The Future of Manual Multitasking: Investigating the Usablity and Effects of Gesture-based Interaction

Duc Anh Le

Abstract— Wiping the finger to show the navigation system, tapping the steering wheel to turn on the A/C and shaking the device to switch to the next song. If these gestures are possible it would be a great step in technology. Why am I talking in the future tense? Some of these just mentioned gestures are already in the testing phase and others are already in use as well. Gesture based interaction is no futuristic topic anymore. Many people are already researching about microinteraction and gesture-based multitasking. In some cases the word "manual multitasking" is mentioned as well. But these words are really powerful, so what can we represent with it? There are many things that play a big role in gesture-based interaction and should be kept in mind while inventing new gesture sets for multitasking usage. What effects does it have on the user or on real life in general? These questions are going to be answered in the following paper.

Index Terms—Microinteraction, Manual multitasking, Gesture-based interaction

1 Introduction

Doing more things at once became more and more important in the last few years. The people are trying to focus on their primary tasks, while still doing other secondary things like drinking coffee while writing a message or telephoning while driving. Some of these secondary tasks require more effort and others require less to be done properly [6]. They can also be really dangerous in sense of safety, if the primary task needs to be done precisely and attentive, for example driving the car. In this situation the driver shouldn't be distracted too much while putting in a disk or checking the navigation system. In the following text I'm going to illustrate the effects of manual multitasking and analyze two different methods to validate the effects of manual multitasking data.

The first of those two methods is the so-called "Manual Multitasking Test (MMT)" [2] . This test "emulates the multitasking situations in everyday" [2] life and shows the manual flexibility of the user while doing the test with 12 defined conditions. Those 12 conditions are manual demands that are needed to do the test, for example "CoffeeMug" is a condition, where the proband has to hold a coffee mug in one hand, while trying to do the primary task, in this case typing a message on a mobile phone. The test was divided in 8 steps of how to use the device for example:

- Using the device with the not preferred hand
- Using parts of the hand for something else
- Using the device with the thumb and the index finger in the same distance
- Using the device with limited movement of shoulder, elbow and wrist

At the end of the test, the experimenters are able to compare the "performance affected by manual constraints" to the "unconstrained performance" [2]. The information gained of this comparison will be gathered in a statistic, to see which constraint came out best and which constraint made the performance worse.

The second method is focused on the "effects of error on device-based gesture interaction" [1]. Everyone knows the situation, when he (to make it easier I'm using the male form for both male and female) holds something in his hand, and wants to use his smartphone at the same time. In some situation it's hard to use the smartphone correctly if the held thing is too big, so what if the person can use a simple gesture to do the input? It's easier said than done. Based on the gestures there are up to 40% error rates in the input. With the help of "Fishing or a Z?" [1] I'm going to take a closer look at different gesture sets in context with manual multitasking.

- *Duc Anh Le is studying Media Informatics at the University of Munich, Germany, E-mail: D.Le@campus.lmu.de*
- *This research paper was written for the Media Informatics Proseminar on "Manual Multitasking", 2015.*

Figure 1. three different input devices, that can be used on a computer [2]

2 Results

2.1 Manual Multitasking Test

Knowing the fact, that the hand has limited resources the MMT tries to gain some information of which type of input works best with manual multitasking actions.

In the first study they tested different pointing methods on a laptop and came to the result that the mouse was the best device for unconstrained use in comparison to the Touchpad (2nd) and the Trackpoint (3rd) *(see figure 1)*. But when it came to the manual constraints the mouse was the worst of all the input devices. This is reflected by the fact, that the mouse uses the whole hand to be used properly, what leads to nearly no space for another object to be interacted with. The condition with holding a small or a medium sized object made the mouse use nearly impossible, while doing the same condition with the Touchpad or the Trackpoint wasn?t a problem at all. In the constrained test the Touchpad came out best because the use of a Touchpad only needs 1 Finger as a resource, what leaves more resources for secondary tasks. The Trackpoint got to the second place.

In the second study the MMT tested the text entry on mobile devices with 3 different types of input. A physical keyboard, a stylus keyboard and a touchpad keyboard. The physical keyboard was the best for the unconstrained use and the touchpad keyboard was the best for the constrained use. The stylus keyboard was the worst in both tests, what is reflected by the fact that some constraints needed the stylus keyboard to be used with one hand. This made the input on a touchscreen with a stylus impossible. Also switching from the preferred hand to the non-preferred hand caused a 20% decrease in performance, which in turn caused a "0.55 drop in workload" [2]. This effect is really important for the designers to create a device, which people are willing to use with manual multitasking.

Summing up those facts the MMT can be a good basis for future experiments and research but is not completed yet, to make accurate performance enhancing design pattern. There is still a gap between manual performance and multitasking, because this test only focuses on the multitasking with the human hand. But it is also possible to do multitasking with gestures.

