

# Performance Evaluation of an xSight 3120 HMD and Usability Analysis of the Phantom Omni in a Thoracocentesis Simulation

Éllen S. Corrêa, Paulo R. Trenhago, Thiago E. A. Oliveira, Jauvane C. de Oliveira

Laboratório de Ambientes Colaborativos e Multimídia Aplicada – ACiMA

Laboratório Nacional de Computação Científica – LNCC

Petrópolis – RJ, Brasil

ellensc@lncc.br, trenhago@lncc.br, thiago.eustakio@gmail.com, jauvane@lncc.br

**Abstract**— This work describes the development of a virtual-reality-based medical training simulation. The system is composed of 3D models of a surgery room, a virtual patient (with the involved internal structures), as well as the software components that implements a thoracocentesis training procedure. This work goes on to evaluate the performance of a xSight 3120 HMD, as well as a usability test of a Phantom Omni to perform the medical procedure. Such usability analysis was drawn out of interviews with volunteers who used the system. The evaluated criteria included the time to complete the task, success rate and user complaints, such as HMD visual discomfort and arm fatigue while using the Phantom Omni.

**Keywords**— *Virtual reality; Head Mounted Display (HMD); Medical training; Pleural tap, thoracocentesis, Phantom Omni.*

## I. INTRODUCTION

Virtual Reality (VR) can be considered a high-level user interface. It is composed of a real-time computational simulation which allows the user to interact with the system using multi-sensorial channels. Such multi-sensorial channels are related to the human senses comprising sight, touch, hearing, smell and taste [1]. Virtual Reality has been applied to several knowledge areas, especially in training simulators.

Training real case scenarios is not always easy to do, either due to financial concerns, risks involved, or simply because there are not enough subjects in which one can practice. Based on such issues, Virtual Reality has been a natural choice for training setups. A flight simulator is a great example of successful training simulation. The success is partially explained by a realistic feeling, handling and behavior of the virtual plane in the virtual environmental conditions. Computational models, including physical models, ensure the realistic handling. As a result, pilots acquire flight experience without the risks to themselves or the equipment. That has been crucial for the training of new pilots. It goes on to allow one to acquire practice even in rare conditions, as the models can mimic the plane behavior in such situations as well.

Much like in the aviation industry, medical applications also involve quite a bit of risk, given that medical failure

caused by the lack of experience of someone who is yet learning a procedure, could cause serious issues to a patient, eventually even death. In a virtual setup the apprentice can try as many times as he/she needs to gain the required experience without such risks. It would be desirable, henceforth, that medical training could be performed in virtual reality systems much like what is already made in flight simulators. Such procedure would reduce the risks for real patients and the amount of medical errors in these circumstances, as they would always be accessed by medics with more extended experience, even though through simulations [2].

Although the usage of VR in training applications has been increasing lately, that is not so trendy in medical applications. That is due to a number of factors, which include economical, technical, scientific and cultural issues involved in the human health care [3]. Nowadays, however, a large amount of the computing and medical communities believe that it is possible to develop medical simulators that are as realistic and efficient as flight simulators.

With the advance of the VR research, some groups question the didactic efficiency of the technique, given that is less frequent a comparative analysis of a VR simulation against applications that use standard (non-immersive) interaction techniques that can be used in a training application of some medical protocol or procedure.

This work takes on the comparison of VR simulations against standard setups to evaluate the performance of an immersive Sensics xSight 3120 Head Mounted Display (HMD) [4] along with a haptic interface Geomagic Phantom Omni [5]. Such analysis was performed through the comparison of the immersive VR setup above against a standard non-immersive LCD-and-joystick based application which interacts with the same virtual environment. A VR simulation was developed as a training application for the Thoracocentesis (Pleural Tap) procedure. The application allows one to use multiple input/output options, which were then tried out by a number of volunteers. The system evaluated the time necessary to complete a given task, as well as success rates and any reported visual discomfort while using the HMD or Phantom. This work also verifies the usability of a phantom device for object

handling, comparing the performance of the users of the thoracocentesis procedure.

In the definition of this work we had a medical doctor with a Ph.D. in human anatomy following the development of the prototype of the Pleural Tap simulation. Such individual advised on how the procedure is performed, the parameters which were important to analyzed in order to determine a success or failure. That led to the development of an automatic evaluation agent, which keeps track of the user actions in real time, generating a performance report.

This work is structured in five more sections, beside this introduction. The next section introduces related work. Section 3 provides basic information related to the pleural tap procedure. Section 4 brings up information about the VR simulation developed for that procedure. Section 5 describes the experiments carried out, as well as the results achieved. Finally, Section 6 concludes this work, drawing up some conclusions and issues that need to be taken into account.

## II. RELATED WORK

There are many VR based medical training applications reported in the literature. The reports include those of immersive tactile-based applications (haptic VR simulations) as well as visually immersive simulations (such as those using a HMD). There is also a large set of training applications for medical procedures which do not rely on special interaction devices. In this section we'll report a number of applications in both extremes, with special focus on three works that exploit either immersive or non-immersive approaches for medical simulations.

The Virtual Environments for Medical Training and Education [6] describes a training system that aims at military medics and paramedics. The most interesting aspect of this system is the visual and haptic immersion. In this system the medic is part of a platoon in a virtual battlefield. The visual immersiveness of the system is provided by a HMD and the tactile immersiveness is provided by an exoskeleton, which also allows the system to keep track of the user body movements.

The system considers procedures that are usual of military battlefields, rather than those usual in hospitals. When a soldier is hit in the battlefield he/she falls and the image of a generic wound is shown in the anatomical area that was hit. When the user get closer to the wounded soldier he/she gets to see a more specific wound that replaces the generic wound that gets displayed when the user is farther away from the victim. One example: if a given soldier gets hit in his/her leg, the medic will see a specific wound in the femur and a tray with appropriated surgical instruments is displayed besides the wound. This system does not teach the medic what he/she needs to do, nor does it evaluate the user performance. It only allows a user, who already knows what needs to be done, to practice the procedure in the simulated battlefield.

Another work we document prioritizes learning and decision taking in a virtual hospital environment. The 3D Interactive Teams (3DiTeams) [7] was built around the Unreal Engine [8]. It was developed in the medical center at Duke

University, in partnership with the HumanSim Corporation [9]. This simulation is capable of mimicking several aspects of a hospital, ranging from medical behavior when receiving patients and the interaction amongst medical teams and other workers at the hospital. The system allows instructors to customize the simulation, through a collaborative module. Such module allows the instructor to modify the vital signs and behavior of a given patient. The system uses mathematical models implemented in a physiological engine. Such models characterizes the cardiovascular, respiratory and reaction to commonly used medications [10]. Finally, the system also uses high quality 3D geometrical models of the medical equipment and a coherent geometry of the human anatomy.

3DiTeams has been developed in a first person game format. The game is structured in three phases: independent learning, coordination and collaboration and reviewing of actions. The Independent Learning phase consists of study of communication concepts and team work, which is done through videos. In the Coordination and Collaboration phase the player joins a multiplayer session which may have up to 32 players. Each player has a specific role, amongst medic doctor, nurse, technician or observer. In this phase the team has to evaluate the clinic case of a patient and to start up the proper treatment, according to decisions taken by each player in real-time. Both video and audio in the virtual world is recorded, including the action of all players. Such video is used in the third phase of the game: action reviewing. Such reviewing allows that each player gets evaluated by specialists, allowing also self-evaluations.

