Employ Gesture Recognition Interface to Screen Operation in Cooking Scenario

Yu-Chun Annester Huang, Po-Hung Chen, Yan-Liang Chen, Saiau-Yue Tsau, Ko-Chiu Wu

*<u>annesterhuang@gmail.com</u> ** <u>leo1989125@hotmail.com</u> ***<u>s444578@</u> **** <u>062842@yahoo.com.tw</u> ***** <u>kochiuwu@mail.ntut.edu.tw</u>

Abstract: Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. (Matthias, 2007) Following the development of the Human computer interaction, people using electronic device for communicating with computer system. However, in some specific cases, operating Touch Screen is not applicable to use, when users' hands are smeared with some substances such as food material while cooking. Therefore, this study developed a set of gesture recognition Interface aiming at cooking scenario in which user can operate electronic receipt without touch screen operation or any traditional input devices. Referring to existing knowledge, a set of hand gestures was developed includes finger-over-the-sensor (Active sensor), top-to-bottom (Turn Off), bottom-to-top (Turn On), Right-to-Left (Slide Page to the Left), and Left-to-Right (Slide Page to the Right). It is expected that this interface of gesture operation can be learned and operated without difficulties in cooking scenario. Finally, an assessment was carried out to validate usability of the gesture recognition Interface. After oral instruction, all subjects of the experiment completed the task without difficulties that means hand motion in this gesture recognition Interface can be intuitively conducted.

Keywords: gesture recognition, cooking scenario, voice search, Voice Navigation

1. Introduction

Design start with the most fundamental interactions in different kinds of scenarios [1], in some specific situation, people either can't use hand to operate the smart device or have to pay attention to things ahead, therefore, gesture recognition interface started to become a trend to solve problems in these scenarios, such as driving and cooking scenario. Referring to existing knowledge of gestures used in applications, smart devices now, a set of hand gestures was developed includes motion in top-to-bottom (volume down), bottom-to-top (volume up), right-to-left (slide page to the left and select previous page), left-to-right (slide page to the right and select next page), and finger-over-the-sensor (select), however, after the users experienced these gesture recognition interfaces, most of them still would like to choose touch-screen to operate because of the frequency misunderstanding between gestures and camera[3], we aimed at this problem, and tried to improve the experience of using gesture control interface by study the voice search interface which is more maturely designed by Google[4], It is expected that this progressive design of gesture operation can be learned and operated more smoothly in cooking scenario.

1.1 Gesture Control

"Customers can well imagine using gesture control on a much border scale then is common today [5], there exists strong evidence showing that future human computer interfaces will allow more natural communication via gestures, because of a good user interface needing to start with the most fundamental interactions in daily life, in some specific cases like (Figure.1)when one's hands are dirty, but they need to apply touch screen of mobile device to help them to cook, they must have a touch-less interface to operate, such as gesture recognition interface or voice recognition interface, however, how to defined a set of common gestures to perfect a careful touch-less interface design is a key, there are five gestures, top-to-bottom (volume down), bottom-to-top (volume up), right-to-left (slide page to the left and select previous page), left-to-right (slide page to the right and select next page), and finger-over-the-sensor (select), which were actually defined as common gestures used in smart appliances nowadays[6].



Figure.1 user's hands are smeared with some substances, having difficulty to operate multi-touch system.

1.2 Voice Search

A decades ago, voice search had been commonly used in a narrow sense like directory assistance or local search, but after natural user interface came out, it lied in the UI and experiences we create for voice search technologies, and how we leverage the potential of voice search technologies is to mirror human capabilities better, such as Siri system (Figure.2) in iPhone and voice search on Google Chrome provides by Google Company. Voice search optimizes the path to expert, applies to given contexts and tasks, and fulfills our needs. For perfect the usability of a voice search interface, a button involving user's permission that allows a system to ask for clarification is necessary, for instance, A button in iPhone to activate the voice search, user has to press it over 3 seconds to conform starting the voice detecting the sensor, It is effective functioning method for avoiding misunderstanding between man and machine's relation also provides the moment that when user can act to operate the function.



