Users' Assessment Based on Interactions with Haptic Devices for Application in Medical Training Simulation

Ronei Marcos de Moraes
Department of Statistics
Federal University of Paraíba (UFPB)
João Pessoa/PB - Brazil
ronei@de.ufpb.br

Liliane S. Machado
Department of Informatics
Federal University of Paraíba (UFPB)
João Pessoa/PB - Brazil
liliane@di.ufpb.br

***Abstract*—Virtual reality have been used to provide training systems in several areas, particularly in medicine. In that area, user's interactions in a virtual environment are modeled and they are compared with predefined classes of performance to know how much users are prepared to perform that procedure on human beings. In this paper is discussed the process of design of a simulators in order to detail how the simulation was implemented to capture interaction data to be used in an assessment.**

Keywords-component; Medical Training; Haptic Devices; Users' Assessment; General Bayesian Networks.

I. Introduction

The training of medical procedures is related to the practice of psychological and psychomotor aspects. The use of the sight and the touch are mandatory in most of the procedures. However, smell and hearing are also important in several of them. Despite of the evolution of medical procedures and also the availability of educational tools, there is a lack of computer-based simulation tools accepted for certification of professionals in Medicine. However, the Medicine does not have well established metrics that can indicate if a person is qualified or not to perform some procedure. In general, new physicians are trained by a set of activities that also include watching and helping real procedures by the supervision of a senior physician.

Critical situations experimentation, repetition, performance standards and well defined goals to be reached are some issues that have been researched in order to define metrics for assessment of physicians. Although simulated scenarios could present less elements than real environments, they have been increasingly used for training. Their development demands a multidisciplinary team to design and observe the most adequate way to present the procedure tasks for users. In those simulators, users actions should be monitored in details and used to verify their performance [5]

Virtual reality (VR) integrates methodologies and technologies able to explore human senses in computer applications and have been used to the construction of simulators for the medical field. Haptics have been frequently used in those simulators of medical procedures to explore touch sense. The human processing of the this sense is accurate and must be carefully planned in the design of medical simulators. Most of haptic devices available can detect six degrees of freedom (DOF) movements and provide force feedback with two or more DOF. However, the sample rate of movements performed with those devices is around 1000Hz. It demands a fast processing and generates a large amount of data at each second.

The main challenge of VR simulators for medicine is to provide realistic tools composed by well defined tasks to be performed by users and decision making components to assess their performance.

Several authors suggests a typical method for medical training assessment based on videotapes post-analysis [5]. However, that methodology is restricted and obsolete didactically. Scott et al. [15] showed that direct assessment can improve student's learning and is better than assessment based on videotapes post-analysis. Gaba et al. [2] showed, using statistical tests, situations in which the agreement between raters varied, which can commit the assessment. As short term memories are somewhat volatile and after training they may be lost, and procedural memories need practice and observation, it is natural that an assessment system should give an answer quickly, to increase the probability that the trainee is able to remember details of their training and better understand the assessment received [11]. Some authors [8] proposed the concepts of online and offline assessment of training. An assessment method can be called online if its answer to user is less than one second, i.e. it is immediate to the user. In another cases, the assessment is called offline. Online methods must be coupled to the training system. For this combination work, an online method must present low computational complexity in order to not compromise the simulation performance. However, they must also have high accuracy that doesn't compromise the quality of assessment [7]. In general, given a user's training data, an online assessment should be performed in less than one second. Since interaction by haptic devices can produce a large amount of data, the assessment methodology must be robust to use all simulation data, including those produced by haptics, in the assessment process to provide accurate results.

In this paper is discussed the process of design of a simulators in order to detail how the system was implemented to capture interaction data to be used in an assessment. Results about the performance of the assessment method are presented.

II. What is necessary to develop a simulator based on VR?

A simulator based on VR must contain the same steps present in a real procedure. The tools (hardware and software) to explore all steps and features in real-time demands a lot of processing and can compromise the real-time performance. Thus, each VR simulators for medicine is defined and developed to deal with specific parts of the human body and can present limitations [10]. Despite of provide more realism in some features than others, the interaction aspects must be preserved as much as possible.

The main challenge related to the development of simulators is the integration of the tasks related to each sense and their synchronization. All information and all application tasks must be synchronized to avoid delays or inconsistencies that can compromise the realism and it can includes methods for collision detection, deformation, stereoscopic visualization and tracking, among others. In Figure 1 presents a diagram with some tasks of a simulator based on VR for medical training.

