

ONLINE TRAINING ASSESSMENT IN VIRTUAL REALITY SIMULATORS BASED ON FUZZY NAIVE BAYES

Ronei Marcos de Moraes¹, Liliane dos Santos Machado²

Abstract — Training systems based on virtual reality are used in several areas. In these systems the user is immersed into a virtual world to have realistic training through realistic interactions. In such training is important to know the quality of user's training. An online assessment system allows the user to improve his learning because it can identify, immediately after the training, where he committed mistakes or presented low efficiency. Several approaches to perform online or offline evaluation in training simulators based on virtual reality have been proposed. In this paper, we present a new approach to online training assessment based on Fuzzy Naive Bayes for modeling and classification of simulation in N pre-defined classes. Fuzzy Naive Bayes is a generalization of Naive Bayes Networks, which are a special case of probabilistic networks that allows to treat continuous variables using fuzzy linguistic values.

Index Terms — Assessment, Fuzzy Naive Bayes, Online Training Assessment, Training Based on Virtual Reality.

INTRODUCTION

Training systems based on virtual reality (VR) have been used in several areas [3]. The user is immersed into a virtual world to have realistic training and realistic interactions. However, it is important assess user's training to know the quality of his skills. Several kinds of training based on VR use to record the user actions in videotapes to post-analysis by experts [2]. In these cases, the user receives his assessment after some time. This is a problem because probably after some hours the user will not remember his exact actions what will make difficult the use of the assessment information to improve his performance. Besides of that, several kinds of training cannot be simply classified as bad or good due to its complexity. Then, the existence of an online assessment tool incorporated into a simulation system based on virtual reality is important to allow the learning improvement and users assessment [8]. An online assessment system allows the user to improve his learning because it can identify, immediately after the training, where mistakes occurred or actions presented low efficiency.

Just a few years ago were proposed the first methodologies for training assessment. Specific assessment methodologies for training through virtual reality simulators

are still more recent. Because VR simulators are real-time systems, an assessment tool must continuously monitor all user interactions and compare his performance with pre-defined expert's classes of performance. By didactic reasons, it is more interesting the use of online assessment tools due the fact that these methods allows the user to easily remember his mistakes and learn how to correct them.

In medicine, some models for online assessment of training have been proposed [8,10,14,16,17,18]. The main problems related to online training assessment methodologies applied to VR systems are the computational complexity and the accuracy. An online assessment tool must have low complexity to do not compromise VR simulations performance, but it also must have high accuracy to do not compromise the user assessment.

Some of the models previously mentioned are based on machine learning and use discretization of continuous variables, as proposed in [18]. They use an assessment system based on Naive Bayes method. However, this approach can cause loss of the information. In this paper, it is proposed a new assessment tool based on Fuzzy Naive Bayes. This approach can lead with continuous variables and can also assist the other requirements of an evaluation system for training based on VR.

VIRTUAL REALITY AND SIMULATED TRAINING

Virtual Reality refers to real-time systems modeled by computer graphics that allow user interaction and movements with three or more degrees of freedom [3,29]. More than a technology, VR became a new science that joins several fields as computers, robotics, graphics, engineering and cognition. VR worlds are 3D environments, created by computer graphics techniques, where one or more users are immersed totally or partially to interact with virtual elements. The realism of a virtual reality application is given by the graphics resolution and by the exploration of users senses. Mainly, special devices stimulate the sight, hearing and touch. As example, head-mounted displays (HMD) or even conventional monitors combined with special glasses can provide stereoscopic visualization, multiple sound sources positioned provides 3D sound, and touch can be simulated by the use of haptic devices [11,25].

Virtual reality systems for training can provide significant benefits over other methods of training, mainly in

¹Ronei Marcos de Moraes, Department of Statistics, Federal University of Paraíba. Cidade Universitária s/n, Castelo Branco, João Pessoa-PB. 58.051-900. Brazil. ronei@de.ufpb.br

²Liliane dos Santos Machado, Department of Informatics, Federal University of Paraíba. Cidade Universitária s/n, Castelo Branco, João Pessoa-PB, Brazil. 58.051-900. liliane@di.ufpb.br

critical medical procedures. In some cases, those procedures are performed without any kind of visualization and the only information received is noticed by the touch sensations provided by a robotic device with force feedback. These devices can measure forces and torque applied during the user interaction [25] and these data can be used in an assessment [8,23]. A specific kind of haptic device, as the presented in Figure 1, is based on a robotic arm and provides force feedback and tactile sensations during user manipulation of objects in a three dimensional scene. This way, user can feel objects texture, density, elasticity and consistency. Since the objects have physical properties, a user can identify them in a 3D scene (without see them) by the use of this kind of device [11]. This is especially interesting in medical applications to simulate proceedings in which visual information is not available. One of the main reasons for the use of such haptic devices in medical applications is their manipulation similarity when compared to real surgical tools.



