Using Gesture-Based Interfaces
to Control Robots

Gabriel M. Bandeira[1]([✉]), Michaela Carmo[2], Bianca Ximenes[1],
and Judith Kelner[1]

[1] Centro de Informática (CIn) - Universidade Federal de Pernambuco (UFPE),
Recife, Pernambuco, Brazil
{gmb,bxhmm,jk}@cin.ufpe.br
[2] Universidade de Pernambuco (UPE), Recife, Pernambuco, Brazil
michaela.carmo@gprt.ufpe.br

Abstract. This paper analyzes human-robot interaction (HRI) to evaluate whether the use of a gesture-based interface is viable for robot control. An experiment was conducted with 19 volunteers. Using a body tracking device, they had to perform gestural commands to remotely control a mobile robot and complete a path marked on the floor. After the experiment, volunteers answered a questionnaire assessing aspects such as system's responsiveness, precision, and triggers to possible physical and psychological discomforts.

The results achieved validated the research aim partially, as it was determined that this control method is viable but only for short-term operations, pointing a necessity to create a more suitable control, less prone to cause user fatigue during long-term use. The developed system was designed not only for the analysis of HRI factors, but also for applications in remote operation contexts, such as industrial maintenance and exploration of inhospitable environments.

Keywords: Human-robot interaction · Human-computer interaction · Natural interfaces · Gestural interfaces · Remote operated robotics · Validation

1 Introduction

In the last few years, a wide range of researches explored usability in the fields of human-computer interaction (HCI) and human-robot interaction (HRI). Usability in this field comprehends several aspects, such as different kinds of command control, information feedback, communicability and its applicability on robotic systems [1–5].

To analyse the usability of a system, the way the interface decodes the commands must be considered. In the HRI field, a robot is usually controlled via classic interaction devices such as keyboards, mice and joysticks [6]. These devices may not be the most suited options for some applications because they work with complex software systems that commonly require prior training, which can be

unpleasant and time consuming. These systems can be simplified with the use of an interface that requires less learning time, thereby improving user experience. Such simplification is made possible with the use of natural interfaces [7]. One option is to operate robots by using gestures, for instance, programming a body tracking device to capture movements the user performs, which correspond to commands recognized by the robot.

Having understood the types of interface currently available, this article analyzes the viability of human-robot interaction (HRI) via gesture-based interfaces [7] using a body tracking device to control a mobile robot. To do this, we built a system with a robot that sends images captured by a camera attached to its structure and is guided remotely by an user, who, in addition to the video stream, also counts on a graphical interface that contains auxiliary information to control the robot.

The aim of this research is to investigate whether the use of a gesture-based natural interface is viable for robot control. A viable interface for robot controlling in this study was considered to be a technically feasible interface that enables the user to control the robot accurately. Beyond that, an ideal gesture-based interface should depend on commands that are both easy to learn and execute, as well as that generate little to no physical or psychological discomfort. Those aspects were assessed via questionnaire with 19 volunteers that were willing to participate in the study controlling the robot and testing the working interface.

The implemented solution was designed not only for the analysis of HRI factors, but also has applications in remote operation contexts. Environments of industrial maintenance and exploration of inhospitable environments are good examples of possible operation fields for this system.

On the next section, we explained the related works that supported this study, containing works related to the interface chosen and the device we used to capture the movements. The following section, divided in two subsections, encloses the methodology followed to make the systems we made to validate this study and the experiment made to validate the same. The last two sections consists of the results achieved and the conclusion we made from this research.

2   Related Work

The development of the robot system and the gesture-based interface was based on a number of works addressing interfaces in HRI that were presented in forms other than the classical mice-keyboard combination. They are not necessarily studies on natural user interfaces, being instead evaluations of desirable aspects of HRI interfaces, and what may render them a higher usability.

The first group of works studied concerned human-robot interaction interfaces that did not use only classic interaction devices. Concerning the interface, a few researches were studied to base our choice. Initially, paper [3] analyzes controlling a robot using a Graphical User Interface (GUI). The lessons learned are deepened in [1], which performed an analysis of interfaces for a remote robot control that reduces task load and fatigue. [2] evaluated a human-machine interface

where a combination of visual and haptic interactions was established to investigate the performance and viability of the proposed solution. Having studied these works, we could understand how each of the solutions could be explored and how a combination of GUI and another types of interface could be integrated. Finally, we found an example of natural user interfaces in the HRI field - an analysis of a semi-autonomous robot control using voice commands [4]. In this case, the robot accepts the commands and then analyzes its vicinity to recognise obstacles and possibly avoid them. This work analyzes the use of natural interface operated exclusively by voice command. As the use of a gesture-based natural interface had not been explored as far as we could find, we approach this aspect in our study. Users' attitude while they are interacting with robots is also a point of interest in the literature, as seen in [8,9]. For that reason, we also assessed this facet in our research.

