



Gesture Elicitation for Augmented Reality Environments

Francisco Vinicius Nascimento da Silva¹(✉),
Francisco C. de Mattos Brito Oliveira¹, Robson de Moraes Alves¹,
and Gabriela de Castro Quintinho²

¹ Ceara State University, Avenue. Dr. Silas Munguba, 1700 - Itaperi,
Fortaleza, CE, Brazil

{vinicius.silva,fran,robson.moraes}@dellead.com,
viniunascimento1@gmail.com

² PUC - RS, Porto Alegre, Brazil
grabriela.quintinho@dellead.com

Abstract. The study of gesture elicitation in Human-Computer Interaction (HCI) aims to improve the interaction with the computational environment, providing improved usability and a consequent increase in performance in executing their tasks. In environments using augmented reality, gestures are preferable over eye gaze, as, in the former, the user does not take the focus off the field of work. However, unnatural motions are more difficult to memorize, and their use puts any gains from using the technique at risk. Thus, it is necessary to elicit gestures to discover which incorporated movement appears naturally when the user is asked to send a specific command to the interface. This work aims to elicit natural gestures for an augmented reality environment. During the elicitation process, research subjects were presented with stimuli to which they should react by performing gestures that represent commands for the interface in an augmented reality environment, such as, for example, moving forward, backward and changing zoom (scale). As a result, 434 movements were found, with an average of 18 movements per person, where nine actions for the right hand and nine for the left hand were separated. In addition, the movements made by the participants were categorized into bi-manuals, those that were worked with both hands to carry out the interaction, uni-manual movements, which is when the participant used only one hand, and symmetrical bi-manual, when the participant performed a move with one hand that depended on the other hand to complete the interaction. After these analyzes of individual movements, an average of 3 movements considered natural for each interaction were obtained, in which the researchers chose the two most performed by each participant; each movement resulted from an analysis carried out considering the naturalness in which the gesture was performed and also the ease that the participant had to do so. The gesture chosen for each command was the one whose category had the highest number of natural occurrences, the one that had the most performed category by the participants.

1 Introduction

Since the studies initiated by Jacob O. Wobbrock, the elicitation of gestures is a technique that allows the collection of input preferences, in their symbolic value, by end-users [22]. This method makes it possible to discover between the input of manual gestures in the computational context, which acts as a popular tool for exchanging information between users and a [19] interface.

The use of gestures proves to be beneficial in performing tasks, especially when the stimulus for their execution is linked to the daily life of users, providing more enriched information, especially when the task requires maintaining focus on the activity while using the application. This is particularly important for Augmented Reality (AR) cases, where movements are performed in an embodied [21] manner.

Gestures have great importance in human communication, as they are considered an extension of their expressions. The movement of the hands during a person's speech, in most cases, presents propositional aspects to the true meaning of the speech that the individual is performing [8]. Human gestural communication can be characterized as natural interaction with its speakers. Each gesture performed can be considered idiosyncratic; for some authors such as [7] this happens due to the state of each person in each person is submitted to the execution of these gestures, such as anger, love, lie, truth, among other feelings.

This work aims to elicit gestures in augmented reality environments. The process of discovering gestures must occur naturally so that each person can learn the proposed gestures to interact with virtual objects such as rotation, translation, zoom. Gestures learned naturally facilitate their identification to carry out the tasks proposed for each user. The elicitation process consisted of experimenting with 25 participants, aged between 18 and 40 years and by male and female gender, 15 men and ten women. The research subjects performed the gestures that best expressed their intentions during the tests and made them aloud. For example, the researcher asks which gesture the participant would perform if he wanted to change the size to scale of a given 3D object immersed in the AR environment, the participant, in turn, would perform a gesture that he thought was appropriate for that proposed interaction, which we call elicitation. Open, where participants had complete freedom to perform any movements that were relevant to them.

In the following section, information about gestures in augmented reality and accessibility environments is shown. In Sect. 3, the theoretical basis for the construction of the work is demonstrated. Section 4 introduces related works. This work methodology is presented in Sect. 5. Section 6 refers to data collection. Data analysis is discussed in Sect. 7. We present the gestures elicited in section eight. The ninth and last section is reserved for discussions and future work.

2 Gestures and AR

The study helped identify two new parametric categories for classifying gestures. Exclusive to AR: while [23] featured gestures in terms of Shape, Nature, Connection, and Flow. In some studies, the authors observed that the size of the AR object affected the number of hands used to manipulate the objects (e.g., for palm-sized or smaller objects, just one hand; for larger objects, two hands) [12].

