

How to Design an User Interface Based on Gestures?

Anna C.S. Medeiros¹(✉), Tatiana A. Tavares²,
and Iguatemi E. da Fonseca¹

¹ Centro de Informática, Universidade Federal da Paraíba, João Pessoa, Brazil
linnamedeiros@gmail.com, iguatemi@ci.ufpb.br

² Laboratório de Aplicações de Vídeo Digital (LAVID), Centro de Informática,
Universidade Federal da Paraíba, João Pessoa, Brazil
tatiana@lavid.ufpb.br

Abstract. The use of our body language to communicate with computer systems is an increasingly possible and applicable feature in the real world. This fact is intensified by the evolution of gesture recognition based commercial solutions. A gesture interface complements or replaces navigation in a conventional interface, it is up to each developer to choose the most appropriate option for their application. When opting for gesture usage, the gestures will be responsible to activate the systems functions. This work presents a gesture development process that can be used to aid the construction of gesture interfaces. The process here described, should help interface designers to incorporate gesture-based natural interaction into their applications in a more systematic way. To illustrate the Process, gestures for the actions “Select”, “Rotate”, “Translate”, “Scale” and “Stop” were developed.

Keywords: Gestures · Process · Gesture interface · Natural interaction

1 Introduction

During the study and research on ways of interaction between humans and electronic devices, using their body as a means, focus of Natural Interaction (NI) study, it was found that there are several possibilities of interaction: through voice commands, haptic devices, olfaction, locomotion, gestures or detection and identification of human body parts such as face, hand, thumb, eye retina [1].

Human-Computer interaction (HCI) studies the communication between people e computational systems, it is situated in the intersection between Behavioral Sciences, Social Sciences and Computational and Information Sciences, involving all aspects related to the interaction between users and systems [2]. The research on HCI stimulated to investigate how to use NI as an alternative to the conventional keyboard and mouse complex, with the intention to make the interaction more simple and intuitive.

Technology is used daily by almost everyone, software solutions are used to perform lots of tasks, therefore simplicity becomes necessary to ease interaction with systems and leads to an easier and more sustainable relationship with the media and technology [3].

With the crescent research in NI equipment, such as Kinect (launched in 2010) and Leap Motion (launched in 2013), accuracy in movement detection is improving, thus providing more possibilities in the use of such hardwares. Older NI equipment such as Kinect, have very affordable prices now, which can create another range of possibilities, for example, its use in public schools.

There are studies that focus on perceptual technology and software development, disregarding elements relative to the determination of gestures, for example: [4], which despite being of great importance to the area, does not seek to understand gestures in search of the most appropriate choice for the interfaces created, which can lead to using tiresome, unintuitive or non-functional gestures and can harm the performance of the application.

The process of associating a gesture to a particular action or function of the system is not trivial, because it must take into account a number of factors such as ergonomics, intuitiveness and objectivity [5, 6], which highlights the importance of a Gesture Development Process (GDP). The Process below described aims to help the development of gestures for user interfaces.

2 Describing the Gesture Development Process

The present research shows the construction of a Process that can assist in deciding which gesture would better represent a certain function (action) of a system. The main references used in developing this process can be seen in [7, 8]. This GDP produces artifacts during its execution, the final product of this Process is a Template depicting the finalist and best gestures for the given functions.

The GDP here described has three stages (Fig. 1). In the first stage, “Define Functions”, the problem domain is analyzed and the functions of the application to be developed are determined. This stage is not concerned with any kind of gesture, only with determining the functions to which you want to assign gestures.

To better the objective of the first stage, understand observe Fig. 2, this image was taken from the game Fruit Ninja [9], in this game players should “slice” fruits as they jump into the screen, on XBOX360 platform, the Kinect device captures user’s gestures in order to make the interaction with the game possible. In this context, the main function of the game is “slice”, but in the game’s menu other functions such as “select” and “cancel” can be observed. So as to exemplify, in the first stage the developer should define which of these functions it would be desired to be triggered by gestures.

The second stage consists of three steps. The first step, “Apply Test Scenarios”, needs volunteers in order to perform tests to generate prototype-gestures for the functions determined in the previous stage. One or more test scenarios, that abstract technical thinking and at the same time contextualize the functions to stimulate the volunteers to execute them, are created. All the interactions at this stage are recorded, the number of participants should be directly proportional to the diversity of the gestures obtained for each function, in other words, as the tests are applied, the more diverse are the gestures for each function, the more volunteers are necessary.