Having defined how the interface would function, it was then necessary to determine how to track users' gestures. A common device used in HCI and HRI to recognise gestural commands is the camera. That happens because the camera can provide plenty of information about the scene, so developers can extract from the environment the exact information needed for their specific aims. That can be seen in an experiment where a hand-shape was recognised using an artificial neural network [10]. Based on this, we decided to use a camera based device, a RGB-D sensor. The RGB-D sensor can be understood in short as a device that captures images and the depth of each pixel of the image.

A research was performed to analyze how effective would be to read gesture commands tracking the user's body. Such a choice and its implications were presented in [10,11], being part of their analysis of user experience when interacting with a machine. Having understood the implications of the reading gestures choice, we could notice a few good aspects on adopting this approach and also some bad ones. In [11], for example, it was pointed some challenges like environmental noise including illumination changes and self occlusion - when one body part is occluded by another one - and how effective would it be to use multiple camera to reduce this last problem.

3   Methodology

The methodology is divided in two parts. The first contains an explanation of the technical aspects of the experiment, covering which modules were necessary to implement the system, why they were necessary, and the components used to build them. The next subsection contains a description of the experiment performed, including details of the interface and metrics used.

3.1   Technique

We adopted a gesture-based natural user interface [6] and analyzed the usability [6,12,13] of the user-robot interaction through a body tracking device. Such a choice was due to the fact that it is important that robot commands are intuitive

and simple, as that provides an easier way to interact with robots. An easier way to operate is consequently more appropriate for people of different ages and skill levels and, moreover, enables robot interaction and interface operation without the need of exhaustive training of the user.

After the theoretical analysis of control interfaces, a prototype that allowed to conduct a practical assessment of gestural interaction was developed. For that, the team implemented two subunits responsible for complemental parts of the system architecture. The first one was called Control sub-system and it ran a software that detected the commands of the user in form of gestures and sent it to the second sub-system, called the Mobile sub-system. The Mobile entity comprehended a robot that executed actions corresponding to the user's commands, received from the Control. The Mobile unit also sent some information about its surroundings to the user by the Control's GUI, as it will be explained later.

The Control sub-system consisted in a personal computer (PC) connected to a Microsoft Kinect platform. The Microsoft Kinect platform was chosen due to its ease of use with to a PC - plenty of APIs and tools are available and both environments are compatible. Besides, Kinect's precision on body tracking and affordability were also considered to settle the choice. The Control sub-system was responsible for analysing the pose of the user and comparing that pose to the ones stored in the memory of the software. During this pose match, if a match occurred, the Control sub-system would send a command to the Mobile sub-system corresponding to the matched pose. On the Control's PC there was also a window on the screen with a few pieces of information sent by the robot from the Mobile sub-system. After that process, the Mobile unit received the command and executed the action related to it. During this execution process, the Mobile also captured images from the environment where it is currently located and sent it to the user.

On the Control unit, there were four available poses that represented four different commands. These poses were defined by different arm positions, which followed the same movement logic of those performed to operate tractor's levers. To move forward, both arms should stay in front of the person, outstretched as in Fig. 1a. To move backwards, both arms should also stay in front of the user's body, but bent, bringing hands closer in as shown in Fig. 1b. Turning right or left was triggered by keeping both arms in front of the user, one of them close to the body and the other stretched out. On the turning commands, if the user's right arm was outstretched and the left was close, the turn would be to the left side as in Fig. 1c. Equivalently, the opposite arm order, the right one close and the left one outstretched as in Fig. 1d, would make the robot perform a right turn.

