A Natural Interface for the Training of Medical Personnel in an Immersive and Virtual Reality System

Alberto Del Bimbo, Andrea Ferracani,
Daniele Pezzatini, and Lorenzo Seidenari

MICC - University of Firenze, Firenze, Tuscany, ITA
delbimbo@dsi.unifi.it,
{andrea.ferracani,daniele.pezzatini,lorenzo.seidenari}@unifi.it
http://www.micc.unifi.it/vim/
Universitá degli Studi di Firenze, Media Integration and Communication Center,
Viale Morgagni 65,
Firenze, Italy

Abstract. In this paper we present an immersive system, developed for the RIMSI project, designed for the training of medical and paramedical personnel in emergency medicine. Virtual reality systems have been used recently in combination with natural interaction systems for patients' rehabilitation but are proving especially useful for educational applications in the field of medical training and teaching. The idea is to mimic real three-dimensional environments and to make them available through natural interaction to enable users to explore, interact, collaborate and assist patients in virtual scenarios that otherwise could not be simulated easily in the real world without an high cost (e.g. to prepare the environment for a plane or a car crash) or possible risks of injuries to actors involved in the simulation (e.g. perilous situations like a gas leak). In addition, these simulation scenarios, using a digital environment, are repeatable and can be recorded, easing the process of errors' highlighting. The RIMSI prototype provides a virtual first aid scenario with interactive 3D graphics which can be controlled and navigated through a natural gesture interface based on Kinect[TM].

Keywords: virtual reality, medical training, body tracking.

1 Introduction

Educational virtual environments are used, through computer applications, in a variety of fields of study to facilitate the learning and the acquisition of knowledge and to allow users to feel immersed in a common environment, in specific situations, in order to ease the exchange of skills. Network systems have been designed, in the recent past, to solve issues related to distance learning, mediated by specialists; but often, in such systems, priority has been given to theoretical rather then practical learning. In systems featuring natural interfaces instead

medical trainees can realistically manipulate 3D anatomical parts of the body, reproduced in graphics, or assist virtual injured people in dangerous scenarios. This aspect is very important in preparing trainees psychologically to deal with real situations.

In the field of medical training the main problem of digital simulation systems emerged in recent years is, on the one hand, the excessive orientation towards the design of too configurable systems, given the technical difficulties due to the correct interfacing between hardware, software and the often complicated medical procedures, difficulties that usually lead to the development of very simple prototypes. On the other hand, the proliferation of very specialized systems has resulted in applications not sufficiently adaptable to several scenarios. Furthermore, for the most part, all these systems are not fully immersive.

The Cybermed framework [6], for example, is a system for medical training via computer network. It has advanced features for configuring some aspects of the software like the number of participants, the manipulation of objects, the type of device (mouse or haptic systems), the definition of the actors for remote mentoring and distance learning (tutor or participants), but does not really allow an exhaustive characterization of the situation or the definition of complex medical procedures. SOFA [2], ViMet [3] and Gipsy [8] frameworks instead are open-source projects which feature an high level modularity and rely on the capability of a multi-model representation of simulation models (deformable models, collision models, instruments) exposing appropriate APIs. Although they easily could allow to simulate a scenario, e.g. a laparoscopy, they neither provides a real natural and immersive interface nor they go beyond enabling the system to respond to punctual stimuli (e.g., in the laparoscopy use case, deformation of the liver and collisions with the ribs). Finally, Spring [7] is a more specific mouse based desktop framework for real-time surgical simulations which allows a basic configurability of patient-specific anatomy.

As regard to more immersive solutions Honey et al. have created a virtual environment in Second Life for hemorrhage management[13]; Cowan et al. [11] have developed a system for the inter-professional education (IPE) for critical care providers where an immersive 3D situation is accessed across the network in a "multi-player online" environment, allowing trainees to participate as avatars from remote locations. These systems have all the benefits of realism but still lack the naturalness of human gestures and actions that can be performed in solutions exploiting natural interaction.