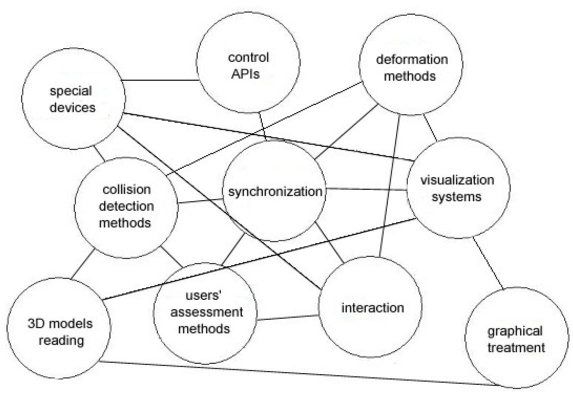

Figure 1. Tasks of a VR simulator and their relation [10, p. 327].

In a traditional bone marrow harvest for transplant, multiple punctures are required (up to 20 per patient) in order to obtain enough bone marrow for the transplant. Each puncture requires a specific application of force to penetrate the bone (to reach the marrow) without overstep it, which can be performed in approximately 5-10 seconds (without obtaining bone marrow), depending on physician's dexterity. However, some special situations can occur and demand specific skills of physician: the repetitive procedure can cause loss of control or fatigue of physician [9], and when performed on a child it must be concluded in a short time due to low dosage of anesthesia.

Another problem is the proximity between a first puncture and the next one (generally about a centimeter far), which requires sufficient dexterity of the operator. The properties of the bone region can change after each harvest, due to the punctures. Additionally, in pediatric oncology is easy to find children with soft bone structures or bones with deterioration caused by disease which demands more attention and dexterity in the harvest procedure [9].

Design and develop realistic procedures is possible and demands planning and committed development teams. The procedure, the goals to be reached and its main requirements are good examples of how the knowledge must be explored and used in the design process. However, according to the features of a procedure, may be necessary to enrich the realism of some functions and do the opposite in others. This can happen to guarantee that the tasks, models or interactions could represent their real representation. Thus, elements not so important for the training simulation can present less details. This kind of decision is discussed and defined by the development team during the design.

The present technologies provide several devices and algorithms to explore human senses, mostly sight, touch and hearing. Although some are still considered expensive for general public, their application in training systems can be justified by the impact of their results. Thus, several of them have been used for stereoscopic visualization, 3D interaction with force feedback and interactive deformation, as example.

The reuse of software and portability in programming of devices have been provided by frameworks. This kind of tool is used for agile development of VR medical applications. However, flexibility allows fast integration of new methods and make frameworks be used also for test of new techniques and methodologies [10].

III. The Bone Marrow harvest Simulator

The bone marrow harvest simulator was developed to allow training of the bone marrow harvest for transplant [16]. The procedure is composed by three steps which were considered in the development of the simulator. The first step is related to the observation of the anatomy, the second is the palpation of the patient pelvic region, and the third contents the punctures for bone marrow harvest. Since the main activity of the procedure is related to the third step, the simulator was designed in details to represent this step. Also, a haptic device was used to provide realism and co-relation between real and virtual environments.

The bone marrow harvest simulated was re-implement in 2005 to test the features of the CyberMed [10] and observe how frameworks can speed up the development process.

Since this simulator included a rigorous process of design, it and the framework CyberMed have been used to test several assessment methodologies. The goal is observe the performance and the efficiency of the assessment methodologies in a simulated training. In this case, the haptic interaction present in the puncture step must be acquired and used to asses the user and indicate their capacity to perform the simulation is a real patient.

An assessment method based on General Bayesian Network was implemented using CyberMed [6] and tested using the bone marrow harvest simulator. This assessment method is described in the next section.

IV. Assessment of Procedure Perfomed by Users

For the reader's better understanding, we provide a brief review about Bayesian networks and general Bayesian networks. A Bayesian network is a probabilistic model, which can represent a set of probabilities distributions from all variables in a complex process and also establish their relationships [14]. Formally, a Bayesian network is defined as directed acyclic graphs, denoted by G and a probabilistic distribution denoted by P. The graph G is a set of nodes and oriented arcs, where nodes represent variables in process and oriented arcs encode conditional dependencies between that variables [14]. The dependencies are modeled by specific conditional probabilistic distributions [4].

Several kinds of Bayesian networks can be found in literature. They can differ on their graph structure and types of dependencies relationship which can model. Cheng [1] proposed a classification for Bayesian networks in five classes, taking into account those characteristics: Naive-Bayes, Tree Augmented Naive-Bayes, Augmented Naive-Bayes, Bayesian Multi-net nets and General Bayesian Networks (GBNs). To choose a specific Bayesian network depends on statistical relationship between variables which describe that process. This is critical, because it changes the final results. The GBN is a generalized form of Bayesian networks, which allows nodes to form an arbitrary graph, rather than just a tree. Another important characteristic is that each child node cannot be connected to the final classes of assessment. In that network, the dependencies between nodes can adjust itself to real dependencies. Thus, it is possible to verify dependencies between variables during network modeling and put them in structure nodes of GBN, which did not occur in other structures [2].