Figure.2 After long pressing the home button on iPhone, the microphone icon will pop up and shine as feedback to notify user, the sensor is activated.

1.3 The Problem of Sensor Mistake of Gesture Recognition

The good communication between man and machine is a key leads to a usable interface operation, however those five gestures above wasn't included a gesture which used for active the sensor that is the reason why sensor is always misunderstanding the command from user, gesture should designed as an input data which can help system to understand the command from user more efficiently and correctly.

2. The Concept

2.1 Motivation

How to make a better experience of using electronic recipe in tablet devices is the big issue of this study, however, in the process of researching similar cases, we found out most of the cooking applications have problem of frequency error communication while user is flipping the pages of recipe, so we focused on finding out one more gesture as solution for activating the sensor, on one hand, this study finally will define a set of gestures specially used in cooking scenario, on the other hand, it will bring out an application as demo version to test if it improve the usability.

2.2 Concept

How many functions should involve in a usable cooking gesture control application? Considering about the distractible environment in the kitchen, making user can both pay attention on cooking and catching information from recipe is necessary, therefore, voice navigation is suitable for taking part in one of the function. Searching for recipe always takes a lot of time in the process of cooking, voice search service can help user searching quickly. According to studies above, the key to operate a gesture control interface more smoothly is the pattern used for activating the sensor, which transferred from a button to a gesture. (Figure. 3)



Figure.3 Active sensor button transferred to Active sensor gesture.

3. Reviews and Analysis

For analyzing a mature interface design and a set of gestures used in kitchen scenario, we reviewed then summarized three cases of gesture and voice control services.

Based on the design directions we mention above, aim at voice search as data searching engine to research related cases, here are examples like (1) Google Voice search (2) Gilt taste (3) Google Maps Navigation, cases analysis initially, an overview of the results is provided.

3.1 User Interaction Experience Exploration

To cook a dish following the recipe is not easy, first the user has to decide one dish to cook, and search for recipe, after cooker found the recipe, user has to pay attention on recipe, remember the steps of recipe, and cooking at the same time, this routine runs over and over again. When we are learning how to cook a dish, an expert stands beside you and telling you steps of recipe is easier than watching a guide video or reading a cook book, because user is busy to pay attention on video recipe and cooking at the same time, the only auditory sense is free for listening to the recipe guide, following the needs above and trying to figure out best service to perfect the gesture interface in kitchen scenario is the main directions.

3. 2 Case Study- Google Voice Search to the Desktop

Google Company launched Google Voice Search(Figure.4) to the Desktop, which is a voice search engine at 14th, June, 2011, this voice searching engine can recognize over 30 kinds of languages, this voice search engine is very useful once user can't spell the vocabulary they wants to search or both of users' hands have been occupied. The full interaction analysis of Google Voice Search on the desktop is below. (Table.1)



Figure.4 Start to search only with your mouth opened, the only thing you have to prepare is a built-in microphone or the one which connected with your computer, start to record the keywords you are going to search by pressing the microphone button on the Google website.

Table.1 Analysis list of Voice search on Google

| Interaction analysis | Input | Output | Behavior | Purpose |
|-------------------------|------------|---------------|---------------|---|
| | Human Oral | Search Result | Speak/ Speech | Using oral search in specific situation |

3. 3 Case Study- Gilt taste

The e-commerce site Gilt Group appears to have combined appealing recipes with a touch-less interface in June. 2012, since tablet computers came on the market, there has been the kitchen equivalent of the space race: competition to develop technology that can keep a cook's sticky, floury hands off the touch-screen. With its new app, user can check the recipe easily (Figure.5). [7] The full interaction analysis of Google Voice Search on the desktop is below. (Table.2)



Figure.5 Using built-in camera of Pad, which will recognize the gestures and movement of users' hand, and flip the pages of the recipe without touching the screen, just lift your hand in front of camera of screen and swipe it from the right to left (like flipping page), the electronic recipe will flipped to the next step of recipe, otherwise, by swiping your hand from left to right, it will flipped back to pervious step.