Several natural interfaces featuring natural interaction systems have been used in medical rehabilitation, although sparsely, leveraging different technologies (e.g. Nintendo Wii, PlayStation EyeToy) or the Microsoft Kinect, as part of physical therapy. Lange et al. [12] of the Institute of Creative Technologies of the University of Southern California, for example, have developed a Kinect-based rehabilitation game "JewelMine" that consists of a set of balance training exercises to stimulate the players to reach out of their base of support. These kind of softwares take advantage from the realism and the naturalness of the situation to

encourage physical exercise but are not common in medical operators training. In this sense they are patient oriented and not trainees oriented.

2   RIMSI: Using Natural Interaction for Medical Training

The use of natural interaction combined with an intelligent and realistic virtual environment allows to improve training systems and to overcome the difficulties which arise, for example, in an haptic enabled environment where it is usually only the teacher that shows the type of movements to be performed on a 3D model [1] and there is only one feedback from teacher to learner via network, as providing all students of a haptic tool would be very expensive.

Furthermore for virtual environments and natural interfaces to be effective, it is essential that the scenarios are very realistic [9]. With scenery we mean the places, the events, but also the virtual humans involved in a particular situation. Such humans should be able to give realistic feedback, express emotions, react to natural stimuli. Many medical simulation centers are currently using patient simulator systems provided with mechanical patients: trainee's task is to carry out some emergency maneuvers on a patient whose vital signs are in the meantime monitored and controlled by training doctors present in another room. The trainees have the possibility to evaluate these values through a vital sign monitor. Their behaviors influence the responses and the reactions of the patient, guided interactively by trainers. The so-called debriefing usually follows this phase. During the debriefing the trainers and the team of trainees discuss actions and mistakes often with the aid of a video record of the whole session: this is the stage where trainees learn more. Laboratories equipped with high and low fidelity medical mannequins are expensive to build and to maintain. Moreover, the traditional classrooms or debriefing for learning have a teacher - centered approach. In the training sessions mediated by mannequins users undergo reactions and behaviors whilst in natural interaction systems individual responsibility, represented by the avatar or character in the game, is more relevant. Our prototype (RIMSI) uses a learner-centered teaching approach where he is the user himself, through interaction, to guide the learning process resulting in a better memorability. As showed by Villneuve et al. the generation of digital natives prefer the use of technology in learning [10].

RIMSI has been designed with the aim of becoming a flexible framework in the context of very specific scenarios. As use case we propose a first aid BLSD (Basic Life Support Defibrillator) scenario providing the software with a configuration interface that allows educators to change several environment variables, such as the location of objects, and some events and actors in the scene, etc. (Figure 1).

The core of the system is on the natural way in which users can exploit the simulation environments. Trainees can move in the space like in real life and interact with realistic objects and avatars (Figure 2). This is very important from the point of view of training as a naturally usable environment eases the learnability and the memorability of the procedures.

Fig. 1. The graphic configuration interface

Fig. 2. Avatars performing actions in the environment

3   Use Cases and Scenarios

Simulation in medical training can address a wide range of aspects, spacing from the application of emergency procedures to the enhancement of a specific skill. Usually, training of specific manual and practical skills is performed via haptic devices that simulates medical tools in a very accurate way or by using body part replication [4] [5]. Since this kind of training require a very accurate evaluation of the results (e.g. a few millimeters error may results in a critical outcome for the patient) it would be very difficult to obtain a good simulation in a complete virtual environment. Anyway, other aspects of medical training such as the application of procedures, environmental awareness or communication with team members can gain a great improvement from virtual simulations. Some medical scenarios, such as Basic Life Support in emergency medicine or sepsis treatment, present a well-known sequence of actions to be taken in order to accomplish a correct patient assistance.

According to this, the proposed system consists of a virtual reality simulator for emergency situations in which trainees are asked to take some actions on the environment and to apply medical procedures on the patient depending on the proposed situation. The system is configurable and it provides a scenario editor through which the medical instructors can define the type of the simulation. It is possible to place patients in different positions in the environment, arrange the environment itself (e.g. pre-hospital first aid, emergency room) and to add to each scenario several additional threats to the safety of patients and medical operators. Furthermore, its also possible to insert different patients conditions, like his state of consciousness and his body pose, and different responses to therapies that require appropriate standardized procedures. In this way, a different virtual simulation can be generated for each training session.

The simulator allows to use Non-Playing Characters (NPCs) in the scenario. NPCs can have different roles: they may be members of the medical team performing the procedure, but they can also be relatives of the patient or simple by-standers. In both cases they are employed to deliver information about the patient or the environment or to give some hints about the correct procedure to the trainees.