Formally, let be the classes of performance in space of decision $\Omega = \{1,...,M\}$ where M is the total number of classes of performance. Let be w_j, $j \in \Omega$ the class of performance for a user and X_i, $1 \leq i \leq n$, represents a node in GBN with n as the number of nodes in a graph. The joint probability distribution in GBN for an event is done by:

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_n \mid X_{n-1}, X_{n-2}, \ldots, X_1 \quad (1)$$

where $P(X_1, X_2, \ldots, X_n)$ is the joint probability distribution and $P(X_n \mid X_{n-1}, X_{n-2}, \ldots, X_1)$ is the conditional probability of X_n conditioned by its predecessor nodes $X_{n-1}, X_{n-2}, \ldots, X_1$.

If the conditional independence between variables is verified, this permits simplifications in equation (1). Then,

$$P(X_1, X_2, \ldots, X_n) = P(X_1 \mid w_j)\, P(X_2 \mid w_j)\, P(X_3 \mid w_j) \ldots \quad (2)$$

The probability nodes are associated to probability distribution. For example, a node A can have a Gaussian distribution and a node B, which depends on A, can have a bivariate Gaussian distribution, with a mean vector and a covariance matrix [3]. The structure of GBN is learned from data, as well as the parameters of conditional probabilities. Using probabilities calculus is possible to find dependencies among nodes in a Bayesian network. If those dependencies are founding and, if is possible to assume Gaussian distribution for nodes, dependencies can be estimated using multivariate linear regression [14]. Scores are used to help estimate the final structure of GBN for each class of assessment. In a first moment a network is created with all independent nodes and an initial score is calculated. Next, all combinations are searched and an arc is designed between two nodes for which an increment of initial score is obtained. Then, the parameters for that nodes set are re-estimated using linear regression. This cycle is repeated until total network score is less than a predetermined value or a fixed number of cycles.

In a usual situation, a physician should calibrate the system previously, according M classes of performance defined by him, where M may be any integer number greater than one. The calibration process consists in to execute several times the procedure and to classify each one according to classes of performance. The number of classes of performance can be defined, e.g. as $M=3$: 1) correct procedures, 2) acceptable procedures, 3) badly executed procedures. So, the classes of performance for a trainee could mean: "you are well qualified", "you need some training yet", "you need more training".

V. Results

The assessment method based on General Bayesian Network was implemented in Cybermed, expanding the framework capabilities. This implementation added a new class in the CybAssess class. In this implementation, the default value for network score is 10^{-4} and the fixed number of cycles is 100. To test the method, four classes of performance were simulated from seven different Gaussian distributions, which were mixture. From that procedure, several kinds of statistical distributions could be generated. Only one restriction was used: at least 30% of data generated by a statistical distribution should be intersection with another one. This makes a more realistic simulation. For each of four classes were generated 30000 vectors, from which 10000 were used to calibrate the system and the other 20000 were used to check the assessment method based on General Bayesian Network. Each vector contained 10000 positions, which is equivalent to a 10 seconds simulation with a haptics device frequency set at its maximum (1000 data/second) in a real situation (see Section II).

A comparison of the classification agreement between assessment method based on General Bayesian Network and the generated data was performed using the Cohen's Kappa Coefficient, according recommended by literature [17] because it is known to be over conservative. The classification matrix obtained is presented in the Table I. The Kappa coefficient was K=99.98%. In 14 cases, the evaluation tool made mistakes and all of them were made in classes of performance C1 and C2. That performance is good and shows that General Bayesian Network is a competitive approach in the solution of assessment problems.

TABLE I. Assessment results.

Real Class of Performance	Class of Performance predicted by Assessment Method based on General Bayesian Network			
	C1	***C2***	***C3***	***C4***
C1	19994	6	0	0
C2	8	19992	0	0
C3	0	0	20000	0
C4	0	0	0	20000

Those tests were performed in a Dual Core computer with 256Mb graphic card and 2 Gb of RAM.. That computer run Fedora Linux 5 operational system and CyberMed version 1.2. The average of CPU time consumed for assessment of a sample vector contained 10000 positions was 0.5 ms, i.e. that system is able to answer in lower than one second, that characterize an online assessment.

VI. Conclusions

The assessment method based on General Bayesian Network was implemented in CyberMed and is available in the last version of frameworks, available for download at Internet. The results of tests using a statistical simulation showed that assessment methodology is a competitive approach in the solution of assessment problems, with high accuracy and low computational complexity.

