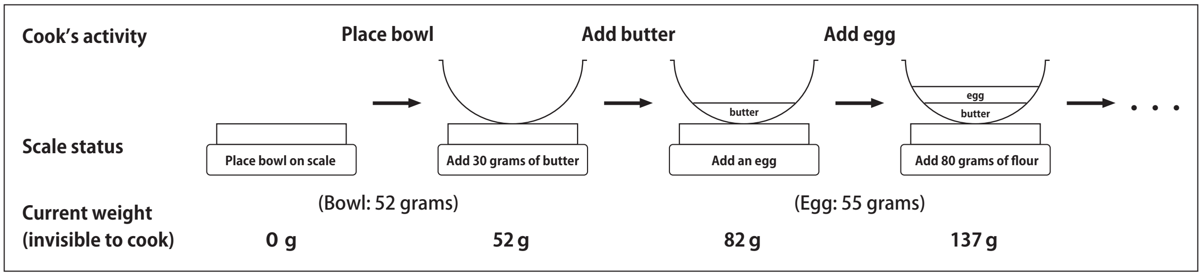

Figure 3. Cooks normally have to calculate required weight. Scaling device integrated in our system does not require calculation or memorize the weight, so that it lessens mistakes.

2-3). The user's situation is recognized based on the basis of objects' positions and weight information. For example, if the step is "Put 50 grams of flour into a bowl", the step moves forward when the user picks up a flour container and puts 50 grams into a bowl. If the user picks up a different container, the system alerts the user that it is wrong. This prevents the users from losing track of what their current progress is and helps users to focus their full attention on the cooking activity itself.

SYSTEM CONFIGURATION

MimiCook consists of a computer, a depth camera, a projector, and a digital kitchen scale (Figure 4-1). The depth camera and the projector are installed above the kitchen counter to detect objects and project a visual guide directly onto the ingredients, utensils and the kitchen counter. A kitchen scale is installed with a Bluetooth chip that sends weight information to the computer in real-time. After a recipe is selected, a step-by-step guide is projected in accordance with the user's current status in the recipe.

Recognition of the User's Cooking Situations

The depth camera recognizes objects' existence at specific places to judge whether the user is following the instructions. The recognition is based on the depth map.

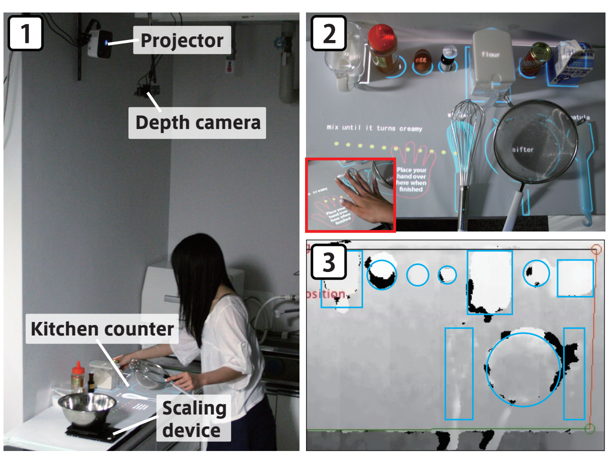

Figure 4. (1) System configuration of MimiCook. (2) Real scene of kitchen counter. (3) Corresponding depth image for recognition.

The fundamental benchmark of the depth is determined by specifying the working area (four points) of the kitchen counter top. Figure 4-3 shows the recognition screen from the depth images. The white parts indicate that they are above the level of the kitchen counter. The blue lines indicate that the object exists. The scaling device is also used to recognize the amount of ingredients currently being used. The objects' movements and current used weight trigger step progressions. For the steps that cannot be judged from these two elements, the user places his/her hand over an image of a hand to move onto the next step (Figure 4-2).

PRELIMINARY OBSERVATION

We conducted a preliminary observation to see whether the system can be operated without any explanation. Two subjects (one male and one female) who did not know about the system prior to the experiment participated in the observation. We asked them to cook the same muffin recipe used in the pilot test. The male participant had never cooked pastries, while the female participant cooks them several times a year. Both participants performed the experiment without any prior explanation on how to use the system. We recorded videos from cameras above and at the side of the kitchen counter to see how participants handle the utensils. Then, we conducted post-experiment interviews to gain more insights into the participants' experiences of using our system.

In the experiment, neither participant made any mistakes, even the male participant who cooked this kind of recipe for the first time. In the interview they both answered that the system is easy to use, so they had no confusion about which utensil to use. However, some instructions were changed too quickly while the user still wanted to see them for longer. For example, when a user was required to put two drops of vanilla oil into a bowl, the step moved forward as soon as the participant picked up the bottle. This inconvenienced the female participant. In addition, when they are pouring liquids such as milk, they tended to pour more than required even though they understood the required amount. They also had difficulty knowing how long they should mix when the instruction only said "Mix well".

The evaluation results show that the system is user-friendly for beginners, but it needs to be improved to explain what to do with ingredients better. For example, it may be more effective to show "Pour slowly" with liquid ingredients.