The Mobile sub-system was composed by a robot that acted according to the commands sent by the Control. To build a suitable robot, it was necessary to have a movement or action module, a perception module and a communication module, beyond the robot's structure. The necessity of the movement or action module was due to the kind of experiment we were performing. Our experiment required that because we wanted a system where the user could control a robot to explore places where (s)he could not go but the robot could. Another necessary
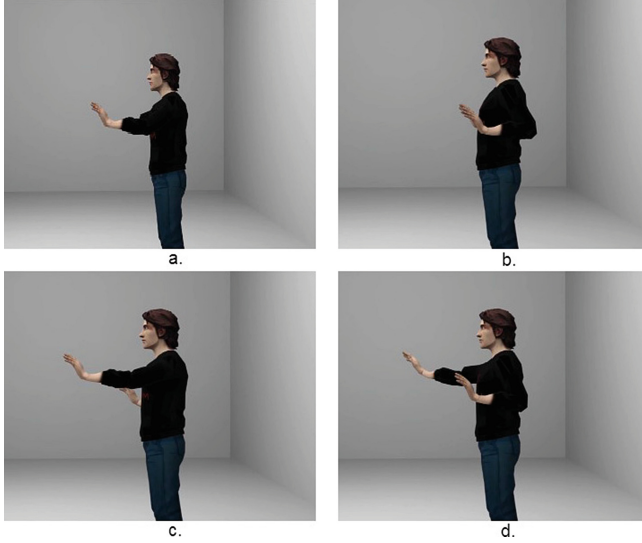
Fig. 1. Avatar performing the four available commands.

module, the perceptive one, was due to the fact that if the robot moves into places where humans cannot go, the human could not directly see the robot nor know how its surroundings are or what the robot could do there. To make the user aware of the robot's environment, (s)he must have some information about it in order to make informed decisions and command it. Finally, the communication module requirement was due to the constant exchange of information taking place between the Control and the Mobile subsystems.

The movement or action module consisted of a LEGO Mindstorms NXT 2.0 set, using its electric motors, wheels and structure components. To make the perception module, a camera was attached to the robot so the user could see exactly what was in front of the robot and make informed decisions. To build a communication module, we needed a fast enough system that allowed the robot to receive commands and also send images of the scene to the user as close as possible to real time. For these reasons, the team chose to use a dedicated wifi network, for a faster connection speed and less latency and packets loss, when compared to a communal use wifi network. The structure of the robot was also constructed with LEGO Mindstorms NXT 2.0 components, using its structure components and connectors. That choice was made to keep the modules of the robot easy to be assembled.

Considering the three modules that compose the robot on the Mobile subsystem, the platform chosen to connect these modules was the Raspberry Pi Model A. The Raspberry Pi was chosen due to its USB interface, programming possibilities for a C++ developer, processor speed and capacity, besides its size and weight. The USB interface was necessary to communicate with the LEGO Mindstorms NXT via USB cable and to connect the platform to the wifi network

via a wifi dongle. The programming possibilities previously pointed out are the availability of the platform to be programmed using the C++ language and the compatibility of the compiler and processor with external libraries. Those external libraries were necessary to do things such as communicating with other modules or capturing images from the camera, while the C++ language was chosen just as a project option selected by the team, since our developers were already used to writing code on that language. The processor speed and capacity were important considering that images were taken and then sent constantly to make the view of the user as close to real time as possible. Last but not least, the size and weight of the platform used in that module integration was crucial because it was a small mobile robot, so none of the platforms could be too big nor too heavy, or the robot's motors could not perform movement.

3.2   Experiment

To test HRI via gesture-based interface, the team conducted an experiment with 19 randomly chosen volunteers from diverse areas of knowledge and an age span of 18 to 35 years old. One third of the volunteers were female and two thirds were male. In the test, the volunteers were required to control the robot aiming to complete a path marked on the floor and then answer a questionnaire to evaluate their experience.

The volunteers were selected randomly in the proximity of the laboratory of the research team, by being offered the chance to take part in a study with a robot. The ones who participated gave their spoken consent. They were from a variety of work areas such as Biology, Secretarial Studies and Physical Education. Regarding their ages, they were between eighteen and thirty-five years old. After chosen, all of them were conducted to the laboratory to perform the test.

To control the robot, the volunteers were presented to the four available commands: moving forward or backwards and turning left or right. The Fig. 1 above consists of parts of a video tutorial shown to volunteers to familiarize them with the system's commands. The complete video lasted less than 20 s, and presented an avatar controlling the robot and indicating the commands being executed so the user could identify and repeat the movements. The video was shown right before the beginning of the test, on the screen of the computer running the Control sub-system, where the user would see the robot informations later, when the test started.