One result of an elicitation study is the production of a mapped set of inputs called a consensus set ([1], and [2]). Most More useful than a single set of mapped inputs is observational data that comes from elicitation studies. This includes insights into the formation of inputs, the times surrounding input generation, and trends in user preferences for inputs and input modalities. An example of these extended benefits is the finding that the size of a proposed gesture is impacted by the size of the displayed object [16]. This work extends previous studies of gesture elicitation in AR [13] by testing additional modalities of isolated speech and multimodal interactions of gesture and speech and allowing gesture proposals for each referent. Furthermore, the set of interactions presented here shows the main proposals allowing a better interpretation of trends in the formation of the gesture [21].

3 Theoretical Foundation

3.1 History of Gestures

Among the fundamental characteristics of human language, learning through interaction is transmitted over generations and varies according to the types of existing human societies. For some authors found in the literature such as [14], the origin of gestures are discussed with four proposed hypotheses, namely: phylogenetic ritualization, which is a process in which communication is displayed as a sign of dominance, which may emerge due to the lack of movements body as they are borrowed from other contexts which are where a communication signal is created by two individuals shaping the behavior of both in repeated instances of interaction over time, ontogenetic ritualization which refers to situations where observers acquire parts from the gestural representation of copying the gestures that are directed to them, social learning by imitation and social learning through negotiation that takes into account the shared understanding of gestural meaning originated from a mutual construction in real-time for both members.

However, speech extends beyond what can be considered an expression of content, a social form of action where gesture plays an important role. One of the first psychological approaches to the modern study of gestures can be found in the work of David McNeill *“Do you think gestures are nonverbal”* [6]. In this work, David McNeill approaches the conceptual content aspects of the utterance, observing through an examination the hand movements considered as components integrated into the utterance produced.

Thus, gestures are an indecomposable part of speech, where McNeill recognized that people, when gesturing, express what he called pragmatic content; that is, gestures are always in conjunction with speech. McNeill further considered that gestural movements conceive a kind of mental image form of the components that people are gesturing about in their speech and conclude in his book *“Hand and mind: What gestures reveal about thought”*, that gestures reveal about thinking, involving cognitive processes in the production of utterances rather than considering gestures as the final component of an utterance in the process of social interaction [7].

3.2 Gesture Elicitation

GES emerged from the need for the importance of understanding different devices and environments; this premeditates the increase in the user’s ability to interact with more complex environments [11]. Some studies found in the literature demonstrate that gestures can be elicited by experts such as [3, 20], while others demonstrate precedence for gestures created by the experiment users themselves, such as the works of [1, 10, 23].

This study aims at mapping the input of technologies through an interactive design, in which these inputs must be discovered by the users of the [21] systems. GES aims at a better understanding of users’ behavior, as, for example, gestures performed on the upper parts of the body are preferred; GES has also seen use for various input domains where it can be found on multi-touch surfaces. As in [1] and mobile devices in [15].

3.3 Concepts and Terminology

Figure 1 shows how the gesture elicitation process takes place. The experiment is conducted by each participant individually, where sessions can last up to 20 min. Participants are submitted to CAD modeling tasks, images, or videos, and later their actions will be captured and analyzed by external researchers.

Some terminology about GES found in the works of [5], and [19] will be shown below.

- Navigation: Navigation refers to the tasks of moving the participant’s viewpoints. The user can perform tasks like zoom and orbit. The participants are positioned so that their field of vision is in a panoramic way in relation to the interactions that are submitted;
- Manipulation: This term refers to methods in which an object’s parameters are changed, such as positioning, orientation, and scale. The categories that were denoted in Fig. ?? were rotated, resized (scale), copied, and moved. For this technique, the participants also positioned themselves so that their vision was in panoramic mode;
- Primitives: In this group are concentrated forms free of manipulation, such as cone, prism, and box. Usually, in this group, static images are shown to the participants. The term freeform is used to refer to 2d objects that contain lines and curves as components;