The third work we will briefly document is a system designed to teach the breast biopsy procedure [11]. Such system is also structured as a game with five levels of difficulty. The medical instruments can be handled both using a mouse or a Phantom Omni. Each level of difficulty is composed of two stages. In the first stage the user interacts with the virtual environment to practice the technique, whilst in the second stage such user has to answer an objective, theoretical, question, which unlocks the next difficulty level, if the answer is correct. This system has, amongst other features, collision detection, deformation of virtual objects. The setup was experimented by a number of users to evaluate both visual and interaction aspects, later going through another user evaluation to determine the usefulness of the virtual environment for medical training.

Both the first and second works listed above are good examples of VR applications for medical training. The first application uses immersive visualization through a HMD, and body tracking to create a realistic experience for the user while in the virtual battlefield, that being considered key point in the learning process of the users. The second work, on the other hand, considers that a LCD screen, mouse and keyboard, or the use of a tablet, may suffice to the learning process. The third work is an example of a work that relies on a haptic device (Phantom Omni) to handle surgical instruments, which is one of the proposals of this work.

In this work, we developed and tested, with a group of volunteers, which interaction devices provide the best experience for the user in a medical simulation. We also take

into account if the choice of technology has any interference with the learning process or in the practice of the tasks in the virtual simulation of a pleural tap procedure.

### III. THE PLEURAL TAP PROCEDURE (THORACOCENTESIS)

The pleural cavity is an anatomic space between the parietal pleura (a layer of biologic tissue that covers the lungs) and the visceral pleura (a tissue that covers the internal portion of the chest cavity). Under normal circumstances, the pleural cavity is filled with a liquid that works as a lubricant, reducing the friction between the pleurae, which is a result of the breathing movements. Pleural effusion is the accumulation of an abnormal amount of liquid between the pleurae, which may be caused by several factors [12]. In Figure 1 we can see the pleurae, with a pleural effusion in the right hand side of the patient.

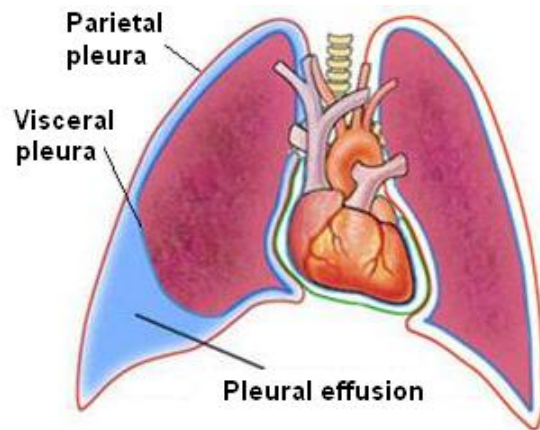


Fig. 1. Illustration showing the pleurae and the pleural effusion [13].

The pleural tap, also known as thoracocentesis, is a medical procedure that accesses the pleural cavity through the thoracic walls, allowing one to collect a sample of the pleural liquid for biopsy purposes. It also allows one to drain any excess of liquid which is in that cavity. The biopsy of the pleurae and the removed liquid aims at allowing further diagnostic of pleural diseases. That is considered a normal routine exam.

In the thoracocentesis procedure, the patient should be properly positioned, most often seated. The skin where the procedure will be performed shall be cleaned up with an antiseptic solution. After that a local anesthesia is applied. Once such initial steps are accomplished, the procedure itself takes place, when a needle is inserted from the skin down to the pleural cavity. The liquid is then pulled up or a catheter is installed [12].

### IV. SYSTEM DEVELOPMENT

A diagram of the proposed system architecture is shown in Figure 2. In the central part of the Figure 2 one can find the software modules that were implemented with the Vizard IDE [14], which is VR specific. The user interaction is carried out by the input/output devices supported: keyboard, joystick and Phantom (left in the Figure 2). The virtual world visualization is done through a HMD or LCD monitor (right hand side of Figure 2). Details of each module, demos, training and

simulation are introduced in the subsections A, B and C that follow.

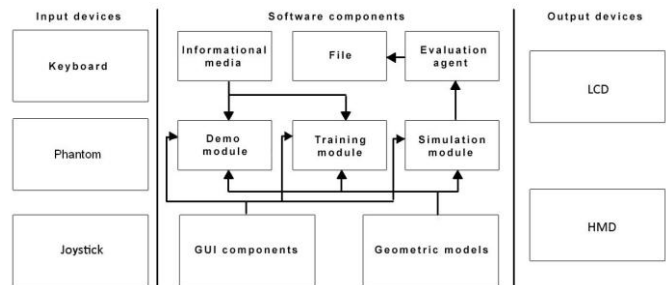


Fig. 2. Hardware and software diagram.

The system allows one to use the following options of input devices: a joystick, a Phantom Omni and/or a keyboard. To navigate in the system one can choose to use either the joystick or the keyboard. The interaction with virtual objects can be performed through the joystick or Phantom Omni. During the execution of the demo and free training modules the user may access informative media through specific keyboard or joystick commands. The system provides two options of visualization output: a LCD monitor, with no stereoscopy, or a HMD, which other than providing a stereoscopic visual output, restricts the user vision to the virtual world alone. Further details of such devices are shown in Subsection D below.



Fig. 3. 3D model of the surgical room.

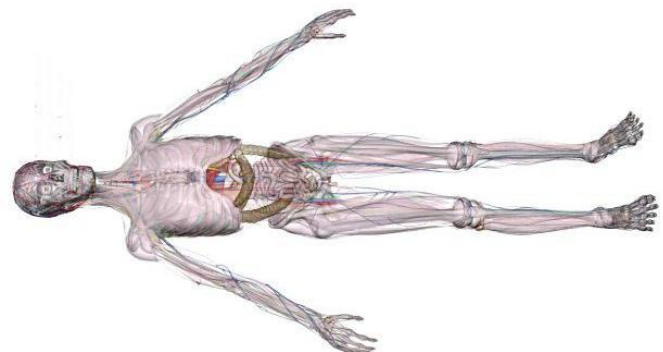


Fig. 4. 3D model of the human body (without the skin).

In order to provide more realism to the simulation, a 3D model of a surgery room has been used in the simulation [15], Figure 3. The room was customized with more appropriated

lightning. A number of surgical instruments used in the thoracocentesis procedure were also added to the scene.

The 3D model of the virtual patient used in the simulation is the Anatomium P1 Set [16], Figure 4, customized to the application at stake, namely optimizing the models of the circulatory, respiratory, skeleton and skin in the regions of interest.

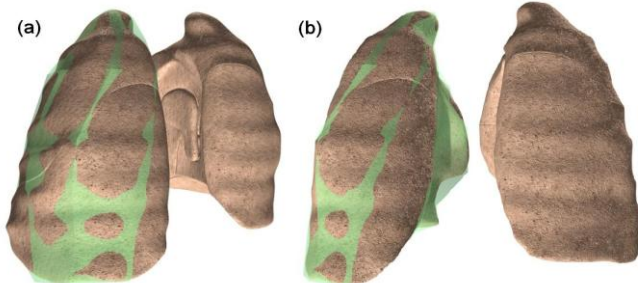


Fig. 5. 3D model of the lungs, along with green approximated geometry used in the collision detection system. Side view in (a) and frontal view in (b).