Since the methodology is available in the framework, it can be used to assess the bone marrow procedure and any other developed with this framework. The data used in the assessment should be provided by one or mode physicians. They must used the simulator to label procedures and insert in the system what is correct and also what is not correct (Figure 2). The several executions and the variability in them will be used by the assessment tool to define the assessment model. This model will be present in the final version of the simulator. Thus, the several users interactions will be used to compose a model also. This model will be compared to the assessment model to classify this user and provide a feedback about his performance.

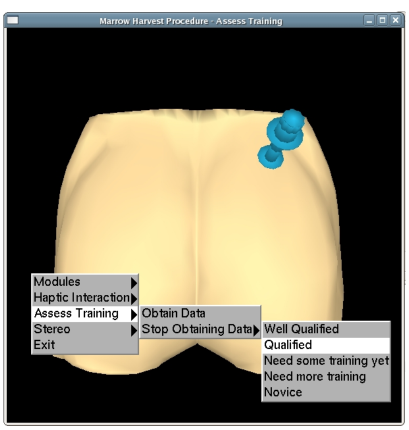

Figure 2. Example of procedure used by physicians to label procedures for the assessment model.

Acknowledgment

This work is partially supported by grants 312375/2009-3 and 310339/2009-0 of the Brazilian Council for Scientific and Technological Development, CNPq. This work is related to the National Institute of Science and Technology for Medicine Assisted by Scientific Computing also supported by CNPq.

References

[1] Cheng, J; Greiner, R.; Learning Bayesian Belief Network Classifiers: Algorithms and System, Proceedings of the Fourteenth Canadian Conference on Artificial Intelligence, 2001.

[2] Gaba D.M. et al. “Assessment of clinical performance during simulated crises using both technical and behavioral ratings”. Anesthesiology, v.89, 1998, pp.8-18.

[3] Johnson, R. A., Wichern, D. W., Applied Multivariate Statistical Analysis, Prentice Hall, 5 edition, 2001.

[4] Krause, P. J., Learning Probabilistic Networks, Knowledge Engineering Review, vol. 13, pp. 321-351, 1998.

[5] Lammers, L. et al. (2008) Teaching and Assessing Procedural Skills Using Simulation: Metrics and Methodology. Academic Emergency Medicine, 15(11): 1079-1087. Wiley.

[6] Machado, L. S.; et al.; A Framework for Development of Virtual Reality-Based Training Simulators. Studies in Health Technology and Informatics, vol. 142, 2009, pp. 174-176.

[7] Machado, L. Moraes, R. “Intelligent Decision Making in Training Based on Virtual Reality”. In: Da Ruan. (Org.). Computational Intelligence in Complex Decision Systems. Amsterdam, Paris: Atlantis Press, 2010, v. , p. 85-123.

[8] Moraes, R. M.; Machado, L. S. “Fuzzy Gaussian Mixture Models for On-line Training Evaluation in Virtual Reality Simulators”. Annals of the International Conference on Fuzzy Information Processing (FIP'2003). March, Beijing, Vol. 2. 2003, pp. 733-740.

[9] Moraes, R.M.; Machado, LS., A new approach for online training assessment for bone marrow harvest when patients have bones deteriored by disease. 2011 Safety, Health and Environmental World Congress, pp.41-45, 2011.

[10] Moraes, R.M.; Machado, L.S., Development of a Medical Training System with Integration of Users' Skills Assessment. Book Chapter. Virtual Reality. Viena: Intech, p. 325-348.

[11] Moraes, R.M.; Machado, L.S., Online Assessment in Medical Simulators based on Virtual Reality Using Fuzzy Gaussian Naive Bayes. Journal of Multiple-Valued Logic and Soft Computing. in press.

[12] Moraes, R.M.; Machado, L.S.; Souza, L.C., Online Assessment of Training in Virtual Reality Simulators Based on General Bayesian Networks. VI International Conference on Engineering and Computer Education, 2009 [CDROM].

[13] Moraes, R.M.; Rocha, A.V.; Machado, L.S., Intelligent Assessment Based on Beta Regression for Realistic Training in Medical Simulators. Journal of Multiple-Valued Logic and Soft Computing. in press.

[14] Neapolitan, R. E. Learning Bayesian Networks, Prentice Hall Series in Artificial Intelligence, 2003.

[15] Scott et al. “Measuring Operative Performance after Laparoscopic Skills Training: Edited Videotape versus Direct Observation”. Journal of Laparoendoscopic & Advanced Surgical Techniques, v.10, n.4, 2000, p.183-190.

[16] Machado, L.S. Et al “A Virtual Reality Simulator for Bone Marrow Harvest for Transplant”. Studies in Health Technology and Informatics, Amsterdam/The Netherlands, v. 81, p. 293-297, 2001.

[17] Duda, R.; Hart, P.; Stork, D. “Pattern Classification”. *Wiley-Interscience*. 2nd ed., 2000.