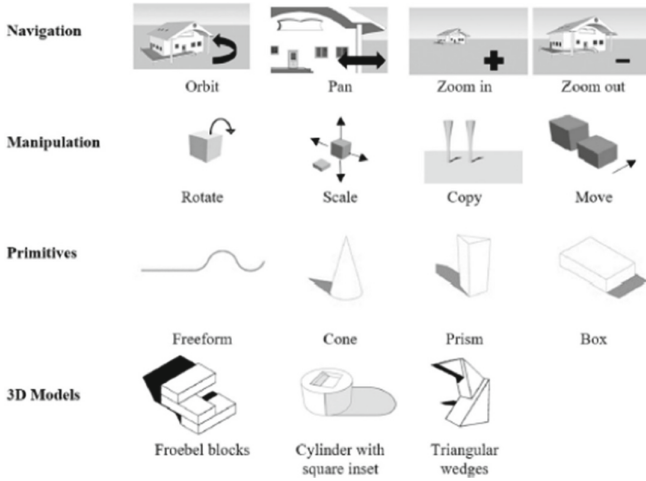


Fig. 1. Methods for the process of eliciting gestures - available in [5]

- Models: Refers to the objects used during the execution of the experiments with the participants. Models can be in either 2D or 3D format and serve as a method of interaction between the interface and users;
- Agreement: Occurs when you have a situation where the gestures of two or more participants are evaluated as being identical according to the rules, criteria, or similarity function. For example, the researcher may consider that the direction of the gesture and the speed are important factors for evaluation in the process of eliciting gestures;
- Agreement Rate: It is the numerical measure that has the function of quantifying the agreement between the elicited gestures of the participants. Example: Of the 20 study participants, three subgroups of sizes 9, 7, and 5 appear for the “Dim illuminates’ so that all participants in each subgroup They Agree. The agreement score [22] calculates $(9/20)^2 + (7/20)^2 + (5/20)^2 = 388$ and according to [18], $(9 * 8 + 7 * 6 + 5 * 4) = (20 * 19) = 353$;
- Command: It is the action that activates the execution of the functionalities in the user interface. It can also be known as a gesture command, where the gesture is represented as an input action for interactions. For example, the researcher asks the participant which gesture can be used to rotate a 3d object, the participant, in turn, would propose a gestural movement that he/she found convenient to perform that task;
- Consensus Rate: A measure of agreement found in [17] that employs dissimilarity and tolerance functions;
- Set of Gestures: It is the set considered with the highest number of the agreement for the referents proposed to the participants. For example, in an experiment with 20 people, 15 participants concluded that hand rotation, with the fingers pointing forward, rotating clockwise, was considered an appropriate gesture to rotate an object;

- Dissimilarity function: Function responsible for calculating the value that is reflected when eliciting two different gestures found in [17];
- Elicitation: It is the process of provoking participants to respond to referents and proposing new gestures to affect these interactions. The elicitation can be closed where the researchers already have a set of predefined gestures that will be later consolidated with the experiment participants in the process of eliciting these and gestures. And then, there is open elicitation where the researchers do not have this predefined set and let the participants decide for themselves which gesture they will perform during the elicitation experiment;
- End User: Final phase of the gesture elicitation process, it can be considered as an interactive device, application, or system for which the elicited gestures will compose the user interaction part;
- Function: Interface resource used to execute commands and relationship between gestures and referents;
- Gesture: A movement of a part of the whole body used to interact with the environment and respond to the action of referents;
- : Interface feature that can be controlled using an automated command or the execution of an elicited gesture. Such as, for example, rotating a 3D object in the augmented reality environment;
- Symbol: Any action that makes sense to evoke a referent in its function form, for example, air gestures, dash-shaped gestures, voice commands, icons, button labels, menu items;
- Participant: Subject who volunteers to participate in the gesture elicitation experiment.

4 Related Work

Multimodal interaction techniques using gestures and voice offer architects and engineers a natural way to create computer-aided design models. The work of [5] conducted an experiment with 41 participants to obtain gesture and speech preferences for references and making these models. The authors presented a compilation of gestures evaluated by experts as the most suitable for articulation, as well as a set of terms and commands extracted from the participants' speech. And finally, they provide recommendations for the design of a CAD modeling system based on gestures and voice for projects.

The work of [21] establishes a better understanding of syntax choices in the types of multimodal interactions reproduced by users' speech and gestures in augmented reality environments. The paper presents a multimodal elicitation study carried out with 24 participants. The canonical referents for translation, rotation, and scaling were used along with some abstract referents (create, destroy, and select). In this study, time windows for multimodal interactions of gestures and speech are developed using the start and stop times for gestures and speech, as well as trigger times for gestures. Finally, trends in the most common proposals for each modality are examined. They are showing that the disagreement between the proposals is often caused by a variation in the posture

or syntax of the hand. This allows you to present aliasing recommendations to increase the percentage of natural user interactions captured by future multi-modal interactive systems.