DISCUSSIONS

MimiCook currently recognizes cooking progress by recognizing the existence of ingredients placed on the kitchen counter. During the design phase, we also considered methods for recognizing cooking contexts with sensors, such as using visual markers and RFID tags. However, they cannot realistically be attached to all the ingredients, especially raw food. It is also difficult to recognize food accurately with image recognition, as colors and shapes can vary widely, even for the same food. Therefore, currently our system asks a chef to place necessary ingredients at the designated positions, guided by projected images.

CONCLUSION AND FUTURE DIRECTION

In this paper, we proposed *MimiCook*, a cooking support system to fill the gap between recipes and actual cooking activities. Based on a user study, our system succeeded in recognizing most of the steps without sophisticated image recognition. Users had no difficulty cooking with the system, and according to their feedback, cooking with our system was smoother than with written recipes.

Our current system has room for improvement in recognizing user's actions, such as distinguishing whether a user is mixing or whipping ingredients. By combining our system with activity recognition, such as that previously studied by [15], it should be possible to detect the detailed process of activities such as sprinkling two drops of vanilla oil, which is currently not supported by our system. It would also enable the detection of objects without setting them in specific places. In addition, projecting videos and pictures onto the counter, like in Video cooking system [2], would help users to cook more complicated recipes.

Finally, we plan to include the "record" part as shown in Figure 5. Recording used amounts automatically is significant in making recipes because recording manually what and how much of the ingredients were used is difficult. Also, people may make mistakes and forget steps when writing down using only their memories.

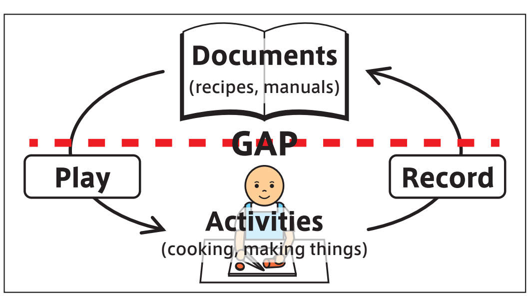

Figure 5. Future direction of *MimiCook*.

REFERENCES

1. Chi, P.-Y., Chen, J.-H., Chu, H.-H., Lo, J.-L. Enabling calorie-aware cooking in a smart kitchen. *In Proc. PERSUASIVE '08* (2008), 116–127.
2. Doman, K., Kuai, C.-Y., Takahashi, T., Ide, I., Murase, H. Video cooking: towards the synthesis of multimedia cooking recipes. *In Proc. MMM '11* (2011), 135–145.
3. Hamaoka, R., Okabe, J., Ide, I., Satoh, S., Sakai, S., Tanaka, H. Cooking navi: assistant for daily cooking in kitchen. *In Proc. ACM Multimedia* (2005), 371-374.
4. Hooper C, J. Preston, A., Balaam, M., Seedhouse, P., Jackson, D., Pham, C., Ladha, C., Ladha, K., Plötz, T., Olivier, P. The French Kitchen: Task-Based Learning in an Instrumented Kitchen, *In Proc. UbiComp '12*, ACM Press (2012), 193–202.
5. Ikeda, S., Asghar, Z., Hyry, J., Pulli, P., Pitkanen, A., Kato, H. Remote assistance using visual prompts for demented elderly in cooking. *In Proc. ISABEL '11*. ACM Press (2011), 1–5.
6. Ji, Y., Ko, Y., Shimada, A., Nagahara, H., Taniguchi, R. Cooking gesture recognition using local feature and depth image, *In Proc. CEA '12*, ACM Press (2012), 34–42.
7. Ju, W. Hurwitz, R., Judd, T., Lee, B. Counteractive: An Interactive Cookbook for the Kitchen Counter. *In Proc. CHI EA '01* (2011), 269-270.
8. Lei, J., Ren, X. and Fox, D. Fine-grained kitchen activity recognition using RGB-D, *In Proc. UbiComp '12*, AMC Press (2012), 208–211.
9. Mennicken, S., Karrer, T., Russell, P., Borchers, J. First-person cooking: a dual-perspective interactive kitchen counter. *In CHI EA '10* (2010), 3403–3408.
10. Siio, I., Mima, N., Frank, I., Ono, T., Weintraub, H. Making recipes in the kitchen of the future. *In Proc. CHI EA '01*, ACM Press (2004), 1554–1554.
11. Suzuki, Y., Morioka, S., Ueda, H, Cooking Support with Information Projection onto Ingredient. *In Proc. APCHI '12*, ACM Press (2012), 193-198.
12. Terrenghi, L., Hilliges, O., Butz, A. Kitchen stories: sharing recipes with the Living Cookbook. Personal and Ubiquitous Computing, June 2007, Volume 11, Issue 5, 409-414.
13. Tran, Q.T., Calcaterra, G., Mynatt, E.D. Cook's Collage: Deja vu display for a home kitchen. *In Proc. HOIT 2005*, 15-32.
14. Uriu, D., Namai, M., Tokuhisa, S., Kashiwagi, S., Inami, M., Okude, N. Experience "panavi,": challenge to master professional culinary arts. *In Proc. CHI '12*, ACM Press (2012), 1445–1446.
15. Ziola, R., Grampurohit, S., Nate, L., Fogarty, J., Harrison, B. OASIS: Creating Smart Objects with Dynamic Digital Behavior, *Workshop at IUI 2011*.