

Identifying Interaction Problems on Web Applications due to the Change of Input Modality

André Constantino da Silva, André Luis Correa Viana, Samuel Gomes de Lima

Instituto Federal de São Paulo – Campus Hortolândia

Hortolândia, São Paulo, Brazil

andre.constantino@ifsp.edu.br, andre_viana40@yahoo.com.br, samuelsmgl@gmail.com

Abstract—New equipments and software for providing different modes for user interaction emerged and became popular in the last decade; they are used in various devices, such as mobile and game consoles, which can display Web pages due to the increasing of the processing power in these devices. Since interaction with each mode has peculiarities, when a user is interacting with a mode not considered in the design time, she might have interaction problems. Here, we present our work about Web navigation with motion sensors; we chose Google Maps, Google Street View, and TelEduc, which is an e-learning environment, to evaluate the use of WiiMote, the motion sensor of Wii console. In this paper, we analyze data from the first case study (Google Maps and Google Street View) and preliminary results from the second one (TelEduc). The collected data confirm our previous findings; adaptations are necessary to users have a good experience when navigating through web pages using a motion sensor.

Keywords—Multimodality; Usability Evaluation; Web Applications.

I. INTRODUCTION

The number of input hardware on desktop computers is increasing, with computers equipped with touchscreen devices, microphones and cameras. New computational devices with input and output hardware are been built; most of them are equipped with processing power enough to render Web pages, allowing users to navigate into Web pages using modalities not available in desktop computers. Websites/portals, even the ones that had been passed by usability testing, have interaction problems when changing the mode of interaction (cross-modality problem) [1]. Since usability is a key factor for system acceptance [2], low usability in a context can cause the user's withdrawal from interacting with an application.

Soon, given the growing use of devices equipped with motion sensors, users equipped with these devices will access Web applications. Thus, it becomes necessary to study problems caused by the changing of interaction modality from keyboard+mouse to motion sensors. Our hypothesis is that for users to have a good experience in navigate through the Web applications using motion sensors some page adaptations are necessary. To suggest adaptations on the Web pages it is necessary to understand the interactions problems that happen due to modality changing. Our general goal is to have a better understanding of the

impact due to the modality changing and to propose guidelines to design user interfaces with high usability for applications that can be accessed by many modalities.

Here, we present our work about Web navigation with motion sensors and some interaction problems that happen when users navigate through Web pages using motion sensors. We planned some case studies based on user tests, where volunteers used the WiiMote, the motion sensor controller of Wii console, to perform some tasks in Google Maps, Google Street View [3] and TelEduc [4]. Section II presents a literature review about the subject. Section III describes our methodology and data analysis. Section IV summarizes the results of our work so far.

II. LITERATURE REVIEW

Nielsen [2] defines the general acceptability of a system, composed by social acceptability and practical acceptability. One of the concepts that composes the practice acceptance is the "usefulness", which refers to use the system to achieve a particular objective. This concept consists of the usefulness and usability. Usability is a combination of five elements: ease of learning, efficiency, ease of remember, probability of the user making few errors and user satisfaction. Since usability is indirect related with the acceptability of a system, we can affirm that the user experience in using a system is a factor that determines the user's system acceptance.

Modality is the term used to define a mode in which a user's input or system's output is expressed. Nigay and Coutaz [5] define modality as an interaction method that an agent can use to reach a goal, and that modality can be specified in general terms as "speech" or in more specific terms as "using microphones". Several modalities have become research topics in recent decades; among them, we can mention the voice, handwriting recognition, touch, and gestures [6]. Bernsen [7] says that there are no two equal modalities; each of them has its own strengths and weakness.