The study by [11] looks at how to find a set of gestures for 3D travel using a multitouch screen and a mid-air device to improve user interaction. A user study was carried out with 30 subjects, concluding that users prefer simple gestures to multitouch. Furthermore, the researchers found that the legacy of the multitouch user is carried over by mid-Air interaction. Finally, the work proposes a set of gestures for both types of interaction.

In the work of [9], a gesture elicitation study is carried out with 25 visually impaired users. Quantitative classification analysis is performed with them, and an ideal set of gestures is obtained. Furthermore, typewriting is proposed using which visually impaired users can interact with computers. In this work, an overview of the gesture selection method is presented, and some important facts about ideal gestures are revealed.

What differentiates our research from other works is the fact that we carried out an open elicitation with the participants, guaranteeing their freedom during the experiment and allowing the gestures to be elicited in the most natural way possible.

5 Our Methodology for Gesture Elicitation in RA Environments

5.1 Test Scenario

The Test Scenario consists of setting up an environment that simulates the interface of an application, with the purpose that the user starts from a questionnaire containing the available functions to interact with the AR environment. The functions are FORWARD, BACK, POUNDS, CLOSE, ROTATE, SCALE, AND RECORD. The complete environment contains a smartphone application, virtual reality glasses, and a camera to capture the gestural movements of the participants. As shown in Fig. 2.

With the application in hand, the user will be submitted to the questionnaire where each question concerns the gesture in which the interaction with the application will be used. All questions will be answered in the form of gestures; that is, each answer consists of a gesture performed naturally. During this stage of answers, the process of open elicitation was carried out, where the user has complete freedom to make a gesture that he/she finds plausible to answer the questions. While the user performs the gestures to answer the questions of the questionnaire, his hand movements will be captured where later, after performing a common factor with the other participants, the assignment of each gesture to the functions contained in the AR application. The capture of the hands is done from the top of the participant so that it simulates their vision. Performing this process, the gestures tend to be performed naturally, making the interaction with the application easier making the participant more comfortable and familiarized while using the application.

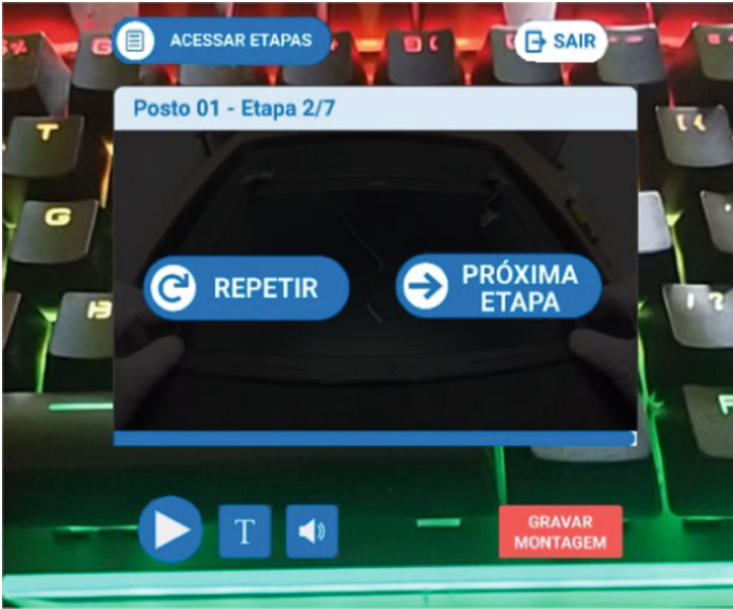


Fig. 2. Application for test scenario and capture of participants' hand movements

The target audience is people with/without knowledge about virtual and augmented reality concepts, as long as they follow the questionnaire didactically. It is not necessary to have extensive knowledge about this area, as the main purpose of this scenario is the process of eliciting the gestures to be used in the AR application. For this, the movements of several hands are captured during the execution of the test scenario.

5.2 Questionnaire for the Gesture Elicitation Process

The applied questionnaire contains questions related to the interaction functions of the AR environment that are present in both video interfaces and 3D models (FORWARD, BACK, POUNDS, RECORD, and CLOSE), where the response form is a gesture performed by the user. Each question consists of a different scenario, where the function intended for the question and the gesture that the user would like to perform to interact with the application are described.

Scenario 01 - Gesture Elicitation for the FORWARD function (Steps) of the AR environment. Question 01: Which gesture would you perform to advance to the next video? Given that: There is a forward function in the application. When: The user performs a gesture that he/she finds coherent for the function to advance. Then: This gesture would be captured.