Table.2 Analysis list of Gilt taste

| Interaction analysis | Input | Output | Behavior | Purpose |
|-------------------------|--|-----------------|---------------|---|
| | Gesture movement (Right to Left/ Left to Right) | Flip the recipe | Swiping hands | Keep cookers' sticky, floury hands off the touch-screen and control the recipe with gesture. |

3. 4 Case Study- Google Maps Navigation

Google Maps Navigation (Figure.6) is a mobile application that comes as a feature of the Google Maps for Mobile application for Android phones, developed by Google in October 28, 2009. This application requires a Internet to access the data (e.g. 3G, 4G, WiFi, etc.) and normally uses a GPS satellite connection to determine its location. A user can enter a destination into the application, which will plot a path to it. The app displays the user's progress along the route and issues instructions for each turn. The full interaction analysis of Google Voice Search on the desktop is below. (Table.3)



Figure.6 After driver type or choose the destination, Google Map will guide you to destination following by the best driving path it arranged, on the way of destination, the Google Map Voice Navigation system will communicate with driver for conforming the exact driving route. [8]

Table.3 Analysis list of Google Map Navigation

| Interaction analysis | Input | Output | Behavior | Purpose |
|-------------------------|--------------|----------------------------|----------|-----------------------------|
| | Type or Oral | Route/ Voice navigation | Sound | Voice navigation for driver |

As studies above, a combination of Voice Navigation, stable gesture control and voice search all together would be a way leads to a successful interactive assistant interface in kitchen scenario.

4. System Design

4. 1 Design Five Gesture Analysis in Interactive Process

Since mobile devices are coming out, people started to download receipts or search online through mobile devices, by using screen touch and gesture control to operate interface is the trend, user used for such like flipping the pages, scaling the information larger or playing the cooking tutorials via mobile devices, however, most of the users faced a common problem that gesture recognition isn't going well, therefore we figure out five gesture to solve the problem, for avoiding mistake sensor, Active Sensor Gesture compared to all the other control gestures, is a stop-gesture, from Active Sensor gesture to Next gesture is like a movement, user doesn't have to remember Active Sensor gesture because of it comes naturally from the first gesture of movement, it's the same situation when it goes to Play Gesture, Off Gesture, Pervious Gesture. (Figure.7)



Figure.7 There are five gestures used in system processing, as we can see the gestures above, they are used for active fuctions, Active sensor(Start to sensor), Next(Flip to next step), Pervious(Flip to last step), Play(Play the voice-instruction), Off(Off the voice-instruction).

4. 2 The Interaction Process of System



Figure.8 There are two options to process the system, as we can see the figure above, they are fuctions of Choose the recpie and Read Recipe & Listen to the recipe, and there are three modes user can operate in this application, the set of choose the recpie and read recipe or listen to the recipe.

In this case, system process devided into two options, choose the recipe and read the recipe. There are two functions to operate choose the recipe, the first is (a)using touch-screen to operate, the second way is to use (b)Voice search to search more quickly, therefore, after (c)uploading the voice data to Google Voice Recognation system online, the (d)voice data will turn to text by calculating and send it back to system for searching.

After user choose the recipe, the system goes to second stage, Read Recipe & Listen to the recipe, open the built-in camera on the Pad (e), user can flip the recipe either by using gesture or (f) (g) touch-screen, because of different disturbance might happened in kitchen scenario (ex. Sticky hands), camera can identify the different gestures for commanding the machine, there are instructions like (h) activating the sensor by stay palm in front of camera for 3 seconds, (i) swiping palm to left and right to flip the recipe and (j) swiping palm (up/down) to turn (on/off) the voice navigation of recipe.(Figure.8)