In our previous work, we cataloged some problems that occur when accessing the TelEduc, an e-learning environment, in touchscreen devices and can occur in other Web applications [1]. These problems were identified through analysis of results generated by inspection methods (methods that involve usability experts) and empirical tests (involve the participation of users), and the identified problems were divided into three categories: problems related with the change of the modality, problems related with the change of platform and problems not related with the change of modality or platform. In another previous work

[8], we presented some results about the use of WiiMote to interact with Google Maps and Google Street View from data collected by questionnaire. The applied opinion questionnaire was composed by 16 (sixteen) questions related to enjoyment, perceived response time of the movement, handling digital artifact activation of the user interface components, selection of options, and navigation in virtual space. Comparing the users opinion, there are a slight difference between navigate in Windows 8 and Linux Ubuntu. The main difference is trigger small buttons and the reason can be related with the used pointer's icon: in the Windows is a half-filled circle and in Linux is the standard mouse icon. The standard icon has a lower visibility, but a higher accuracy. Related to interaction problems due to the size of the interface elements and the high sensitivity of the control (performing pointer movement when pressing the button) points that the need to do some adjustments for a better interaction with this device on the Web. About the how easy was to trigger small buttons of the Google Maps' user interface, 3 volunteers said that it was easy most of the time, 2 volunteers said that it was easy half of the time, and 1 volunteer said that it was easy in few times. About how easy was to use the virtual keyboard to type a text, 3 volunteers said that it was easy every time, 2 volunteers said that it was most of the time, and 1 volunteer said that it was half of the time. Here, we present results from video analysis that support our previous findings.

III. METHODOLOGY

Our methodology is based on an empirical method, user tests in laboratory, to identify interaction problems. We planned two case studies, one for Google Maps and Google Street View, and another for TelEduc. We finished the first one case study, where six volunteers navigated via the applications. The second case study is in-progress; we just did the planning and a pre-test.

The main activities of the planning phase of both case studies were: (i) selection of the tools and features, and specification of tasks to be performed by volunteers, (ii) creation of the profile and opinion questionnaires, (iii) definition of hardware and software to be used, (iv) preparation of the environment, and (v) creation of the consent term.

In the execution phase, six volunteers performed the defined tasks, while the interaction was recorded; then, the specialists analyzed these videos. Before starting to interact with the application, the volunteers filled up the profile questionnaire, and after session test, the volunteers filled up the opinion questionnaire. The profile questionnaire was composed by 13 questions about age and genre, the previous experience in using the selected applications, asking about how long and frequency the volunteer had used a Wii console, and information about which sites, browsers, frequency of use of the Google Maps and Google Street View, and where the volunteer usually access Internet from.

A. Materials and Tools

The used hardware (Figure 1) was a motion sensitive controller of the Wii console, the WiiMote, connected



Figure 1. Equipments used in the case study: infrared sensor, 19'' display and a computer.

through Bluetooth with a computer (a HP Touchsmart tx2-1040br). A 19 inches monitor was used to have a good readability in a distance required to use the motion controller, usually two steps back.

The infrared bar was placed on the top of the monitor to have a higher accuracy in the recognition of movements. The infrared bar was built with a total of 8 infrared LEDs, but tests showed that there is a better accuracy of the movements when only two LEDs are connected; so, we used this configuration.

The mapping between input data from the Wii control to mouse actions enables to use this device to interact with a desktop computer. This mapping was done on Linux through WMInput application, available in the package CWiid version 0.6.00. We select this set of software (SO plus browser and input data receiver program) due to our previous experience [8] that showed fewer differences in navigate in Linux and Windows using the WiiMote. The main difference is related to the icon of the mouse pointer: in the WMInput (Linux) application, this is the standard one, while Win Remote application (Windows 8) uses a half-filled circle.

To supply the lack of a keyboard next to the user, we used a virtual keyboard fixed in the bottom of the screen occupying about two-thirds of the screen. In the Linux SO, we used the Onboard virtual keyboard and, in the Windows SO, we used the standard virtual keyboard. To navigate through the Internet, volunteers used the Google Chrome in both SOs.

B. Case Study 1: Navigation on Google Maps and Google Street View

Google Maps [3] is a Web application that provides information about geographical regions around the world, offering aerial and satellite views. Considering its functionalities, we defined two tasks to be performed by the volunteers: Task 1 - Find your home and point it; Task 2 - Show the route from your home to college. Google Street View is integrated with Google Maps and offers street views for some cities, function accessed by the use of *Pegman* icon in the Google Maps. Considering its functionalities, we wrote two tasks to be performed by the volunteers: Task 3 -

Join Google Street View and show the facade of your home;
Task 4 - Show the route from your home to the nearest market.