Scenario 02 - Gesture Elicitation for the BACK function (Steps) of the AR environment. Question 02: What gesture would you perform to go back to the previous video? Given that: There is a back function in the application When: The user performs a gesture that he/she finds coherent for the function to return. Then: This gesture would be captured.

Scenario 03 - Gesture Elicitation for the POUNDS function of the AR environment. Question 03: What gesture would you perform to activate the video pounds function? Given that: There is a function to enable video library in the application When: The user makes a gesture that he/she thinks is coherent for the function to activate video pounds. Then: This gesture would be captured.

Scenario 04 - Gesture Elicitation for the RECORD function of the training carried out within the AR environment. Question 04: What gesture would you perform so that the system could start recording your actions? Given that: There is a function of recording training in the application When: The user performs a gesture that he/she finds coherent for the function of recording training. Then: This gesture would be captured.

Scenario 05 - Gesture Elicitation for the CLOSE function of the training carried out in the AR environment. Question 05: What gesture would you perform so that you could EXIT the augmented reality application? Given that: There is a function to close training in the application When: The user performs a gesture that he/she thinks is coherent for the function of closing training. Then: This gesture would be captured.

Scenario 06 - Gesture Elicitation for ZOOM/SCALE function in 3D objects in the Ar environment. Question 06: What gesture would you perform to zoom in and out of the 3D object displayed on the interface? Given that: There is a pause function in the application. When: The user makes a gesture that he/she thinks is coherent for the zoom function. Then: This gesture would be captured.

6 Data Collection

The elicitation process consisted of carrying out an experiment with 25 participants, aged between 18 and 40 years, and also by male and female gender, 15 men and ten women. During the tests, the research subjects performed the gestures that best expressed their intentions and made them out loud. For example, the researcher asks which gesture the participant would perform if he wanted to change the size to scale of a given 3D object immersed in the AR environment, the participant, in turn, would perform a gesture that he thought was appropriate for that proposed interaction, which we call elicitation. Open, where participants had complete freedom to perform any movements that were relevant to them.

Figure 3 represents a graph with the ages of the participants. Participants were divided into groups of 18–29 years, 30–39 years, 40–49 years, and over 50 years.

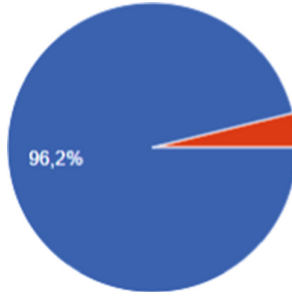


Fig. 3. Age of participants

Figure 4 represents data with participants who have already used devices that work with virtual/augmented reality shown in blue and participants who have never used these devices shown in red.

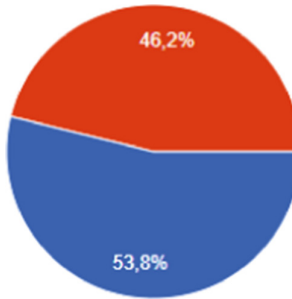


Fig. 4. Use of virtual/augmented reality devices (Color figure online)

As a result, a total of 434 movements were found, with an average of 18 movements per person, where nine movements for the right hand and nine movements for the left hand were separated. In addition, the movements made by the participants were categorized into bi-manual, those that were worked with both hands to perform the interaction, uni-manual movements, which is when the participant used only one hand, and symmetrical bi-manual, when the participant performed a movement with one hand that depended on the other hand to complete the interaction.

7 Data Analysis

After completing the questionnaire and capturing the participants' hand movements, the videos were submitted on situated analysis techniques added to the Elan program. ELAN is the acronym for EUDICO Linguistic Annotator; this tool was created for multimodal analysis, for complex labeling of audiovisual resources. It has a layer-based data model that supports multi-level annotations, and multi-player time-based media [4].

With Elan, it was possible to add annotations during the execution flow of each video. The notes contain descriptions of the movements made by each participant, such as for the gesture of moving forward, descriptions were made of how the participant performed the movements of his hands from the beginning to the end when he lowered his hands to perform the next movement. See Fig. 5.