We also developed 3D models of approximations of the organs involved in the surgical procedure. Such simpler models are crucial to speed up the collision detection between surgical instruments and the anatomic models. These 3D models are transparent and co-located with the 3d organ models (juxtaposed). The two objects become integrated, as shown in Figure 5 with regard to the left lung. We applied a green color to the approximated geometry in order for it to become visible. Such procedure was used when it was not doable to perform collision detection with a given complex structure in real time. The simplified structures were developed using the software 3ds Max [17].

#### A. Demonstration Module

The demonstration module carries out the whole pleural tap procedure without interaction by the user. The user is just a spectator, being allowed free navigation in the environment, which allows him/her to watch the procedure from any chosen angle. In such mode a virtual medic performs the medical procedure, in all its steps, through pre-defined movements. The user is allowed to interrupt the procedure, go back or forward the execution time, as well as to access auxiliary media. He or she cannot interfere with the medic actions though.

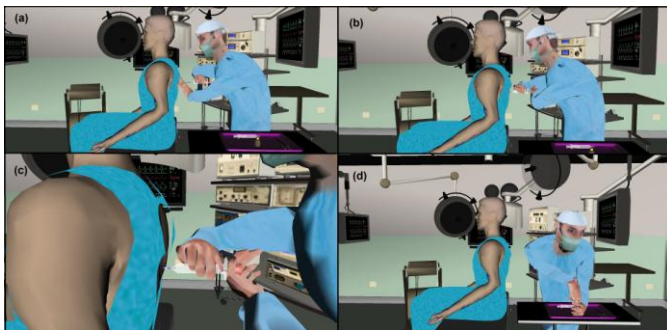


Fig. 6. Virtual medic carrying out the pleural tap procedure: (a) needle insertion; (b) aspirating the pleural liquid; (c) detail of the aspiration of liquid; (d) draining the syringe content in a container.

Figure 6 illustrates the sequence performed in the demonstration of the pleural tap procedure. The sequence showcases the final stage of the procedure, the removal of the liquid from the pleural cavity. Figure 6-a shows the insertion of the needle, Figure 6-b displays the aspiration of liquid, which is further detailed in Figure 6-c in close up. Finally, Figure 6-d gets the liquid pushed from the syringe and deposited in a proper container.

Besides the ability to interrupt, rewind, restart or fast forward the action, another option available for the user in the demo stage is the ability to change the transparency of the clothes, skin and organs of the virtual patient, as seen in Figure 7. This option exposes the anatomic structures involved in the procedure, allowing the user to see what happens in further detail while the procedure takes place. The transparency adjustment can be performed either through the joystick or the keyboard.

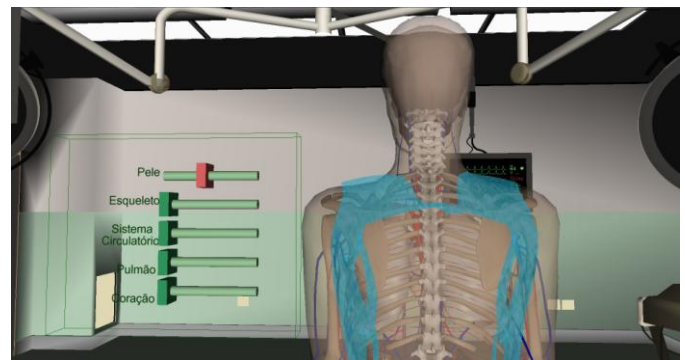


Fig. 7. Transparency selection in the model of the virtual patient..

#### B. Training Module

The training module enables the user to perform, himself/herself, all actions that were shown in the previous module. The user interacts with the system through the input devices available. This is the module that allows the user to practice, with support of the system, all steps related to the thoracocentesis procedure. Amongst the actions the user can perform in this module are: to move freely throughout the virtual scenario, to handle several tools, medical equipment and utensils. There is no pre-defined objective to be performed, neither there is a chronometer counting the time spent by the user to perform any specific action, be it related or not to the pleural tap procedure.

In order to allow the user to practice the pleural tap procedure, a number of methods were developed to allow the user to handle the three surgical instruments involved in the procedure, namely: asepsis gauze, anesthesia syringe and pleural liquid removal syringe. The handling of the instruments can be performed through a joystick or Phantom. In order to change the device being handled the user needs to press specific buttons in either device. The movements that are allowed include rotation in all axes, as well as translation. The joystick interface also allows one to select the speed in which the rotation and translation are performed. There is also a textual tutorial, accessible through keyboard keystrokes, with information regarding the pleural tap procedure.

The proper behavior of the system requires an efficient collision detection mechanism, along with proper ways to show the user that a given target structure was reached. In that regard, we developed an indication mechanism that changes the color of part of the syringe, the piston, to indicate when the needle reaches the proper structure. When the user is performing the pleural liquid collection step, for instance, the piston changes color from blue to purple, indicating that the needle has reached the pleural cavity. Figure 8-a and 8-b show, respectively, such color change. If the needle hits structures that are not related to the procedure, like bones or lungs, the error is signaled by the piston changing its color to red, as shown in Figure 8-c.

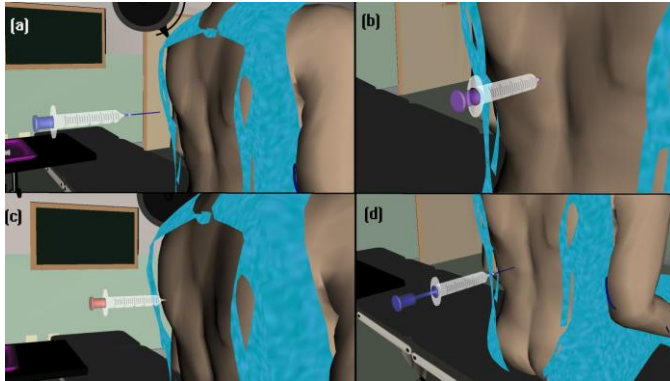


Fig. 8. Syringe piston changing color in the pleural liquid collection stage: (a) blue at procedure start, (b) purple when the right position is reached, (c) red when a wrong structure is hit and (d) blue again when the procedure is complete (needle removed from the patient body).

### C. Simulation Module

The simulation module is the most complex portion of the system. It provides a simulation setup where the user can be, to some extent, evaluated by the system itself. Basically, this module extends the training module with the addition of an evaluation agent which is capable of monitoring the user actions and determining if the procedure was performed correctly. The agent analyzes the order of the actions performed, the time spent in each step, as well as if the medical instruments were properly handled. In this module all resources that provided support to the user are turned off. That includes the transparency control and the auxiliary media, amongst others. The user had to, henceforth, perform the actions on his/her own, autonomously.

*a) Evaluation Agent:* The criteria verified by the evaluation agent in all three stages of the procedure (asepsis, anesthesia and pleural tap) are: the order of execution of the stages, the total time spent in each stage, as well as the correctness of the user actions in each stage.