Six volunteers who participated (three volunteers for each operation system) have on average 17 years old; one woman (volunteer 1) and five are males (volunteers 2, 3, 4, 5 and 6). All volunteers declare that does not have a Wii console, and only two volunteers have used sometime, although without frequently (volunteer 2 used for about 1 year, and volunteer 5 among 3-6 months). All volunteers said they access Internet, five of them reported using more than 10 hours per week, while one declared access it for about 5-10 hours a week. All volunteers declared access in their homes, and four declared that access at school. A volunteer reported using the Internet at friends and family home, and two volunteers declared access in places that offer free access through Wi-Fi or through 3G network. Three volunteers reported using Mozilla Firefox browser, while the other four ones reported using the Google Chrome browser. All volunteers reported had been used Google Maps; volunteer 5 informed that uses about once or twice a month and two other volunteers reported using more than three times per month (volunteers 2 and 3). About Street View, three volunteers reported had never been used, while one declared having used only once (volunteer 4), and two other volunteers reported using once every two months (volunteers 3 and 5).

About Task 1 - Find your home and point it, the users can use two different strategies or a combination of them: i) navigate in the map (volunteers 2, 3, and 5); or ii) use the keyboard to type the address (volunteers 1, 4, and 6). Analyzing the recorded video, volunteers had difficult to trigger the zoom buttons; this was observed mostly in interactions where volunteers used the first strategy due to the need to navigate and zooming into the map. Volunteer 2 tried 20 times to trigger one of the zooming buttons (zoom in or zoom out): 9 hits and 11 errors. To confirm that she was pointing out the correct place on map, volunteer 2 triggered the Satellite View feature. In this interaction, the user tried to trigger the feature 2 times (1 hint and 1 error). Volunteer 3 chose the same strategy; but, to trigger the zoom in feature, she did a combination of commands instead of only pressing the zoom in button in the user interface. The commands were press the left button, point to the zoom in option and press the right button. Due to the difficult in identify which buttons had been mapped to the mouse's left click and right click, the volunteer tried 12 times to zoom in (9 hints and 3 errors due to press the wrong button). She did an error in more 2 times, when she tried to trigger the zoom in feature but when she pressed the WiiMote's button, she moved the mouse cursor and the system recognized as image selection. Volunteer 5, who preferred to trigger the zoom buttons on the user interface, tried 11 times to trigger these buttons (8 hints and 3 errors).

About the interaction of the volunteers that preferred to type the address, volunteer 1 typed an address with 29 letters and triggered 37 buttons due to some wrong letters typed (4 wrong typed keys and more 4 on backspace button). Analyzing the wrong letters, we noticed proximity between

the typed wrong letter and the correct one; e.g., the volunteer wanted to type the letter 'a' and button 's' was pressed. This user also used the zoom buttons in their interaction, trying to trigger the zoom 8 times (3 hints and 5 errors).

Volunteer 4 tried to complete the task 1 by typing the address but, due to the difficult to complete it, she changed to the map navigation way. Using the typing address strategy, this volunteer typed 13 correct words of the desired address and tried to select one of the suggested addresses. Since she did not selected the desired address due to press the wrong button of the WiiMote (pressed the wrong button three times), she decided to complete the address, clicking on address bar and typing more 2 letters of the virtual keyboard. Finally, she pressed the 'enter' button of the virtual keyboard. The system showed the place of the typed address with an information window. The volunteer closed this window and the system returned to the previous position (before the address selection). The volunteer decided to change the address and typed one with 14 letters, all correctly typed. To adjust the view, the volunteer used the map navigation with zooming buttons. She moved the map three times, and tried to zoom 8 times, but pressed 3 times the wrong WiiMote's button.

Volunteer 6 typed an address with 27 letters (no errors). The system suggested some streets based on previous interaction, and volunteer selected one. Due to the selected street is not the desirable one, the volunteer edited the address clicking in the text. In the first time, the volunteer selected a letter. She tried again just to put the text cursor in the correct place. After, the volunteer completed the address with 10 letters (the volunteer pressed 14 keys, due to press 2 wrong keys and need to erase them pressing the backspace button 2 times). Similar to volunteer 1, the wrong pressed keys were next to the desirable key: volunteer pressed 'k' key instead of 'l' key; and volunteer pressed 'u' key instead of 'i' key.