During the second step, “Analyze and Register Recordings”, an artifact is created containing the analysis and registers of the identified gestures in the recordings. In this

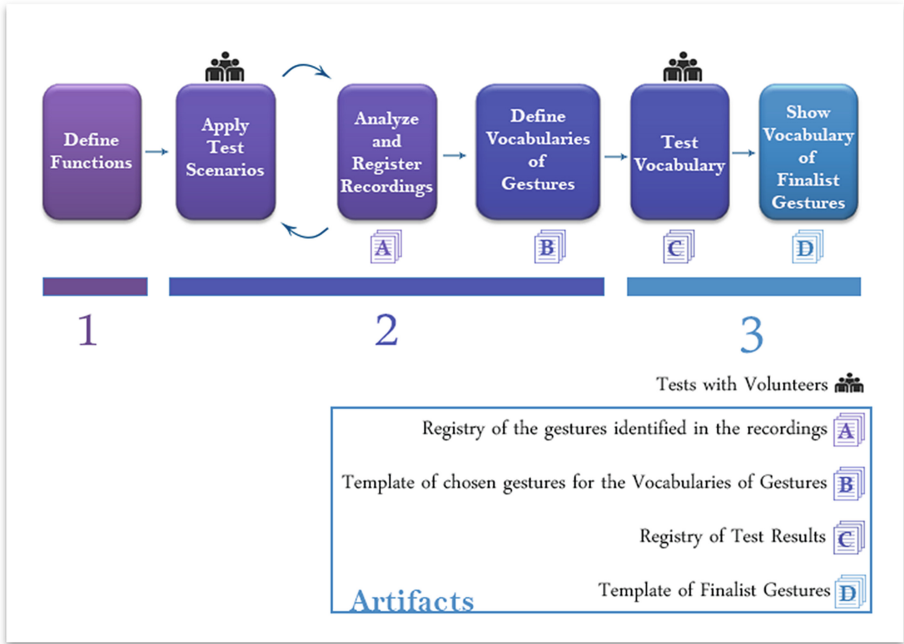


Fig. 1. Illustration of the process here presented



Fig. 2. Fruit Ninja

step a decision must be taken: if the results are not considered sufficiently favorable, it is necessary to go back to the previous step, but now with prototypes-gestures as final suggestions for volunteers, in order to refine the gestures development process. Otherwise, go forward to the third step.

In the third step, “Define Vocabularies of Gestures”, the most favorable gestures are determined, based on the previously produced artifact, taking into account the most frequent, the ones chosen by the designer and those whose movement and posture do not compromise physical structures and articulations. A template of the chosen gestures is then generated to be used in the next stages.

It is worth mentioning that one can determine various gesture-vocabularies. At the end of this second stage two artifacts should have been produced: the registers of the identified gestures in the recordings and the template of the chosen gestures.

The third and last stage has two steps, In “Test Vocabulary” the process evaluates the N gesture vocabularies through three tests: Attribution of Semantics, Memory and Stress Tests, as described in [7]. Each test has a score and at the end it will be possible to decide the best vocabulary through the score comparison. In the case of only one vocabulary the score it will be used to infer its quality.

In the following tests, the N vocabularies obtained are tested separately, the same volunteers evaluate all vocabularies, it is important to vary the vocabulary with each volunteer will start evaluating, in other words, in the case of being two vocabularies, a volunteer will first evaluate the vocabulary A then B, and other volunteer would start with B and later evaluate A, trying to make the numbers of volunteers starting with A equal to the ones starting with B.

The Attribution of Semantics Test consists in presenting to the volunteer a template of the Gesture Vocabulary (GV, meaning a group or list of gestures) and a list of functions, however it's not revealed the correspondence between the gestures and functions. It is then asked for the volunteer indicate which gesture corresponds to each function. The score is the sum of wrong guesses, divided by the number of gestures.

The Memory Test measures the familiarity of the user with the gesture. Only after the user is acquainted with the gestures is that he will be able to maintain the focus in the task at hand rather than to how operate the interface. First it is shown a GV to each volunteer, later, a slideshow with all the functions, staying 2 s in each function. The volunteer is then asked to do the correspondent gesture of that function, until being able to get all functions right, when mistaken the presentation is restarted and the vocabulary is revised. The score of the memory test is calculated by the number of restarts necessary, resulting in a measurement of the difficulty for a new user to become familiar with the gesture vocabulary.

The Stress Test shows the volunteers a sequential list of gestures. Each volunteer should repeat the sequence a determined number of times, enough so that he can infer any possible discomfort. At the end of the test, each volunteer is questioned about how stressing is each gesture, giving a general classification for them: “no problem”, “slightly tiresome”, “annoying”, “painful” or “impossible”.