The analysis of this criteria is performed through the monitoring the start and end collision events between objects. The system registers collision events between anatomical structures and surgical instruments. Such information is stored in a data structure that is later sent through a post-processing by the end of the simulation. The results from such analysis

are then stored in a text file. One example of the data stored in such text file is shown in Table I.

TABLE I. MODEL OF REPORT FILE WITH USER DATA.

User Report			
STAGE	SUCCESS	TIME (sec)	ORDER
Asepsis	True	30.00	1
Anesthesia	True	45.00	2
Thoracocentesis	False	41.00	3

To further illustrate this mechanism, let us consider an example. In order to determine if a user has performed the procedure correctly, the system needs to interpret some events in the right sequence. First and foremost, collisions need to be detected, in the right order, between the needle and the correct anatomic structures: skin, muscle and pleural cavity, without collisions being detected with bone or lungs. Secondly, the syringe piston needs to be pulled while the needle is within the pleural cavity. Thirdly, the syringe, and its needle, needs to be removed, terminating the collision that was detected at the first step in sequence: pleural cavity, muscles, skin. Moreover, the syringe shall be inserted in the unloaded state and removed in the loaded state. If all such events occur in the proper order, the system records “true” in the SUCCESS field of Table I, the total time to perform each step of process is registered in the “Time” field of the table and the field ORDER holds the order in which the stages have been performed.

### D. Graphical Output Devices

One of the goals of this work is to evaluate the performance of a Sensics xSight 3120 HMD while generating stereoscopic imagery to the user. Also, we wish to evaluate whether or not such imagery improves the performance of the user while carrying out his/her tasks in the pleural tap procedure. For that reason, two options of graphical output are available in the system. We have a simple 2D monoscopic view of the 3D virtual world, generated in a standard LCD monitor, as well as a stereoscopic output generated via two separated 2D images targeted to each of our eyes, which are fed to the two input channels of the HMD. The last setup provides a 3D immersive view of the virtual world to the user. The graphical connectivity of the system is shown in Figure 9.

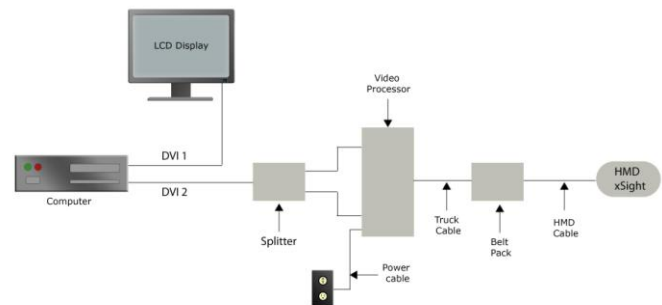


Fig. 9. Graphical output of the system for a Sensics xSight 3120 HMD. Adapted from [18].

The HMD used in this work has a graphical resolution of 1920 x 600 pixels per eye, with 24 bits per pixel (color). The horizontal field of view is adjustable between 72° and 120° and the vertical field of view between 30° and 45°, with overlap of 53°. The distance of projection from the eyes is equivalent to 1m, using lenses between the eyes and the HMD displays.

The LCD monitor used was a LG Flatron w2353v, whose resolution is 1920x1080 pixels, with 32 bits per pixel (color) and a diagonal dimension of 22", without stereoscopy.

### E. Input Devices

The input devices used in this work include Joystick, keyboard and a Geomagic Phantom Omni. Both the Joystick and keyboard may be used to navigate and interact with the virtual environment. The Geomagic Phantom Omni is a tactile device that works both as an input and output device. One of the features of the Phantom is to keep track of both position and orientation (6 Degrees of Freedom – DOF) in the space. It is handled through a pen-like structure (called stylus) that is attached to the tip of the haptic device. In the virtual world, the stylus may represent other objects, like a syringe in our application. The position and orientation are determined from the status of encoders linked to three electric step motors. The movement is restricted by the size of the Phantom Omni, which was designed for hand movements. Another feature of the Phantom is the ability to activate the step motors to mimic interaction forces that the user shall feel according to the actions performed in the virtual world. That allows the user to feel in his/her own hand what he/she would feel in similar situation in real life. The force feedback mechanism heavily enriches a simulation. In this work we only use the Phantom as an input device though.

## V. RESULTS

In this section, we introduce the results of the tests performed for the thoracocentesis procedure. The tests aimed at verifying the effectiveness of the system in training and evaluating a group of volunteers while learning a medical procedure. We also estimated the improvement in the performance of volunteers when a HMD or Phantom was used as visualization and interaction devices, respectively. The tests were performed with twelve volunteers who had no previous contact with medical or nursing training. That was observed aiming at minimizing the influence of pre-existing knowledge in the learning experience. Four different tests were carried out, namely: (A) Initial test, (B) Second test, (C) Third test and (D) fourth test.

### A. Initial Guided Test

The goal of this first test is to evaluate the immediate exposition of users to the virtual training and the HMD. The criteria verified has to do with any discomfort the users may report while using the HMD. In order to perform this test, we modified the software module responsible for running the simulation. During the test all helping functionality was enabled to guide the user in the stages of the medical procedure. Such behavior allowed the user to start using the system quickly, without the need for a long training on the system usage itself, which would expose the user to the HMD

before the simulation of the medical procedure would take place.

In order to perform this test the set of volunteers was split in two groups of six users each. The first group started the system using the HMD and later used the LCD while the second group started using the LCD and only after that got to use the HMD. Table II shows how the experiment was conducted with the two groups. The graphical output was alternated aiming at minimizing possible interferences in the results due to a longer or shorter exposure to the training setup.

TABLE II. EXPERIMENTAL SETUP FOR THE INITIAL TEST.

Group	Demonstration using the LCD monitor (1x)	Training with the HMD (2x)	Simulation with the HMD (1x)	Training with the LCD (2x)	Simulation with the LCD (1x)
<b>Group 1</b>	Demonstration using the LCD monitor (1x)	Training with the HMD (2x)	Simulation with the HMD (1x)	Training with the LCD (2x)	Simulation with the LCD (1x)
<b>Group 2</b>	Demonstration using the LCD monitor (1x)	Training with the LCD (2x)	Simulation with the LCD (1x)	Training with the HMD (2x)	Simulation with the HMD (1x)

The test is composed of two phases for the two groups: training of the required abilities in the pleural tap procedure and; simulation to evaluate the knowledge acquired in the previous phase. The training phase consisted of two identical training sequences, including asepsis, anesthesia and pleural tap. The simulation phase was composed of a single try. Immediately after finishing this training cycle the user repeated the same cycle again, but using the other graphical output option.

The data collected during the simulation phase were the total time spent by each user in the asepsis, anesthesia and pleural tap. We also collected user reports with regard to any visual discomfort with the HMD, as well as that related with the quality of the rendering with regard to: color identification, which is related to any difficulty to notice the change in colors in the pistons of the syringe; object border identification, which is related to the presence of patterns due to graphical aliasing or any other difficulty to visualize the needle. The user reports were collected through a standard questionnaire answered by each user after the test sequence. The questionnaire is composed of questions with scales varying from 0 to 10, where 0 means that it is not possible to differentiate the colors while 10 means that the colors change are easily noticed.