We present a possible use case scenario: a basic life support and defibrillation (BLSD) procedure with a patient inside a house with a gas leak. In this scenario the trainee, associated with the doctor character, can perform the full procedure assisted by a NPC, the medical assistant. As the simulation begin, the assistant deliver informations about the house situation (e.g. she claims she is smelling gas). The player can navigate in the virtual room to check the environment safety or go directly to the patient to check his conditions. In order to perform a correct procedure on the patient, the player should locate the gas source, block the leak and open a window before a certain amount of seconds. When the room is safe, the assistant give feedback to the player, so he can approach the patient in order to check his state. Then, to complete the procedure, the trainee have to choose between different action proposed by the simulator and perform them in the correct order and within a certain time. Errors are reported by the medical

assistant with audio e textual feedback. Timers are used in order to assess the player performance and to consider the need of different procedures. To finalize the procedure, the player has to grab the proper tools and activate them in the correct way. As an example, if a defibrillator is needed, the player has to grab it, turn it on and wait until it is completely charged before activating it. The simulation ends successfully if the player completes the procedure correctly or unsuccessfully if he commits too many errors or delays in taking decisions.

4   Navigation and Interaction System

The proposed system is a free-roam game. It means that there is no predefined order in which actions need to be taken in order to obtain accomplishments. This is also a way to improve realism since, as in real life, a user may decide to assist a patient before another or check the environment to assess its safety before beginning the medical procedures. Obviously, some actions need to be taken in the correct order to perform a correct training session. As an example, if multiple patients are in the scene, the more severely injured should be assisted first.

The navigation in the RIMSI virtual environment is made possible by a gesture driven Kinect^{TM}interface, through which the variation of user posture is evaluated in order to interact with the scenario. Simple or complex gestures can be recognized with the use of a depth camera and a skeletal tracking system [14]. Based on information extracted form the depth and RGB cameras, the Microsoft Kinect APIs track the position and orientation parameters of all the joints of the skeleton model in real time. For every gesture that users can perform in the simulator, we defined a simple Finite State Machine (FSM) and implemented the relative FSM recognizer. In general, gesture recognition relies only on some specific body part, so for every gesture we only consider significant skeleton joints.

Defined gestures can be found in Table 1, along with details about considered skeleton joints and actions performed by the virtual character.

Gestures have been dened according to the requirements of the simulation context. Since the training sessions should consist in applying the correct procedures and not in learning the semantic of predefined gestures, users should interact in the virtual environment as they would do in real life. Simplest actions and behaviors can be easily translated in natural body gestures or poses. As an example, the rotation of the the torso will cause a rotation of the character, while the action of bending will be used to check the conditions of the patient.

However, not every phase of the medical procedure can be associated to natural gestures. When the trainee has to face a choice between different options, virtual menus will appear to allow the player to choose the action to perform. We defined a "pointing" gesture that produce a cursor on the virtual environment, allowing user to select one of the options by maintaining the persistence over it. Figure 3 shows some phases of the interaction.

Table 1. Associations between gestures and character actions

Gesture	Body pose	Actions
Take / Grab	One of the hands moves in front of the body (on Z axes)	The character grab objects in front of him, open/closes handles
Rotate	The rotation of the torso is evaluated using x,y values of users shoulders	The camera rotate in order to explore the scene
Walk-in-place	Sequences of x,y values of legs' joints are evaluated to check several phases of the walk	The character moves along the direction he is facing
Bend	Sequences of decreasing y values of chest and knees	The character bends on his knees in order to examine the patient
Point	User's hand in front of his body, x,y values are normalized respect to the user's position	A cursor is visualized in order to allow multiple choices

The only limitation of the standard Kinect SDK has is the lack of a hand pose recognition algorithm. The recognition of the hand pose could improve the system capabilities, allowing to recognize on/off actions such as activating, grabbing or manipulating an object in the 3D world. We propose to employ an improvement to this system based on a robust hand pose classifier we developed [16]. The proposed module is trained on a large dataset (30k+) of hands in open and closed status recorded at several distances, from eight subjects of different genders. This vision module improves the responsiveness of the interface and enables the activation and manipulation of 3D objects using informations about the state of the hands of the player. In Figure 4 an example of the detector output is shown.