At the end of all tests, in “Show Vocabulary of Finalist Gestures”, the vocabularies can be compared through their scores and classifications, then it would be drawn a final template for the gestures of the chosen vocabulary, so as to aid the implementation of these gestures to their respective functions on the application.

3 Applying the Gesture Development Process

To demonstrate the application of the GDP described above, the idea was to develop gestures for common functions in 3D environments, the chosen functions were “Rotate”, “Scale” and “Translate”, as these are the main two-dimensional and three-dimensional geometric transformations [10], and “Select” and “Stop”, because they are related to the beginning and end of interactions in general.

The first stage is to select which function will receive gestures as triggers. In this case; “Rotate”, “Scale”, “Translate”, “Select” and “Stop”.

In the “Apply Test Scenarios” step of the second stage, a case scenario application was created to stimulate the volunteers in the production of gestures for the functions mentioned before (Fig. 3). The Cube application featured a colorful cube that could be selected, rotated, translated, scaled and stopped (when stopped, no interactions could be performed with the cube), all through the mouse and keyboard input. The idea was that the volunteers could observe this interactions, then they were asked to perform gestures that, in their minds, would seem intuitive to trigger the before functions. This test was first carried out with 12 volunteers, all students from Computer Science bachelor, all tests were recorded.

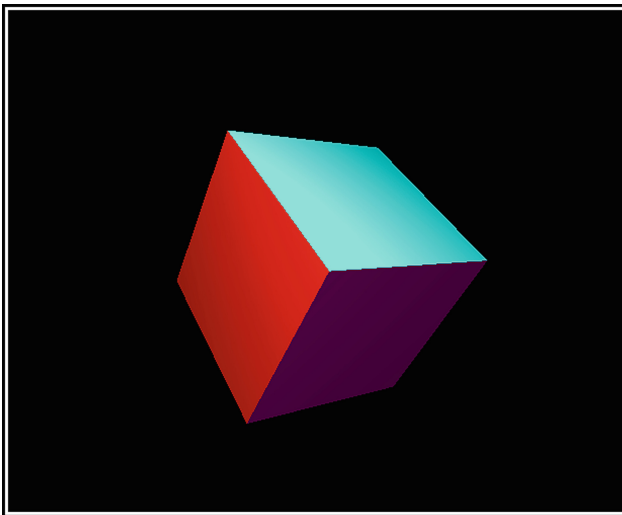
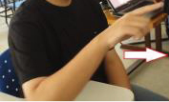

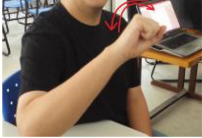


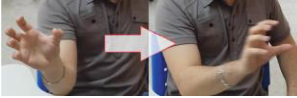





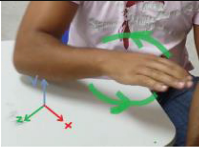


Fig. 3. Cube application


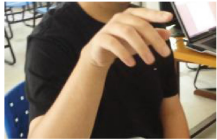









With the recordings from the first round of tests, it was time to proceed to the second step, “Analyze and Register Recordings”. Here the recordings were analyzed and the gestures captured were registered as shown in the Tables 1 and 2. The most popular gestures were S1, R3, T2, E1 and P2. The decision here was made in favor of another round of tests, in order to see if there would be many more new gestures or if 12 volunteers were enough.

Table 1. “Select”, “Rotate”, “Translate”

Select	Rotate	Translate
 “Click” on the object S1	 Move hand around R1	 Move hand around T1
 “Click” on the object S2	 R2	 T2
 “Capture” the object S3	 R3	 Point to the object and then point to where it should be moved T3
	 R4	 With hand facing the floor, move hand around freely T4
	 With hand facing the floor, move hand around in X and Z R5	

The first step was then repeated with another 12 volunteers, also students from Computer Science, only now the tests would provide the gestures in Tables 1 and 2 as options at the end of each test, so that the volunteer could change its choice if he preferred one of the gestures presented.