The graph shown in Figure 10 shows the average time spent by the users to perform each task. In such graph we can notice that the average time required is slightly different depending on the graphical output device in use. However, in order to properly verify that there is indeed a difference between the averages, we need to perform a statistical test. Due to the non-Gaussian characteristic and the small size of the sample, we chose the Wilcoxon test [19] for coupled samples in order to determine relevant sample differences. In the graph shown in Figure 11 we have a subjective evaluation of the users with regard to the color quality and object border visibility in both output devices. The percentage of users who reported visual discomfort while using the HMD is shown in the graph of Figure 12.

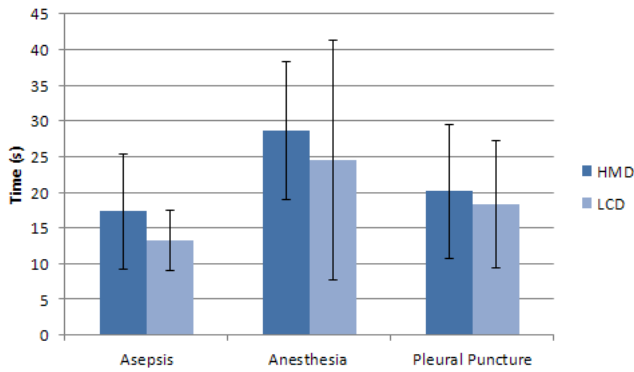


Fig. 10. Averages and standard deviation in timings required to complete each part of the pleural tap procedure.

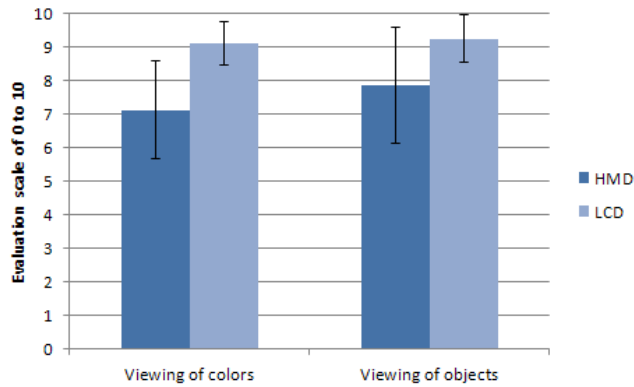


Fig. 11. Subjective levels of quality in color definition and object border visualization in the rendered virtual environment.

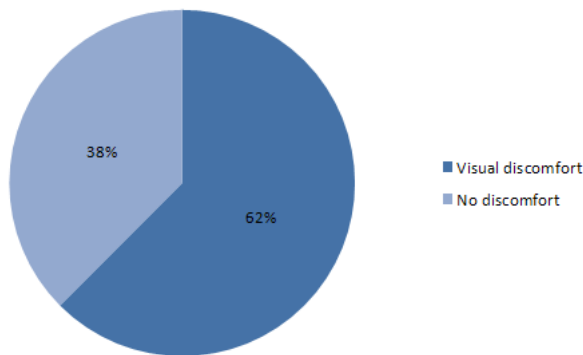


Fig. 12. Fraction of users who reported visual discomfort while using the HMD in the initial test.

The statistical test revealed that for the data shown in Figure 10, the samples do not show relevant difference, with 95% of reliability. That means that, although 62% of the users reported some visual discomfort, they all finished the procedure with satisfactory results in all stages of the pleural tap procedure. We shall remember that in this test the computer helps the users. The system informs when the virtual objects that represent the medical instruments reach the appropriate

anatomical structures, which help the users in the evaluation of the depth they needed to push the needle until it reaches the pleural cavity. Technical details related to the causes of visual discomfort from the HMD usage can be found in [20] and [21]. With regard to the graph shown in Figure 11, the statistical test showed relevant differences only in the criteria to differentiate the colors, which suggests some deficiency in color accuracy when the HMD is used.

*B. Second Test*

In this second test we evaluate the performance of the users while carrying out the pleural tap procedure without the system help. We also verify any difference in success rates when we change the graphical output device. For this test we used the same twelve volunteers, organized according to the experimental setup reported in Table III. Values 1 through 6 represent the order in which the activities were performed by the users of each group.

TABLE III. EXPERIMENTAL SETUP FOR THE INITIAL TEST.

<b>Group 1</b>	1. Demonstration in LCD (1x)	2. Training in LCD (20 - 40 min)	3. Simulation in LCD (1x)
	4. Demonstration in HMD (1x)	5. Training in HMD (20 - 40 min)	6. Simulation in HMD (1x)
<b>Group 2</b>	1. Demonstration in HMD (1x)	2. Training in HMD (20 - 40 min)	3. Simulation in HMD (1x)
	4. Demonstration in LCD (1x)	5. Training in LCD (20 - 40 min)	6. Simulation in LCD (1x)

The graph shown in Figure 13 shows information regarding the fraction of users that reported some visual discomfort while using the HMD in this second exposition to the device. This result shows a tendency to reduce the visual discomfort once the users spend more time wearing the device, i.e. the users have a tendency to adapt themselves to the HMD as time passes by. In this test the users had more time to practice the protocol before they had to run the simulation, in addition to the expertise already acquired in the first test.

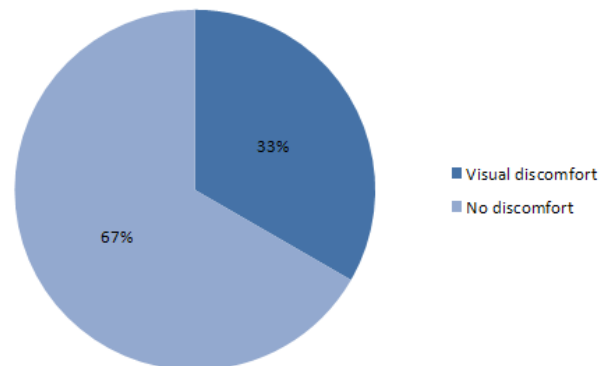


Fig. 13. Fraction of users who reported visual discomfort while using the HMD in the second test.

Another difference in this test is that the system performed automatic monitoring of the behavior of the users during the

simulation. The data related to the success rate for both devices are displayed in Figures 14 and 15. The success rate for the HMD reached 75%, while the LCD users reached only 33%. That means that the success rate is more than twice as much for users who interact with the system using the HMD when compared with those who use a LCD monitor. A detail worth mentioning is that the two tasks that require more ability from the users, anesthesia and pleural tap, add up 100% of the errors performed by the users. There were no errors in the asepsis stage. This fact confirms the hypothesis that the stereoscopic vision is an important feature when the user carries out complex tasks.

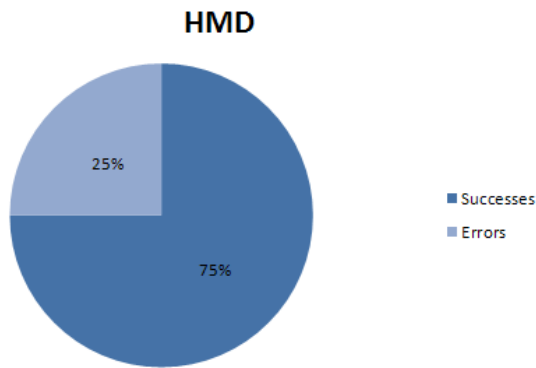


Fig. 14. Success rate in the second test for users wearing a HMD.