4. 3 Interactive Operation Process

Except of gesture control, user can also use tablet multi-touch screen to operate, besides, in the stage 1 choose recipe, user has two search modes to choose, first mode is to touch the catalog bar on the right side of Pad, and pick up one dish you would like to cook by touching one's photo, secondly, press the microphone icon on the search bar, and tell the Pad which dish you are going to search, after few second, there will be a vocabulary recognized by voice search shows on search bar, finally user has to conform is it the same with the one you said or not, finally, press the search button and start to search, after user chose one dish on the menu, the ingredient page will show on the screen, now it goes to stage 2, at this moment, lift you hand in front of camera for few second to active the camera sensor, and naturally swipe hand to the left and the recipe will goes to the next step automatically, also swipe hand to the up to turn on the voice-guide of recipe, swipe hand to the down to turn off it off. (Figure.9)



Figure.9 "WAVE COOKER" Interactive operation process

5. Scenario

For understanding the usages of "WAVE COOKER" system, one scenario is described as followed (Figure.10) to get a personal cooking.



Figure.10 Anne is a 24-years-old student who usually cooks at home, one day, she was going to prepare for her dinner, but she had no idea what to cook, so (1)she turned on (2)WAVE COKER via Pad, she used (3)voice search to search Pasta by speaking to Pad, then the systems (4)shows many pasta dishes on the screen, (5)she chose Tomato Pasta to cook tonight, before cooking, she operated this application with touch-screen, in the middle of cooking, she (6)stayed her hands in front of camera for 3 seconds and (7)swiped her hand to the left, recipe flipped to next step automatically, then (8)she swiped her hand up for playing the voice-recipe, after she listening the recipe, she (9)followed what Wave Cooker said, finally, by using "WAVE COOKER", she finished a tasty tomato pasta without any interrupt.

6. Conclusions and Discussion

After surveyed with 10 subjects, they mentioned that the volume of voice navigation wasn't loud enough once it's too noisy in the kitchen, and we still found out the Google search engine does have issues with special characters, unusual names and URLs, however most of the user represented that "WAVE COOKER" is better than other cooking applications they've used before and did improve the usability of a gesture control interface.

References

- [1] Daniel Wigdor, Dennis Wixon (2011) Brave NUI World, : Designing Natural User Interfaces for Touch and Gesture Available at < http://books.google.com.tw/books?id=ID0L0EI79-YC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false> [Online PDF] [Accessed 25 March, 2013].
- [2] Alexander F., Brian. M(2011) Searching the Web by Voice [Online PDF]. Available at http://people.csail.mit.edu/milch/papers/gvs.pdf [Accessed 18 March 2013]
- [3] YOUSE GmbH (2011) Untouched possibilities [Online PDF]. Available at <<u>http://www.youse.de/documents/Kompetenzen/Publikationen/YOUSE_HMI-</u> concept_gesture_control_in_Bathroom_and_Kitchen_- short_version.pdf> [Accessed 10 March 2013]

- [4] Emily P. (2012) Google Takes on Siri With Updated iOS Voice Search App. Available at <<u>http://mashable.com/2012/10/30/google-voice-search-ios/</u>> [Accessed 10 March 2013]
- [5] YOUSE GmbH (2011) Untouched possibilities. Available at <<u>http://www.youse.de/documents/Kompetenzen/Publikationen/YOUSE_HMI-</u> concept_gesture_control_in_Bathroom_and_Kitchen_-_short_version.pdf> [Accessed 3 March 2013]
- [6] Sangmee L., Kwangsu C. (2013) Integrative Gesture UIs for Appliance. HCI 2013.(No. R31-2008-00010062-0)
- [7] Coleen. T (2011) *Gilt Taste's iPad App Has A 'Touch-Free Recipe Mode' To Keep Screens Clean*. Available at <<u>http://techcrunch.com/2012/06/27/gilt-taste-ipad-app-touch-free-recipe-mode/</u>>[Accessed 9 March 2013]
- [8] Stan. S (2009) Google Maps Navigation Becomes Reality on Android. Available at < <u>http://mashable.com/2009/10/28/googles-turn-by-turn-navigation/</u> > [Accessed 20 February 2013]