Table 2. “Scale”, “Stop”

Scale	Stop
 <p data-bbox="164 388 517 502">Join fingers to close in and open hand to close out Move arm as well E1</p>	 <p data-bbox="635 414 952 467">“Click” away from the object P1</p>
 <p data-bbox="158 626 517 714">Only moves arm to indicate close in and close out E2</p>	 <p data-bbox="776 684 811 714">P2</p>
 <p data-bbox="317 846 352 878">E3</p>	 <p data-bbox="776 855 811 885">P3</p>
 <p data-bbox="164 1054 517 1173">Join fingers to close in and open hand to close out Do not move arm E4</p>	 <p data-bbox="776 1090 811 1120">P4</p>
 <p data-bbox="317 1323 352 1354">E5</p>	 <p data-bbox="676 1314 911 1375">“Click” on the Object P5</p>
 <p data-bbox="711 1525 876 1587">Put hand down P6</p>	


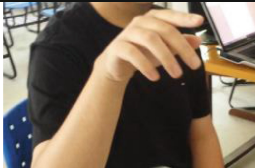
Going forward to step 2, the registry showed one new gesture for “Select”, two new gestures for “Rotate” and “Translate”, and three new gestures for “Scale” and “Stop”. The most popular gestures now were still S1, R3, T2, E1 and P2. This was a good indicator that the first 12 volunteers produced a good variety of gestures, yet it was decided to do a last round of tests with students from different areas, to see if the results would be alike.

Back in to the first step another 12 volunteers repeated the last test, but now with more gesture options at the end of each test. The volunteers from this third round of tests were students from several majors including: Electrical Engineering, Law, Public Administration and Medicine. Going forward to step 2, the registry showed two new gestures for “Select”, “Rotate” and “Scale”, four new gestures for “Translate”, and no new gestures for “Stop”. The most popular gestures with this last group of volunteers were S1, R3, T2, E4 and P2. But in the total 36 volunteers the most popular gestures were still S1, R3, T2, E1 and P2. Seeing as that from 12 to 36 volunteers only one gesture changed (E1 to E4) the results were satisfactory and it was decided to proceed to the next step “Define Vocabularies of Gestures”.

It was selected two gestures vocabularies; one containing: S1, R3, T2, E1 and P2, the most popular choice between the 36 volunteers, and another containing: S3, R7 (Table 3), T1, E4 and P8 (Table 3), less popular choices but considered good to be compared to the first vocabulary, in the next stage. The template produced in this third step can be seen in Tables 4 and 5.

In the third stage, Attribution of Semantics, Memory and Stress tests were performed with Gesture Vocabularies 1 (Table 4) and 2 (Table 5). This stage counted with 19 volunteers, students from several majors: Civil Engineering, Medicine, Computer Science, Electric Engineering, Law, among others.

Table 3. Gestures R7 and P8

Rotate	Stop
	
R7	<p data-bbox="595 1238 992 1294">“Click” on the object, so that if its selected, it will stop any future interactions with it</p> <p data-bbox="777 1330 809 1354">P8</p>

In the Attribution of Semantics test the templates (Tables 4 and 5) and a list with the functions before commented were presented to each volunteer. It was then asked for the volunteers to indicate which gesture corresponds to each function. The results for each GV can be seen in Table 6. Bear in mind that the score is the sum of wrong guesses divided by the number of gestures, so a perfect score would be zero.

Table 4. Gesture Vocabulary 1



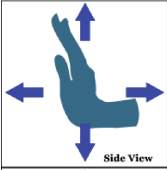
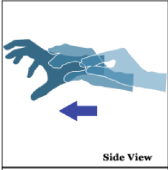
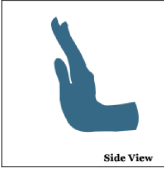


Select	Rotate	Translate	Scale	Stop
 <p>Top View</p>	 <p>Side View</p>	 <p>Side View</p>	 <p>Side View</p>	 <p>Side View</p>
	 <p>Side View</p>			
	 <p>Side View</p>			

Table 5. Gesture Vocabulary 2

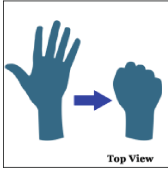

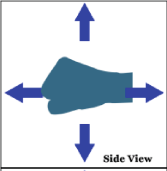
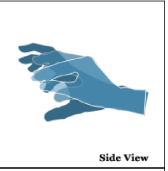



Select	Rotate	Translate	Scale	Stop
 <p>Top View</p>	 <p>Top View</p>	 <p>Side View</p>	 <p>Side View</p>	 <p>Top View</p>
	 <p>Top View</p>			
	 <p>Top View</p>			

Table 6. Results from attribution of semantics test

	GV 1	GV 2
Average score	1,4	1,8
	In 19 volunteers there were a total of 7 wrong guesses of the 5 functions	In 19 volunteers there were a total of 9 wrong guesses of the 5 functions

The Memory test showed to each volunteer the GVs 1 and 2 and a slideshow with one slide for the name of each function (“Select”, “Rotate”, “Translate”, “Scale” and “Stop”)., the slideshow would linger 2 s on each slide. When reading the name of the function in a slide, the volunteers were asked to do the correspondent gesture of that function, if they did a wrong gesture the presentation was restarted.