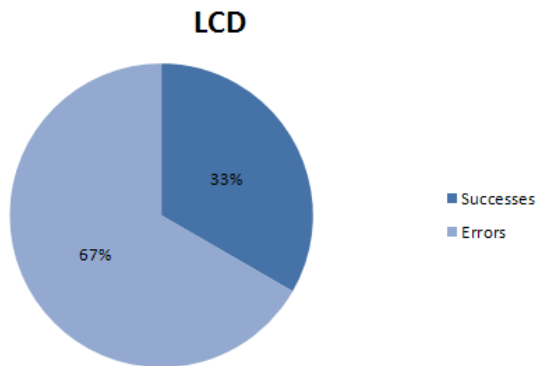


Fig. 15. Success rate in the second test for users using a LCD monitor.

Figure 16 shows the average time spent by the users while carrying out the several steps of the medical procedure. A Wilcoxon test has been applied on the acquired data aiming at identifying relevant differences between the timings measured for the HMD and LCD devices. The result indicates, with 95% reliability, that only the timing measured with regard to the pleural liquid collection are distinct ( $p\text{-value} = 0.003418$ ), while the remaining timings do not represent relevant differences. We shall also take into account the error rates in the anesthesia procedure, which is distinct for the HMD and LCD. 71% of the errors performed in the anesthesia procedure occurred with the LCD, which added up to more than twice as many as those for users wearing the HMD, which reached only

29%. Finally, these results strengthen the hypothesis that the performance of the users improves when he/she uses a HMD with a stereoscopic graphical output.

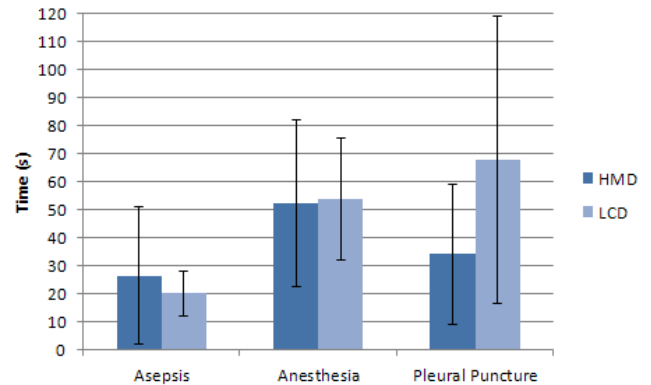


Fig. 16. Average and standard deviation for timings related to the execution of the three main tasks of the pleural tap procedure.

C. Third Test

The experimental setup of the third test is identical to that of the second test. The criteria evaluated are also the same. The goal of this new test is to evaluate the possible improvement in user performance after further training. If any improvement is reported, we will verify if the success rates have any improvement. Finally, we observe if the level of discomfort with the HMD has any change after further user exposition to the device.

The data obtained shows that the error rate is stable when the user wears the HMD, i.e. there is no significant difference in the user performance, as reported in Figure 17. However, the error rate diminished considerably when the users interact through a simple LCD monitor, which is reported in the graph displayed in Figure 18. That suggests that with longer training, the difficulties caused by the lack the sensation of depth (provided by the HMD stereoscopy) is overcome at least partially, once the users gets more hours of training. The discomfort rate reported by users wearing the HMD remains stable at 33%.

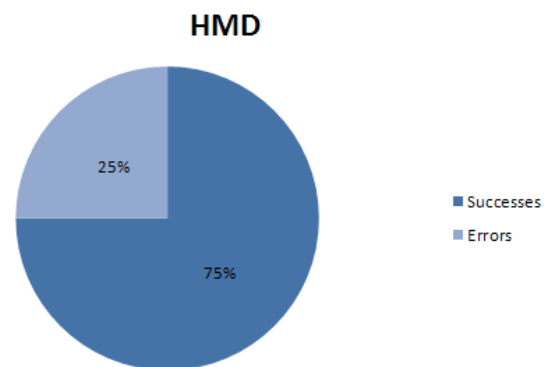


Fig. 17. Success rate in the third test when the user wears the HMD.



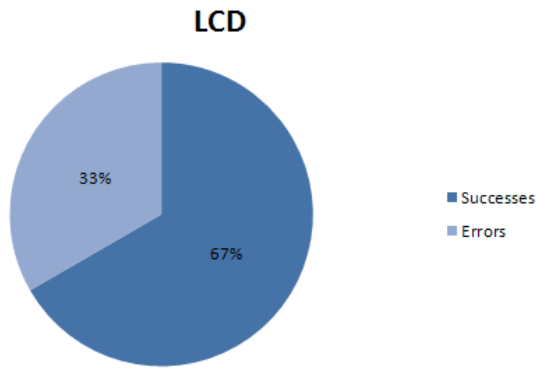


Fig. 18. Success rate in the third test when the user uses a LCD monitor.

Figure 19 shows the average time and standard deviation for times measured during the user interaction in this third test. The Wilcoxon test shows that the only stage that displays a relevant difference is the asepsis (p-value = 0.0210), while the anesthesia (p-value = 1) and thoracocentesis (p-value = 0.9097) stages do not display relevant differences in the third test, adopting 95% reliability. That affirms the hypothesis that longer training makes up for the lack of depth sensation of the LCD when compared with a HMD. The issue can be disregarded if you increase the training time in the simpler setup for all three stages related to the thoracocentesis procedure. We can then conclude that the performance of the users in all steps of the procedure, in this last test, have been satisfactory for both output devices. We had an important reduction in error rates for users of the LCD, who reached similar rates of users of the HMD in those steps that require greater precision.

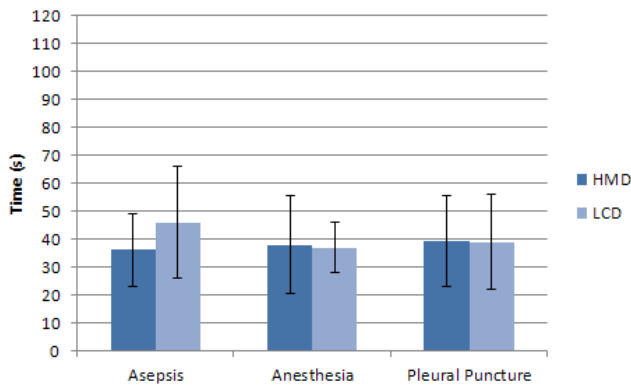


Fig. 19. Averages and standard deviation for timings measured while users perform the actions related to the thoracocentesis procedure.

In this third test there is uniformity both in the error rates as well as in the averages and standard deviations of average execution times measured in the anesthesia and puncture stages. With regard to the asepsis, there has been an increase in the time spent in the procedure when compared with the previous tests. That may indicate greater worry by the users in this important step of the procedure. Although the timings here have increased, one can observe that the standard deviation is similar amongst the two devices (about 13 second for the HMD and about 19 seconds for the LCD) when compared to the

previous test that had distinct standard deviations between the devices, with about 24 seconds for the HMD and 8 seconds for the LCD.