As soon as the tutorial was over, the user started the experiment trying to control the robot, moving it through the complete path (Fig. 2). The path marked on the floor with delimiting lines was shaped so that volunteers had to use at least three of the four available commands to complete the task. The length of the path was defined considering a 1-min inferior time limit based on the time development team's members themselves took to arrive at the end of the path with the robot. Figure 3 depicts one of the tests executed by the team to analyze the time necessary to get to the end of the path. The path length and complexity also aimed to spark feelings in the user, such as exhaustion, frustration, accomplishment, relaxation or pleasure, and serve as extra indicators

Fig. 2. Representation of the demarked path on the floor.

Fig. 3. One test executed by the team to analyze the necessary time to get the robot to the end of the path.

of the interaction validity. Objective metrics regarding system's precision and ease of operation included registering the number of times the robot was driven outside the marked path, as well as an overall average of the time volunteers needed to complete the path.

After having the robot complete the path, the volunteers answered a questionnaire evaluating their experience. On this evaluation, volunteers shared not only the amount of time needed to finish the experiment or how many times the robot got out of the delimiting path lines, but also their degree of difficulty to learn the movements and to perform them. They were also encouraged to talk about their impression of the robot's movement precision and the time for the actions to be executed by the robot after the commands were performed, which measure the system's responsiveness. We also tried to analyse the feelings the volunteers had during the experiment, be them related to controlling a robot or to the user control interface. Considering the feelings related to the interface,

we asked the user to explain the ergonomic and psychological discomforts that may have arose from the experiment so that we could analyse their experience, identifying points of improvement for the experiment.

4   Results

The results of this research are a synthesis of the experiment's validation analysis. They were interpreted as positive in general, since 84 % of the volunteers considered that natural control commands were indeed easy to both learn and perform, one of this study's main goals when analysing the usability of the system.

At the beginning of the validation, before controlling the robot through the designated track, all volunteers watched a tutorial and 78,9 % of them have considered it reasonably easy, easy or very easy to learn the movements. After that, 94,7 % volunteers managed to finish the 5,4 meters path, having done so in an average time of 3 min. Figure 4 shows the time spent by users to complete the course in reference to the group average and the time spent by developers.

Concerning the control of the robot vehicle, 84,2 % of the interviewed considered it easy to perform the command movements and the same percentage found that the robot operation with this type of control was accurate or very accurate, with no negative considerations from the users concerning lack of precision due to the technical limitations of the body tracking device.

On the feelings experienced during the trial, 85 % of the participants considered the activity not physically tiresome. It was also perceived that, in general, there was no psychological discomfort, with 79 % of participants stating that the
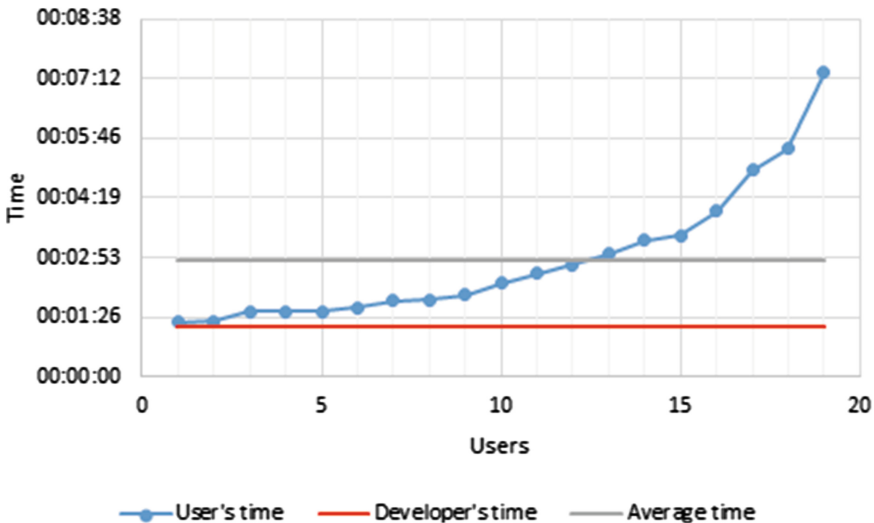

Fig. 4. Time spent by users to complete the course.

experience was pleasant, reporting feelings like amusement, relaxation and satisfaction. Some of them also described emotions related to the technical aspect, such as feeling innovative and immersed.

However, approximately 63 % of the volunteers claimed that even though they considered the activity pleasant or neutral and not tiresome during the experiment, they would not consider viable to operate the robot for long periods of time. That consideration was due to some degree of physical fatigue or ergonomic discomfort caused by the movements chosen to trigger the actions on the robot.