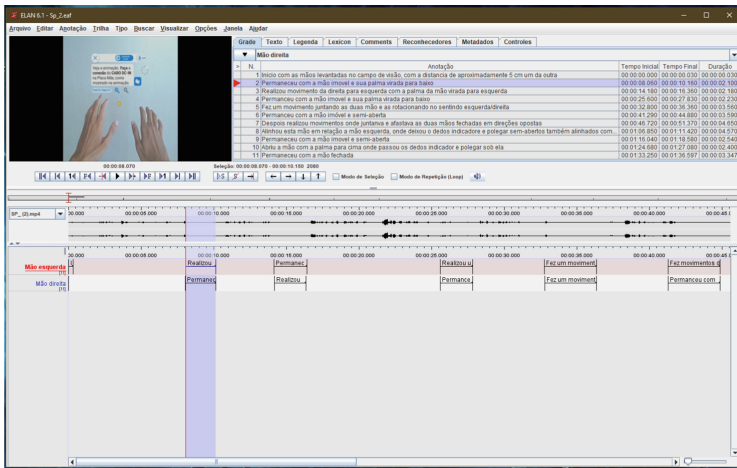


Fig. 5. Elan annotation interface

Figure 6 represents an example of how annotations are extracted from participants' videos. Participants were renamed Research Subject (SP), where each SP has its respective referents, such as the forward function and, for this function, the description of the movements of each of the hands. Each hand has its own distinctive character, where we color the movements categorized through a common factor derived from the videos capturing the movements of the hands of the SPs as well as the notes extracted from Elan.

The researchers organized by color each movement performed by the SPs. And after this separation, the elected movements were those that obtained the greatest common factor on the part of the participants. The descriptions were separated for each hand individually, showing from the beginning of the elicitation where the SP starts by raising the hand until the end when he lowers it at rest.

Aplicação do Cenário de testes para elicitación				
SUJEITOS DE PESQUISA	Sexo	Ação	Mãos	Descrição do movimento
Sujeito de Pesquisa 1	MASCULINO	INICIO	Esquerda	Iniciou com a mão semi fechada com aproximadamente a distância de 5cm entre elas, depois começou a afastá-las para realizar os movimentos.
			Direita	
		AVANÇAR	Esquerda	Realizou o movimento de apontar para frente movimentando a mão da direita para esquerda.
			Direita	
		VOLTAR	Esquerda	Realizou o mesmo movimento de apontar, porém mudando a direção, desta vez da esquerda para direita.
			Direita	
		SELECIONAR	Esquerda	Continuou com a mão no mesmo estado.
			Direita	
		ROTAÇÃO	Esquerda	Virou a palma em direção a mão esquerda e fechou a mão.
			Direita	
		ESCALA	Esquerda	Fez movimentos de zig-zag com a mão semi-aberta em paralelo com a mão esquerda para frente e para trás.
			Direita	
		LIBRAS	Esquerda	Fez movimentos de zig-zag com a mão semi-aberta em paralelo com a mão direita para frente e para trás.
			Direita	
		GRABAR	Esquerda	Fez movimentos de abertura da mão para frente, afastando-a levemente para direita e tomando distanciamento da mão esquerda.
			Direita	
		ENCERRAR	Esquerda	Fez movimentos de abertura da mão para frente, afastando-a levemente para esquerda e tomando distanciamento da mão direita.
			Direita	

Fig. 6. Annotations extracted from videos

Figure 7a represents the most common hand movements performed among all SPs during the execution of the gesture elicitation experiment representing the forward referent, and Fig. 7b demonstrates the common factor of annotations for the referent return.

For example, the gesture for the function of advancing, the researchers chose six types of movements in common by the participants. For the advance function, the gesture in which the SP makes the movement with the open hand from right to left was the most common execution among the others, demonstrating a natural movement because it emerged from a provocation made to the SP, and he performed his movement naturally.

The same occurs for Figs. 8a, 8b, 9a and 9b.

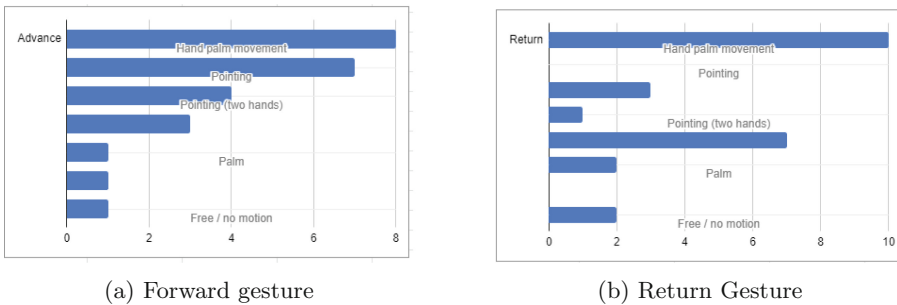
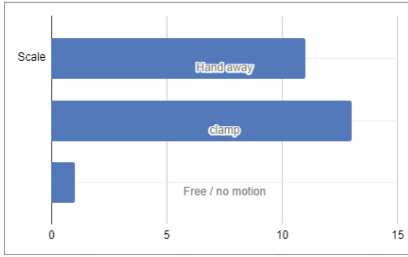
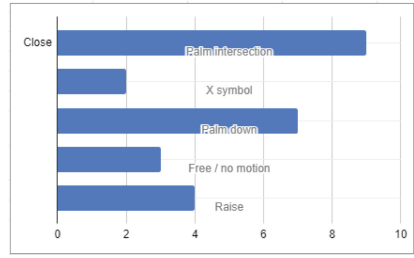