If we compare the three tests performed with the volunteers, we can observe great reduction in visual discomfort using the HMD if we compare the first with the second test, as reported in Figure 20. If we compare the second and third tests, however, the discomfort rate stabilized, which suggests that the time needed for a user to adapt him/her with that device was that of the first + second tests.

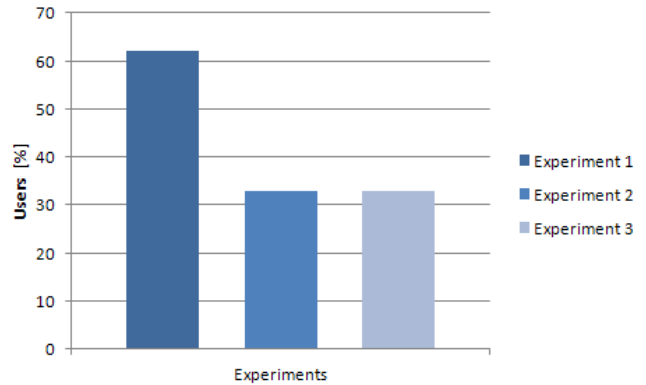


Fig. 20. Percentage of volunteers that reported visual discomfort wearing the HMD in the three tests performed.

With regard to user error rates, we can observe that with the HMD they are the same in the second and third tests, as shows the graph in Figure 21. However, the LCD users improved their error rate, diminishing the levels from the second to the third experiment. The first test was aided by the system, reason why no error report was generated.

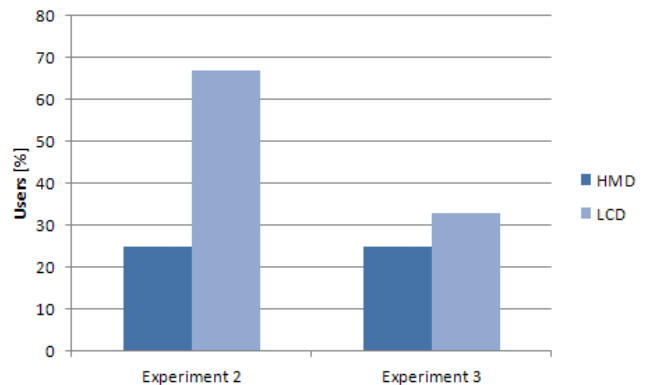


Fig. 21. Error rates in the medical procedure using both visual output devices in the second and third tests.

*D. Test Comparing Input Devices*

This test aims at evaluating the usability of a Phantom Omni device as an option for the joystick. The test was performed with the same group of volunteers. The evaluation was performed through the comparison of success rates achieved when a Phantom Omni or a joystick were used as input device to handle the instruments in the medical applications. The test setup was that reported in Table IV.

TABLE IV. EXPERIMENTAL TEST SETUP TO COMPARE INPUT DEVICES

Group	Demo of the protocol	Training with Phantom	Simulation with Phantom	Training with Joystick	Simulation with Joystick
Group 1	Demo of the protocol	Training with Phantom	Simulation with Phantom	Training with Joystick	Simulation with Joystick
Group 2	Demo of the protocol	Training with Joystick	Simulation with Joystick	Training with Phantom	Simulation with Phantom

The volunteers carried out the medical procedure using both input devices. The test was then timed and the results achieved are represented in the graph shown in Figure 22. Such graph shows that the use of the Phantom Omni reduced the average time needed for a user to perform each step, which may be considered an improvement against the joystick. Following the tendency of the averages, the standard deviation also shows reduction when the Phantom is used in the anesthesia and puncture stages. The Wilconox test statistical analysis shows that there are relevant differences between the two input devices for all three steps of the medical procedure, if we assume a 95% reliability (asepsis p-value = 0.0002441, anesthesia p-value = 0.02454 and puncture p-value = 0.02954).

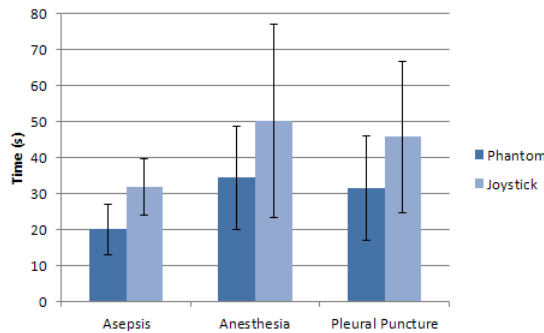


Fig. 22. Averages and standard deviation for the pleural tap procedure for Phantom Omni and Joystick input devices.

Error rates are also distinct. The Joystick led to a 14% error rate, as shown in Figure 23, while the Phantom Omni led to a 20% error rate, as reported in Figure 24.

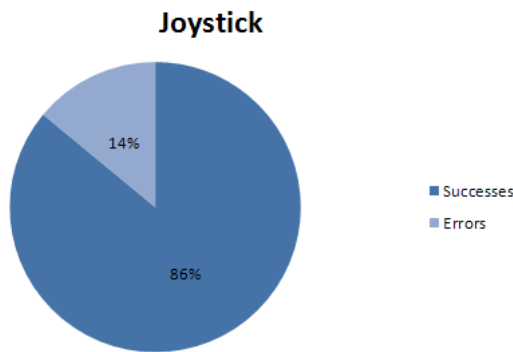


Fig. 23. Success rate in the medical procedure with a Joystick.

The hypotheses for this difference are: (1) the volunteers gained greater familiarity with the Joystick and (2) the Phantom Omni is more realistic, requiring that the users have greater eye-hand coordination. The Joystick doesn't rely on the user's arm movement as the Phantom does. We do not investigate these hypotheses in this work.

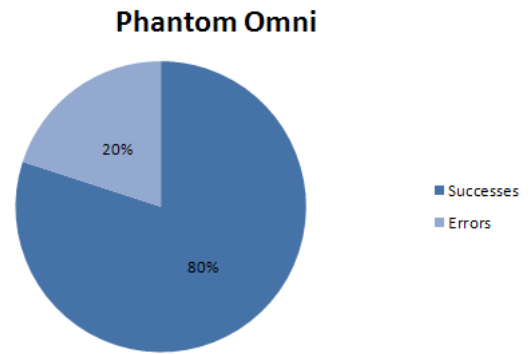


Fig. 24. Success rate in the medical procedure with a Phantom Omni.

Although the error rate is greater when the user operates the system through a Phantom Omni, the majority of the users reported that they prefer to handle the medical instruments with the Phantom, as reported in Figure 25. That reinforces the idea that the use of the Phantom Omni turns the simulation more realistic. However, 14% of the volunteers reported some discomfort or pain in the arm or hand while handling the phantom. Further evaluation of ergonomics is called for if the device is to be used in long periods.

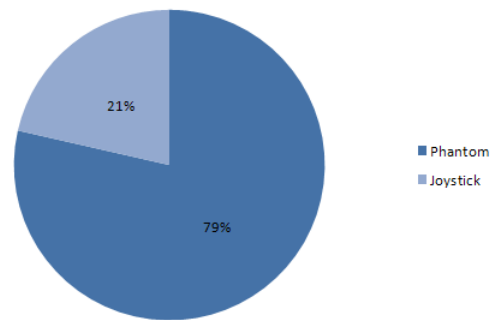


Fig. 25. Preference of volunteers with regard to the input device options.