5   Conclusion

Our research assessed human-robot interaction through a gesture-based natural interface which worked with a body tracking device. It made possible for users to control a remote robot car in order to complete a path marked on the floor. The results allowed us to conclude that an improvement on users' experience and control intuitiveness occurs when using gestural interfaces. They also indicated that the loss of external command devices (keyboard, mice) had no effect on user's ability to learn or reproduce the movements that served as controls to the robot.

Despite the expected loss of precision in control due to the inherent inaccuracy of present-day body tracking devices, robot operation was not compromised by these limitations. The users found the sensor readings accurate and fast. The system also did not present any delays while processing, reading or interpreting the commands declared by the volunteers through gestures.

Additionally, the potential physical fatigue after long periods of operation highlighted by the users indicates that there is a fine line between the feasibility and infeasibility of gesture-based interfaces in the robot control field. This suggests there is a need to deepen the knowledge about the ergonomics of gestural commands in order to refine the movements used to control the robot.

The achievement of part of the research's aims, including intuitiveness on the use of gestural control and the precision on the control expressed by the volunteers, supports the use of gestural interface on the HRI field. With these results, the next steps of this study will focus on finding interaction methods in which users feel comfortable and that also minimize physical fatigue caused by the command movements, thus making the operation of machines using natural gesture-based interfaces for long periods viable.

References

1. Singh, A., Seo, S.H., Hashish, Y., Nakane, M., Young, J.E., Bunt, A.: An interface for remote robotic manipulator control that reduces task load and fatigue. In: Proceedings IEEE International Work. Robot Hum. Interact. Commun., pp.738–743 (2013)

2. Tanzini, M., Tripicchio, P., Ruffaldi, E., Galgani, G., Lutzemberger, G., Avizzano, C.A.: A novel human-machine interface for working machines operation. In: Proceedings IEEE International Workshop on Robot Human Communication, pp. 744–750 (2013)
3. Mayora-Ibarra, O., Sucar, E., Aviles, H., Miranda-Palma, C.: From HCI to HRI - Usability inspection in multimodal human - Robot interactions. In: Proceedings IEEE International Workshop on Robot and Human Interactive Communication, pp. 37–41 (2003)
4. Rao, A., Kumar, S., Renu, A., Nandi, G.C.: MILO - mobile intelligent Linux robot. In: Proceedings IEEE INDICON 2004. First India Annual Conference 2004 (2004)
5. Shackel, B., Richardson, S.J.: Human Factors for Informatics Usability. Cambridge UP, Cambridge (1991). Print
6. de Souza, C.S., Leite, J.C., Prates, R.O., Barbosa, S.D.J.: Interação Humano-Computador: Perspectivas Cognitivas e Semióticas. In: Fuks, H. (Org.) Anais das Jornadas de Atualização em Informática. Rio de Janeiro: Edições EntreLugar, 1999, pp. 420–470 (1999)
7. Wigdor, D., Wixon, D.: Brave NUI World: Designing Natural User Interfaces for Touch and Gesture. Elsevier Science, USA (2011)
8. Walker, R., Bartneck, C.: The pleasure of receiving a head massage from a robot. In: Proceedings IEEE International Workshop on Robot and Human Interactive Communication, pp. 807–813 (2013)
9. De Graaf, M.M.A., Ben Allouch, S.: The relation between peoples attitude and anxiety towards robots in human-robot interaction. In: Proceedings International Workshop on Robot and Human Interactive Communication, pp. 632–637 (2013)
10. Wei, J.W.J., Qin, H.Q.H., Guo, J.G.J., Chen, Y.C.Y.: The hand shape recognition of Human Computer Interaction with Artificial Neural Network. In: 2009 IEEE International Conference Human-Computer Interfaces, and Measurement System 04 (2009)
11. Cho, N.-G., Lee, S.-W.: Incorporating global and local observation models for human pose tracking. RO-MAN, 2013 IEEE, pp. 25–30 (2013)
12. Prates, R., Barbosa, S.: Avaliação de Interfaces de UsuárioConceitos e Métodos. An. do XXIII Congress, p. 14 (2003)
13. Prates, R.O., Barbosa, S.D.J.: Introdução Teoria e Prática da Interação Humano Computador fundamentada na Engenharia Semiótica. Jornadas de Atualização em Informática, JAI 2007, pp. 263–326 (2007)