As an example, in the BLSD scenario, the player could grab the electrodes one at a time or with both hands by simply pointing at them in the 3D space and closing the hand(s), and then place them on the patient chest by opening his hand(s).

The vision module represents a hand, segmented with a skeletal tracker, with a regular grid of 5 SURF[15] descriptors. This feature is fed to a nonlinear SVM classifier that predicts the state of the hand. This module has a very low overhead in terms of computation, in fact a detection can be obtained in less than a frame, although this detection can be noisy. The classifier is cascaded with a temporal Kalman filter that outputs a smooth estimate of the hand state. The σ parameter can be varied obtaining a trade-off between responsiveness and robustness of the system. Overall the hand status recognition module has an extremely high reliability with an accuracy of to 98.95%.
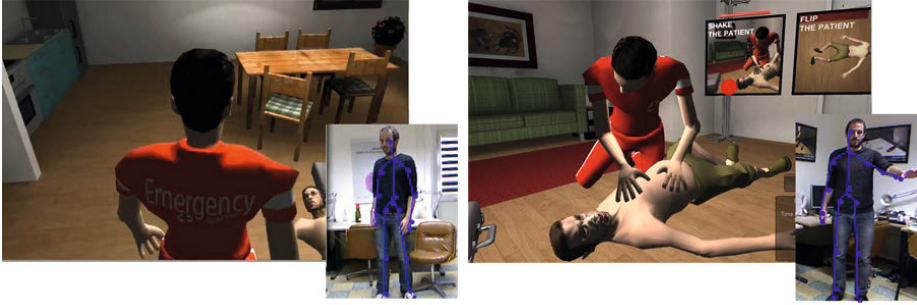
Fig. 3. Avatars performing actions based on user's gestures. On the right, a virtual menu with multiple options is shown.
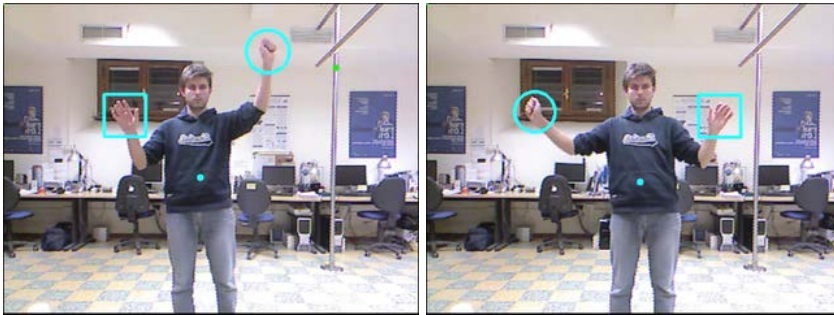

Fig. 4. Hand status detection: square (open hand), circle (closed hand)

The game is based on the Unity3D game engine while gesture detection module has been developed in C# using the official Microsoft Kinect SDK. Scenes and models have been realized with Maya.

5   Conclusion

The RIMSI prototype (RIMSI) is part of the so-called serious games technologies for the improvement of skills and knowledge in medicine. Serious games provide an interactive environment where students and teachers are represented by avatars. Currently most of the simulated training sessions in medicine occur in real environments such as laboratories but more realistic digital simulations that facilitate learning in a safe environment can be implemented through emerging technologies such as tracking and VR environments.

RIMSI features a 3D virtual environment, high resolution graphics and natural interaction interfaces to improve the effectiveness of the training of medical personnel in emergency situations. It provides an immersive and intelligent digital environment which can respond to natural gestures and actions of more team members in a simulated situation of emergency. The system has been developed to improve the learning methods currently in use and to solve some common issues.

In particular RIMSI allows to reduce costs of simulations, optimize some critical phases,like debriefing, with a better management of the time spent and of the use of personnel in the centers of simulation. Furthermore, compared to simulations with mannequins, the RIMSI system allows: 1) to have an intelligent system that responds automatically to user actions, 2) have continuous feedback, 3) to simulate an immersive environment such as that with the use of mannequins, but at a lower cost, enhancing learnability and memorability of procedures 4) to be corrected on the fly and have system's feedback in realtime (the debriefing phase assumes instead a period of time between the wrong action and the correction).