About Task 2 - show the route from your home to college, the volunteers needed to do map navigations using a combination of click and hold the left button over the map and moving the hand in the desire direction. Volunteer 1 did this combination to move the map 23 times, volunteer 2 did 8 times, volunteer 3 did 81 times (more 2 errors), volunteer 4 did 7 times, volunteer 5 did 25 times and volunteer 6 did 38 times. Some volunteers (1, 4, and 5) used the zoom features to perform the Task 2. Volunteer 1 tried to trigger the zoom feature 10 times (4 hits and 6 errors), volunteer 4 tried 4 times (4 hits) and volunteer 5 tried 4 times (3 hits and 1 error).

To perform Task 3 - Join Google Street View and show the facade of your home, volunteers needed to trigger the *Pegman* button, hold it and release it in the map. Before it, the volunteer could be some adjustment in the map view. After, volunteers needed to navigate to find the façade of her home. Volunteer 1 tried 16 times until trigger the *Pegman* function, volunteer 2 tried 4 times (3 errors), volunteer 3 tried 2 times (the first one was considered an error due to the volunteer released in the wrong place), volunteer 6 hit it in the first time.

To navigate in the Google Street View, volunteers 1 and 3 did not need to navigate, and volunteer 2 triggered 5 times (2 hits and 3 errors). Volunteer 6 hit the trigger in the first time (1 hit and 0 error). About the view position, volunteer 1 triggered 5 hits, volunteer 2 triggered 4 hits, volunteer 6 did 2 hits, and volunteer 3 did not need to change the position's view.

To complete Task 4 - Show the route from your home to the nearest market, volunteers needed to navigate into the virtual world and change the position's view. About the navigation, volunteers 1 and 3 used the direction pad of the WiiMote, pressing 1 and 41 times, respectively. Volunteer 2 tried to change the position 15 times (4 errors), and volunteer 6 tried 20 times (2 errors). The total of trying was 35 (29 hits and 6 errors), without count the use of direction pad. To change the position's view, volunteer 1 hit 4 times, volunteer 2 hit 6 times, volunteer 3 hit 4 times and volunteer 6 hit 4 times. The total of hit was 18 times. No error was identified.

Table I summarizing the collected data. Volunteers 1, 2, 5, and 6 tried to trigger one of the zoom buttons (both have the same size, so we computed together) in Task 1, and volunteers 1, 4, and 5 in Task 2. They tried zooming 57 times, and performed 31 hits and 26 errors. Volunteers 1, 2, and 6 used the virtual keyboard in their interactions, pressed 80 keys: 68 hits, 6 errors and more 6 to erase the wrong typed letter. To trigger the *Pegman* button, volunteers tried 23 times (4 hits). Trigger the zoom buttons can be considered with low efficacy due to the half of the time the users do an error; but trigger the virtual keyboard's key had a better efficacy. We believe the cause is the buttons' dimensions. About map navigation, volunteers did 182 times the combination of pressing the button and moving the hand and made 2 errors.

TABLE I. HITS AND ERRORS IDENTIFIED IN VOLUNTEERS' INTERACTIONS

	Command				
	Trigger zoom in and zoom out buttons	Typing in the virtual keyboard	Trigger the Pegman button	Navigate into the virtual world	Change the position view
Trials	57	80	23	41	29
Hits	31 (54.38%)	68 (85.00%)	4 (17.39%)	32 (78.05%)	29 (100%)
Errors	26 (45.62%)	6 + 6 (15.00%)	19 (82.61%)	9 (21.95%)	0

C. Case Study 2: Navigation on TelEduc

Online systems that support e-Learning through the Web are called e-learning environment systems or Virtual Learning Environments (VLE) or Learning Management Systems (LMS). An e-learning environment system is an application that uses the Web infrastructure to support teaching and learning activities, designed to support a variety of users and learning contexts. This environment is composed of tools that allow users to create content,

communicate with other users, and manage the virtual space, e.g., chat, forums, portfolios, and repositories. One example is TelEduc, a Brazilian e-learning environment.