Fig. 7. .

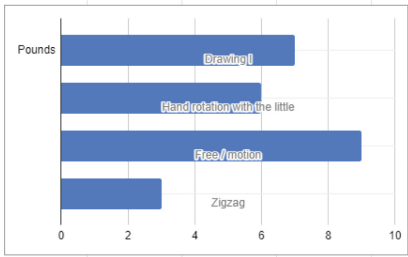


(a) Scale Gesture

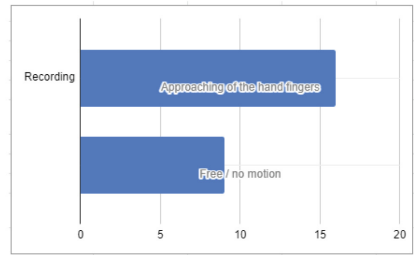


(b) Close Gesture

Fig. 8. .



(a) Pounds Gesture



(b) Recording Gesture

Fig. 9. .

8 The Gestures Elicited

After analyzing the individual movements, an average of 2 movements considered natural for each interaction were obtained, in which the researchers chose the two most performed the participant; each movement resulted from an analysis performed considering the naturalness in which the gesture was performed and also the ease of use. That the participant had in doing so. The gesture chosen for each command was the one whose category had the highest number of natural occurrences, the one that had the most performed category by the participants.

Next, we have the demonstration of the main gestures elicited from the application of this elicitation experiment. The images contain the name of the gesture and its movement being performed. The choice of the two most elected movements is due to the Libras referent having been elicited only two movements, so for the sake of standardization; the researchers decided to leave only the two most common movements for each gesture.

In Fig. 10a the first gesture elicited for the advance function is shown, where the participant performs the movement of moving the open hand to the right, in Fig. 10b the second gesture for the advance function is demonstrated, where the participant points the finger to the right. In Fig. 11a there is the first gesture elicited for the return function, in which the hand movement is performed towards the left, in Fig. 11b we find the gesture for the return function with the participant performing the left-pointing motion. In Fig. 12a, the participant performs the first gesture elicited for the scaling function, moving his hands away in opposite directions, while in Fig. 12b, the participant performs the second gesture elicited with pinch movements. In Fig. 13a, the participant performed the first gesture elicited for the recording function, joining the index and thumb fingers of the hands, and in Fig. 13b the second gesture elicited for the recording function. In Fig. 14a for the Libra function, the participant makes a movement with the symbol of L in the first elicited gesture, and in Fig. 14b, the participant makes an anticlockwise movement with the little finger upwards in the second elicited gesture. In Fig. 15a the participant performs a zig-zag movement with his hands for the first gesture elicited for the end function and Fig. 15b the participant lowered his hands in the second gesture elicited for the end function.



(a) forward gesture 1



(b) forward gesture 2

Fig. 10. .

(a) Return gesture 1



(b) Return gesture 2

Fig. 11. .



(a) Scale gesture 1



(b) Gesture scale 2

Fig. 12. .



(a) Record gesture 1



(b) Record gesture 2

Fig. 13. .



(a) Pounds gesture 1



(b) Pounds gesture 2

Fig. 14. .



(a) Close gesture 1



(b) Close gesture 2

Fig. 15. .

9 Discussion and Future Work

It is concluded that the elicited gestures occurred naturally; that is, each gesture was generated from a stimulus caused by the methodology of the experiment applied to the participants. The researchers' analysis also showed more than one gesture for each task proposed by the questionnaire while recording the videos; that is, each task can be performed by more than one elicited natural gesture, the open elicitation process provided more naturalness and ease during the execution of gestures by the participants, making the interaction process as comfortable as possible during the application of the experiment and the use of this type of interaction in augmented reality applications. As future works, we propose the expansion of gesture research with the PCD public (person with disabilities) so that we can elicit other types of gestures and also prove the accessibility of gestures currently discovered.