The score of the memory test was calculated by the number of restarts necessary until the volunteer got all gestures correct. This was carried out twice for each volunteer since there was two GVs. The results can be seen in Table 7. The best result would be zero.

Table 7. Results from memory test

	GV 1	GV 2
Average score	1	2
	1 restarts out of 19 volunteers	2 restarts out of 19 volunteers

The Stress Test showed the volunteers a sequential list of the gestures from each GV. Each volunteer was then asked to repeat the sequence 50 times, for each gesture in each GV. At the end of the test, each volunteer was questioned about how stressing was each gesture, giving a general classification for each of the gestures in both GVs: “no problem”, “slightly tiresome”, “annoying”, “painful” or “impossible”. The result can be seen in Tables 8 and 9.

Table 8. Results from stress test - GV 1

	Select	Rotate	Translate	Scale	Stop
“no problem”	19	15	19	10	19
“slightly tiresome”		4		8	
“annoying”				1	
“painful”					
“impossible”					

Table 9. Results from stress test - GV 2

	Select	Rotate	Translate	Scale	Stop
“no problem”	15	10	18	16	19
“slightly tiresome”	4	5	1	3	
“annoying”		4			
“painful”					
“impossible”					

In the end the Gesture Vocabulary 1 was slightly better than the GV 2. As this work is still in progress the next step will be to implement these gestures in the Cube application (Fig. 3), using the Leap Motion device and with aid of the final template (Table 4).

4 Final Considerations

The interaction style supported by Natural Interaction devices, such as Leap Motion and Kinect, has a wide variety of potential applications, well beyond those described in this paper. When the interaction style is based on gestures, gesture design and development is among the issues to be addressed.

This paper presented a strategy to guide designers throughout the user interfaces based on gestures process. Nowadays, gestures are a realistic solution for user interfaces. Devices like Leap Motion and Kinect make it easier to develop this kind of user interfaces. The challenge is how to design them. This is the main goal of this paper.

In order to do that, the GDP depicted here presented a guide with all the steps to identify and specify gestures. Also, verification techniques are presented. A conceptual proof of the proposed strategy was verified using the Cube application. Also, a prototype using the Leap Motion device was implemented.

So, we turn back to our initial question: How to Design an User Interface Based on Gestures?

We are still thinking of ourselves as working stations, primarily based on computers devices. But things are changing and the physical world's functionalities are devices to access computer's functionalities. This paper approached our first big step in this direction: the gestures design. So, we emphasize the point of view of Pierre Wellner: "Instead of making us work in the computer's world, let us make it work in our world." [11].

References

1. FUTURELAB. Interação Natural (2014). <http://www.futurelab.com.br/site/interacao-natural/>
2. SBC. Sociedade Brasileira de Computação (2014). http://www.sbc.org.br/index.php?option=com_content&view=category&layout=blog&id=45&Itemid=66
3. Valli, A.: The Design of Natural Interaction (2006)
4. Almeida, F.B.: Sistema Interativo Baseado em Gestos para Utilização de Comandos no Computador. UFPE (2013)
5. Bacim, F., Nabyouni, M., Bowman, D.A.: Slice-n-Swipe: A Free-Hand Gesture User Interface for 3D Point Cloud Annotation. Virginia Tech, Minnesota (2014)
6. Jeong, S., Jin, J., Song, T., Kwon, K., Jeon, J.W.: Single-Camera Dedicated Television Control System using Gesture Drawing. Sungkyunkwan University, Suwon, South Korea (2012)
7. Nielsen, M., Störing, M., Moeslund, T.B., Granum, E.: A procedure for developing intuitive and ergonomic gesture interfaces for HCI. Aalborg University, Aalborg (2004)

8. Choi, E., Kwon, S., Lee, D., Lee, H., Chung, M.K.: Towards successful user interaction with systems: Focusing on user-derived gestures for smart home systems. Pohang University of Science & Technology (2014)
9. HALBRICK STUDIOS. The Greatest Fruit-Slicing Game in the World! (2010). <http://fruitminja.com/>. Accessed 2014
10. Foley, D., van Dam, A., Feiner, S., Hughes, J.: Computer Graphics Principles and Practise. Addison-Wesley Publishing Company, Reading (1997)
11. Wellner, P.: Interacting with paper on the DigitalDesk. Commun. ACM Special Issue Comput. Augmented Environ. Back Real World **36**(7), 87–96 (1993)