Due to the amount of tools available at TelEduc, we chose 6 tools (mail, agenda, pool, support material, portfolio and message board) and defined nine tasks to users perform: i) login, ii) read mail messages, iii) delete mail messages, iv) read the course's agenda, v) vote in a pool, vi) watch a video, vii) post a portfolio item, viii) publish in the message board, and ix) logoff. Questions that composed the opinion questionnaire were about the difficult to: i) read text in the screen, ii) trigger icons such as navigator icon, iii) trigger the scroll bar, iv) the menu, v) the keys of the virtual keyboard, vi) put the keyboard focus on the text field, and vii) trigger small buttons, such as the minimize button.

We performed a pre-test to evaluate our questionnaire for the second case study. Analyzing the recorded interaction, we found difficult to trigger small buttons and links in words, to put the mouse pointer in a specific place and trigger the scroll bar due to the small size of its buttons and its width. Another problem happens when the user will press the button A: the mouse pointer moves. This problem was identified in another user test; but, in this case, the problem happen on trigger links: the user needs to try many times to reach her goal.

IV. FINAL CONSIDERATIONS

From our previous work [8], we can perceived that using motion sensor to navigate on Web page is fun and the users might feel motivated to adopt this interaction modality. In this paper, we presented the analysis of the recorded interactions, and identified that volunteers had difficulties to trigger small buttons and to typewrite text. This analysis confirms our hypothesis that adaptations are necessary to users have a good experience when navigating through web pages using a motion sensor. We believe that to minimize the number of interaction problems that happen due to the size of the user interface elements or the high sensitivity of the control (performing pointer movement when pressing the button) is necessary to do some adjustments, such as increasing the width and the height of the elements. We plan to investigate solutions for the identified problems and develop tools that aim to collect and analysis the generated data in similar case studies.

As future work, we plan to continue our study by analyzing the interaction using a Kinect movement sensor instead of WiiMote. To compare the data collect in the cases study, we plan to apply the same questionnaires and develop guidelines for design Web applications which users can access them with keyboard+mouse or motion sensors.

ACKNOWLEDGMENT

The authors thank the CNPq and IFSP for financial support.

REFERENCES

- [1] A. C. da Silva, F. M. P. Freire, and H. V. da Rocha, "Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments," Proc. of 6th

- International Conference on Advances in Computer-Human Interactions (ACHI 2013), IARIA, Feb. 2013, pp. 243-249.
- [2] J. Nielsen, *Usability Engineering*, San Francisco, CA: Morgan Kaufmann Publishers Inc., 1993.
- [3] Google Inc., Google Maps. Available from <http://maps.google.com/>.
- [4] TelEduc Project. Home | teleduc.org.br. Available from <http://www.teleduc.org.br/>.
- [5] L. Nigay and J. Coutaz, "A Generic Platform for Addressing the Multimodal Challenge," *The SIGCHI Proceedings of the 13th Conference On Human Factors in Computing Systems (CHI 1995)*, ACM Press, May. 1995, pp. 98-105.
- [6] B. Dumas, D. Lalanne, and S. Oviatt, "Multimodal Interfaces: A Survey of Principles, Models and Frameworks," *Human-Computer Interaction*, D. Lalanne and J. Kohlas, Eds. Berlin: Springer Berlin / Heidelberg, pp. 3-26, 2009.
- [7] N. O. Bernsen, "Multimodality Theory," *Multimodal user Interfaces: From signal to interaction*, D. Tzovaras, Ed. Berlin: Springer, pp. 5-28, 2008.
- [8] A. C. da Silva, A. L. C. Viana, and D. Marques, "Identifying Interaction Problems on Internet Navigation Caused by Change of Input Mode: a study about motino sensor controller, Google Maps and Google Street View," *Proc. of 8th International Conference on Interfaces and Human Computer Interaction 2014, Game and Entertainment Technologies 2014 and Computer Graphics, Visualization, Computer Vision and Image Processing 2014, IADIS*, Jul. 2014, pp. 271-275.