2.2 Technologies and techniques

Gesture based interaction has two types of recognition. The first type uses cameras to film and recognize the movement of the user, and the second type uses the "accelerometer and the gyroscope sensors" [1] of the device to recognize the movement. Both types have their pros and cons but they both adept to the same idea of gesture interaction with some kind of device. But before we start to invent our own gesture sets, we should start by looking at the error rates of gesture recognition and their effects on the user. To do that, we have to classify the different types of gestures in their manner of performing. The first type of gestures are the symbolic gestures [3] that have a specific culturally meaning. As an example we can take the American Sign Language (ASL) gestures, which are used to communicate with deaf people. Another type of gestures are the pantomimic gestures [3]. These are used to show someone an invisible tool or object with some specific movements that is usually used with that object. Digging deeper into the topic of gestures we need to specify the different sets of gestures. On the one side there are the alphabetic gestures, which are movements that should draw letters of the alphabet. On the other side there are the mimetic gestures that are movements from real-life actions, like handshaking or fishing. Comparing those two in the test from [1] showed that the subjects had to put in more effort to perform the alphabetic gestures correctly so that the computer will recognize it. For the mimetic gestures this wasn't the case. The subjects weren't frustrated at the low and medium error rates while doing the mimetic gestures [3], but while performing the alphabetic gestures the frustration rate was relatively high at all error rates [3]. The result of this effect is that people feel more helpless if they continue failing the alphabetic gestures, because there are less variable parameters for these gestures. With mimetic gestures people tend to vary their performed gestures more until it gets recognized. But these explorations only take place, if the recognition failed in first place. In conclusion this means that mimetic gestures are more immune to recognition errors than alphabetic gestures. Another finding was that if a gesture was performed quickly and failed, the subjects performed that gesture again but with a slower pace. The other way round the subject tended to do it faster if he started with slower version. Also if the subject managed to succeed in a series they are going to repeat the same gesture pattern in later blocks as well. But too much success ended up with less effort from the subjects, which resulted in incomplete gesture execution.

Figure 2. Performable microgestures while steering a car with a palm grasp [6]

3 Discussion

Combining those two studies now in sense of "manual multitasking" brings us to the conclusion that there are a lot of ways to do two things at the same time, but not all of them result in good performance and efficient resource usage. Coming back to the example of the car driver it's not easy to implement gesture interaction without limiting concentration on the street or on the drive process itself. Another thing to keep in mind are the restricted resources while holding the steer wheel. *(see figure 2)* But it still leaves some place for microinteraction. Potentially still moveable hand parts are the thumb and the index finger [6]. The first possible method that wouldn't distract the driver too much would be some kind of finger gesture recognition. In the first study we learned that using devices with less resources resulted in the best performance, for example using the touchpad only requires one finger and leaves the most opportunities for multitasking action to take place. In this case there is a possibility to build in a gesture recognizer for little finger gestures while driving with functionalities that could help improving concentration of the driver. For example we could replace the volume control with two gestures of the index finger (left wipe to turn the volume down, right wipe to turn the volume up). Other possible gestures while driving could be tipping or dragging [6]. But executing those gestures always leave a chance of lowering the performance of the primary and also the secondary task. By focusing the multitasking actions on the motor and cognitive efforts it is possible to determine it on a specific model. The Wickens' Multiple Resource Theory (MRT) says that cognitive resources can be used parallel if there is a different input method of both tasks [5]. Knowing this it is definitely possible to carry out the defined gestures while driving.

Another use-case for gestures are public displays. The people on the streets mostly have a primary task in their minds like shopping or going to work, when they walk past a public display. While probably holding their shopping bags in their hands most people can't use touch-based public displays and tend to move on. But if the interaction is based on gestures the public display has the advantage, that it doesn't require touch inputs on the screen, so that people are more likely to use it [4]. For the hygienic reasons it would be positive as well. Also the users doesn't need to come close to the screen to interact with it, what also helps with noticing the interactivity of public displays, when people interact unintendedly with it while passing by. [4]

4 Conclusion

Summing up all the points that were mentioned there is a huge benefit in doing secondary tasks for instance while driving. If the driver is able to control the navigation system or automotive functions in the car without releasing the steering wheel, there is a positive impact in the security of car driving. But the compatibility of gesture based multitasking interaction is highly dependent on the primary task, because the primary task also defines which parts of the hands are still free to be used for gesture interaction. That's why it is highly recommended to test every possible option of grasp of a device or object before we can make a proposition about the possible manual multitasking features. Another way could be the use of short interruptions for the secondary task. In [6] they said that doing the secondary task while stopping the primary task for a second could benefit for the primary task as well. This could be used with pens as an example. While drawing the person is able to change the color of the pen. The flow of drawing wouldn't be interrupted if the color change option were done with a small button tap. The experts also said that the palm grasp is the best grasp for dual-task scenarios, because in this activity the precision needed for executing the primary task is in most cases low and takes less cognitive load than pad ore side grasp tasks [6]. Furthermore gestural interfaces for multitasking activities shouldn't be seen as a replacement for other interfaces. They should be seen as an option for inventing new types of interaction and can be used as an expansion to the already existing interfaces. Also different interaction types should use different recognition technology. Cameras are limited in their function of recognizing tap-interactions or pressure based interactions and should only be used for gestural recognition. For force-based interaction there is the possibility of using sensors that can measure the applied force. All this said there is still a lot of space for future research and I'm really excited for the things that will come in the future in context with microinteraction and manual multitasking.

REFERENCES

[1] A. E. Ali, J. Kildal, and V. Lantz. Fishing or a z?: Investigating the effects of error on mimetic and alphabet device-based gesture interaction. In *ICMI '12 Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 93–100, 2012.

[2] A. Oulasvirta and J. Bergstrom-Lehtovirta. Ease of juggling: Studying the effects of manual multitasking. In *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3103–3112, 2011.

[3] B. Rime and L. Schiaratura. *Fundamentals of Nonverbal Behavior*. Cambridge University Press, 1991.

[4] R. Walter, G. Bailly, and J. Müller. Strikeaposte: Reavealing mid-air gestures on public displays. In *CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013.

[5] C. D. Wickens. *Processing resources in attention*. Academic Press, 1984.

[6] K. Wolf, M. Rohs, A. Naumann, and J. Müller. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-dependent requirements. In *INTERACT'11 Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part I*, pages 559–575, 2011.