Acknowledgment. This study was an initiative of Dell Computers and Universidade Estadual do Ceará. In addition, developed by Lead - Research, Development and Innovation Center Dell.

References

1. Buchanan, S., Floyd, B., Holderness, W., LaViola, J.J.: Towards user-defined multi-touch gestures for 3D objects. In: Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces, pp. 231–240 (2013)
2. Cohé, A., Hachet, M.: Understanding user gestures for manipulating 3D objects from touchscreen inputs. In: GI 2012-Graphics Interface, pp. 157–164. ACM (2012)
3. Freeman, D., Benko, H., Morris, M.R., Wigdor, D.: ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, pp. 165–172 (2009)
4. Fuldain, J.S.: Elan analisi multimodalerako lanabesa. In: I. Jardunaldiak Ondareaz eta Hezkuntzaz (2012)
5. Khan, S., Tunçer, B.: Gesture and speech elicitation for 3D CAD modeling in conceptual design. *Autom. Constr.* **106**, 102847 (2019)
6. McNeill, D.: So you think gestures are nonverbal? *Psychol. Rev.* **92**(3), 350 (1985)
7. McNeill, D.: *Hand and Mind: What Gestures Reveal About Thought*. Bibliovault OAI Repository, vol. 27. The University of Chicago Press, Chicago (1994). <https://doi.org/10.2307/1576015>
8. Müller, C.: How recurrent gestures mean: conventionalized contexts-of-use and embodied motivation. *Gesture* **16**, 277–304 (2017). <https://doi.org/10.1075/gest.16.2.05mul>
9. Modanwal, G., Sarawadekar, K.: A gesture elicitation study with visually impaired users. In: Stephanidis, C. (ed.) *HCI 2018. CCIS*, vol. 851, pp. 54–61. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-92279-9_7
10. Morris, M.R., Wobbrock, J.O., Wilson, A.D.: Understanding users' preferences for surface gestures. In: Proceedings of Graphics Interface 2010, pp. 261–268 (2010)
11. Ortega, F.R., et al.: Gesture elicitation for 3D travel via multi-touch and mid-air systems for procedurally generated pseudo-universe. In: 2017 IEEE Symposium on 3D User Interfaces (3DUI), pp. 144–153. IEEE (2017)

12. Pham, T., Vermeulen, J., Tang, A., MacDonald Vermeulen, L.: Scale impacts elicited gestures for manipulating holograms: implications for AR gesture design. In: Proceedings of the 2018 Designing Interactive Systems Conference, pp. 227–240 (2018)
13. Piumsomboon, T., Clark, A., Billingham, M., Cockburn, A.: User-defined gestures for augmented reality. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013. LNCS, vol. 8118, pp. 282–299. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40480-1_18
14. Prieur, J., Barbu, S., Blois-Heulin, C., Lemasson, A.: The origins of gestures and language: history, current advances and proposed theories. *Biol. Rev.* **95**(3), 531–554 (2020)
15. Ruiz, J., Li, Y., Lank, E.: User-defined motion gestures for mobile interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 197–206 (2011)
16. Tarre, K., Williams, A.S., Borges, L., Rische, N.D., Barreto, A.B., Ortega, F.R.: Towards first person gamer modeling and the problem with game classification in user studies. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, pp. 1–2 (2018)
17. Vatavu, R.D.: The dissimilarity-consensus approach to agreement analysis in gesture elicitation studies. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)
18. Vatavu, R.D., Wobbrock, J.O.: Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 1325–1334 (2015)
19. Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R.D., Wobbrock, J.A.: A systematic review of gesture elicitation studies: what can we learn from 216 studies. In: Proceedings of ACM International Conference on Designing Interactive Systems (DIS 2020) (2020)
20. Wigdor, D., Wixon, D.: *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Elsevier, Amsterdam (2011)
21. Williams, A.S., Garcia, J., Ortega, F.: Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation. *IEEE Trans. Visual Comput. Graphics* **26**(12), 3479–3489 (2020)
22. Wobbrock, J.O., Aung, H.H., Rothrock, B., Myers, B.A.: Maximizing the guessability of symbolic input. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1869–1872 (2005)
23. Wobbrock, J.O., Morris, M.R., Wilson, A.D.: User-defined gestures for surface computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1083–1092 (2009)