## VI. CONCLUSION

This work introduced a VR system for medical training of the pleural tap procedure. The system is composed of 3D models of a virtual surgery room, a virtual surgeon, and a virtual patient, including the appropriated internal organs. The system can be used as a start point for the simulation of other medical procedures. Some software modules were built aiming at helping the user to learn the procedure at stake. Such modules also worked to evaluate the user actions while performing the procedure. We also evaluated the usability and possible improvement in user performance when non-conventional devices, such as a HMD and a Phantom Omni, were used instead of the usual keyboard/mouse, joystick and standard LCD monitor options. In order to evaluate the system, as well as the options of input/output devices, twelve volunteers who had no medical expertise were invited to use the application.

Based on the results achieved, we can assert that the modules implemented are efficient in providing training to the users with regard to the pleural tap medical procedure. The

volunteers succeeded in performing the procedure in the right order and without skipping steps. Moreover, we could notice that as users spent more time training in the system, there is a clear tendency for them to reduce mistakes and improve accuracy. Such indications suggest that the system has been effective in providing basic training of the procedure. It is also able to properly evaluate the user performance.

With regard to the usage of the Sensics xSight 3120 HMD, the results suggest that it contributes in the initial stages, as it improves the performance of the users. Such improvement was observed specially in the anesthesia and pleural tap stages, which are the stages that require greater precision by the user, even though the HMD resulted in some visual. We can also observe that although the HMD imposed a series of restrictions to the simulation, it does not restrict the system usability, even improving the training performance in some cases.

The results of this work also demonstrate that if one needs to use a simple LCD monitor as graphical output, similar final performance can be reached if compared with users that used the HMD. The users do need a longer training time, but the final performance gets to be equivalent. This is important, as it allows developers to design training simulators that rely on simpler devices, eventually improving their graphical output whenever possible.

With regard to the Geomagic Phantom Omni, the results indicate that it reduces the time needed for a user to perform a given task, even though it leads to a small increase in the error rate. However, the Phantom provides more realism in the execution of the tasks, i.e. the real procedure cannot be performed with a joystick, so that the phantom allows the user to act closer to what he/she needs to do in real life, even though the Phantom is not precisely similar to the actions, as the stylus does not mimic properly the behavior of a syringe. It is still better than a joystick. The volunteers chose the Phantom as the more realistic device. A customized interface, replacing the stylus, could further improve its similarity to the real setup.

Improvements to the system can be made by adding some features, such as: extending the hospital environment; implementing more medical procedures as well as procedures to be performed by other health professionals, like nurses and technicians; adding a physiological engine capable to simulate general anesthesia as well as patient reactions to medications; allowing the system to be collaborative, i.e. to allow several users to jointly perform a given procedure (nurses, technicians, surgeons, etc.) in a collaborative virtual environment – CVE. On the other hand, we can verify the usability of more advanced devices, such as the Cyberforce/Cybergrasp/Cyberglove system along with a HMD, for a fully immersive user interaction with the virtual environment.

#### ACKNOWLEDGMENT

This work was financially supported by the Brazilian Ministry of Science and Technology CNPq agency (Processes 309318/2011-4 and 573710/2008-2), State of Rio de Janeiro

FAPERJ agency (Processes E-26/103.241/2011 and E-26/170.030/2008), as well as Brazilian Ministry of Education Capes agency and the PCI/LNCC Program. We also thank the volunteers who helped us participating in the tests we report here.

#### REFERENCES

- [1] G. C. Burdea and P. Coiffet, "Virtual Reality Technology". Jonh Wiley & Sons, Inc, 2003.
- [2] A. Pazin Filho and S. Scarpelini, "Simulação: Definição". Medicina Ribeirão Preto. Vol. 40, No.2, April/June 2007, 162-166.
- [3] A. Ziv, P. R. Wolpe, S. D. Small, and S. Glick, "Simulation-Based Medical Education; an Ethical Imperative". Academic Medicine. Vol. 78, No. 8, Agosto 2003, 783-788
- [4] Sensics, "Products: xSight Panoramic HMDs". Available at: <<http://sensics.com/products/xSight>>, February 2012.
- [5] Geomagics, "The Geomagic Touch Haptic Device". Available at: <<http://geomagic.com/en/products/phantom-omni/overview>>, August 2013.
- [6] R. M. Satava and S. B. Jones, "Virtual Environments for Medical Training and Education". Presence – Teleoperators and Virtual Environments. Vol.6, No. 2, April 1997, 139-146.
- [7] J. M. Taekman, N. Segall, G. W. Hobbs and M. C. Wright, "3DiTeams – Healthcare team training in a virtual environment". Anesthesiology, Vol. 107, October 2007, A2145.
- [8] Unreal Technology, "Unreal Engine". Available at: <<http://www.unrealengine.com/>>. May 2012.
- [9] HumanSim, "Medical Education and Training". Available at: <<http://www.humansim.com/>>. May 2012.
- [10] Y. Fukui and N. T. Smith, "Interaction among Ventilation, the Circulation, and the Uptake and Distribution of Halothane – Use of a Hybrid Computer Multiple Model: I. The Basic Model". Anesthesiology, 54:107-118, 1981.
- [11] R. S. Torres, H. H. Bísaro, L. V. Araújo and F. L. S. Nunes, "ViMeTGame: A serious game for virtual medical training of breast biopsy". SBC Journal on 3D Interactive Systems, volume 3, number 3, 2012.
- [12] R. Sales and R. Onishi, "Toracocentese e biopsia pleural". Jornal Brasileiro de Pneumologia, 32:170-173, 2006
- [13] MD. Saúde, "Derrame Pleural – Sintomas e Tratamento". Available at: <<http://www.mdsaude.com/2010/04/derrame-pleural-tratamento-sintomas-e.html>>. July 2012.
- [14] WorldViz, "Vizard: virtual reality toolkit". Available at: <<http://www.worldviz.com/products/vizard4/index.html>>. January 2012.
- [15] TurboSquid, "Hospital Surgery Room". Available at: <<http://www.turbosquid.com/FullPreview/Index.cfm/ID/438222>>. February 2012.
- [16] Anatomium 3D, "The P-Set Series". Available at: <<http://www.anatomium.com/n-p1.html>>. January 2012.
- [17] Autodesk, "3ds Max". Available at: <<http://usa.autodesk.com/3ds-max/>>. February 2012.
- [18] Sensics, "xSight and piSight - Head Mounted Displays User's Guide". Sensics, October 2010.
- [19] S. Siegel and N. J. Castellan Jr. "Nonparametric Statistics for The Behavioral Sciences". 2 ed., Editora McGraw-Hill, 1988.
- [20] R. A. Siscoutto, F. Szenberg, R. Tori, A. B. Raposo, W. Celes and M. Gattass, "Estereoscopia". In: Realidade Virtual: Conceitos e Tendências. Livro do Pré-Simpósio SVR 2004, Cap. 11, p. 179-201. Eds: C. Kirner e R. Tori. Editora Mania de Livro, São Paulo, 2004.
- [21] J. V. Monteiro, "3D na oftalmologia". Universo Visual. Available at: <<http://www.universovisual.com.br/>>. March 2012.