Interactive sessions with RIMSI can also be recorded and used as training materials. In this way a large database can be build quickly to evaluate errors and the most common difficulties of trainees. This data, in addition to increase the effectiveness of training, can be useful to assess how well these simulation environments work and how they can be improved. In fact, future work will include the testing of the prototype developed by the research unit and subsequent focus groups to discuss critical points.

References

1. Gunn, C., Hutchins, M., Stevenson, D., Adcock, M., Youngblood, P.: Using collaborative haptics in remote surgical training. In: Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC 2005), Italy (2005)
2. Allard, J., Cotin, S., Faure, F., Bensoussan, P.J., Poyer, F., Du-riez, C., et al.: SOFA - an Open Source Framework for Medical Simulation. Medicine Meets Virtual Reality 15, 1–6 (2007)
3. Oliveira, A.C.M.T.G., Botega, L.C., Pavarini, L., Rossatto, D.J., Nunes, F.L.S., Bezerra, A.: Virtual Reality Framework for Medical Training: Implementation of a deformation class using Java. In: Proceedings of the SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (SIGGRAPH 2006), Hong Kong, pp. 347–351 (2006)
4. Basdogan, C., Ho, C.-H., Srinivasan, M.A.: Virtual environments for medical training: graphical and haptic simulation of laparoscopic common bile duct exploration. IEEE/ASME Transactions on Mechatronics 6(3), 269–285 (2001)
5. Coles, T.R., Dwight, M., Nigel, W.J.: The role of haptics in medical training simulators: a survey of the state of the art. IEEE Transactions on Haptics 4(1), 51–66 (2011)
6. Sales, B.R.A., Machado, L.S., Moraes, R.M.: Interactive Collaboration for Virtual Reality Systems related to Medical Education and Training. Technology and Medical Sciences, 157–162 (2011)
7. Montgomery, K., Bruyns, C., Brown, J., Sorkin, S., Mazzella, F., Thonier, G., Tellier, A., Lerman, B., Menon, A.: Spring: A General Framework for Collaborative, Real-time Surgical Simulation. In: Medicine Meets Virtual Reality (MMVR 2002), pp. 23–26 (2002)

8. Goktekin, T.G., Çavuşoğlu, M.C., Tendick, F., Sastry, S.: GiPSi: An Open Source Open Architecture Software Development Framework for Surgical Simulation. In: Cotin, S., Metaxas, D. (eds.) ISMS 2004. LNCS, vol. 3078, pp. 240–248. Springer, Heidelberg (2004)

9. Lok, B., Ferdig, R.E., Raij, A., Johnsen, K., Dickerson, R., Coutts, J., Stevens, A., Lind, D.S.: Applying virtual reality in medical communication education: current findings and potential teaching and learning benefits of immersive virtual patients. Virtual Real. 10(3), 185–195 (2006)

10. Villeneuve, M., MacDonald, J.: Toward 2020: Visions for nursing. Technical Report. Canadian Nurses Association, Ottawa, Ontario, Canada (2006)

11. Cowan, B., Shelley, M., Sabri, H., Kapralos, B., Hogue, A., Hogan, M., Jenkin, M., Goldsworthy, S., Rose, L., Dubrowski, A.: Interprofessional care simulator for critical care education. In: Proceedings of the Conference on Future Play: Research, Play, Share (Future Play 2008), pp. 260–261. ACM, New York (2008)

12. Lange, B., Koenig, S., McConnell, E., Chang, C., Juang, R., Suma, E., Bolas, M., Rizzo, A.: Interactive game-based rehabilitation using the Microsoft Kinect. In: IEEE Virtual Reality Short Papers and Posters, VRW (2012)

13. Honey, M.L.L., Diener, S., Connor, K., Veltman, M., Bodily, D.: Teaching in virtual space: Second Life simulation for haemorrhage management. In: Ascilite Conference, Aukland (2009)

14. Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., Moore, R.: Real-time human pose recognition in parts from single depth images. Commun. ACM 56(1), 116–124 (2013)

15. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. Computer Vision and Image Understanding 110(3), 346–359 (2008)

16. Bagdanov, A.D., Del Bimbo, A., Seidenari, L., Usai, L.: Real-time hand status recognition from rgb-d imagery. In: International Conference on Pattern Recognition (2012)