FIGURE. 1
HAPTIC DEVICE USED IN VR SYSTEMS.

ASSESSMENT IN VIRTUAL REALITY SIMULATORS

The assessment of simulations is necessary to monitor the training quality and provide some feedback about the user performance. User movements, as spatial movements, can be collected from mouse, keyboard and any other tracking device. Applied forces, angles, position and torque can be collected from haptic devices [25]. Then, virtual reality systems can use one or more variables, as the mentioned above, to assess a simulation performed by user.

Some simulators for training present a method of assessment. However they just compare the final result with an expected result or are post-analyses of videotape records [2]. Recently, some models for offline or online assessment of training have been proposed, some of them use Discrete Hidden Markov Models [23] or Continuous Hidden Markov Models [24] to modeling forces and torque during a simulated training in a porcine model. Machado et al. [8] proposed the use of a fuzzy rule-based system to online assessment of training in virtual worlds. Using an optoelectronic motion analysis and video records, McBeth et al. [12] acquired and compared postural and movement data of experts and residents in different contexts by use of

distributions statistics. Moraes and Machado proposed several methods for online assessment [10,15,16,18,19]. They also proposed a methodology to automatically assess a user's progress to improve his/her performance in virtual reality training systems [17] using statistical measures and models (time dependent or not) as well as a fuzzy expert system. After that, Morris et al. [20] suggested the use of statistical linear regression to evaluate user's progress in a bone surgery.

In this paper, we propose a new system for assessment based on Fuzzy Naive Bayes classifier. This system can perform an online training assessment for virtual reality simulators. The system uses a vector of information, with data collected from user interactions with virtual reality simulator, and these data are compared by the assessment system with M pre-defined classes of performance. To test the method proposed, we are using a bone marrow harvest simulator [9]. This simulator has as goal to training new doctors to execute the bone marrow harvest, one of the stages of the bone marrow transplant. The procedure is done blindly, performed without any visual feedback, except the external view of the patient body, and the physician needs to feel the skin and bone layers trespassed by the needle to find the bone marrow and then start the material aspiration (Figure 2). The simulator uses a robotic arm that operates with six degrees of freedom movements and provides force feedback to give to the user the tactile sensations felt during the penetration of the patient's body [9]. In the system the robotic arm simulates the needle used in the real procedure, and the virtual body visually represented has the tactile properties of the real tissues. The assessment tool proposed supervised the user movements during the puncture and evaluated the training according to M possible classes of performance.

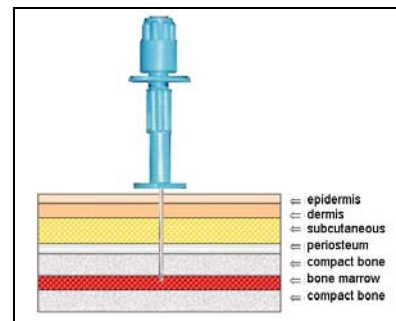


FIGURE. 2
THE TISSUE LAYERS TRESPASSED BY NEEDLE IN A BONE MARROW HARVEST.

ASSESSMENT TOOL BASED ON NAIVE BAYES

This section presents the method for training assessment, based on Fuzzy Naive Bayes. For reader's better understanding, we first present a short review about Naive Bayes method [18]. After that, we present the Fuzzy Naive Bayes method.

Naive Bayes Method

Formally, let be the classes of performance in space of decision $\Omega = \{1, \dots, M\}$ where M is the total number of classes of performance. Let be $w_i, i \in \Omega$ the class of performance for an user. A Naive Bayes classifier computes conditional class probabilities and then predict the most probable class of a vector of training data X , according to sample data D , where X is a vector with n features obtained when a training is performed, i.e. $X = \{X_1, X_2, \dots, X_n\}$. Using the Bayes Theorem:

$$\begin{aligned} P(w_i | X) &= [P(X | w_i) P(w_i)] / P(X) \Leftrightarrow \\ &\Leftrightarrow P(w_i | X_1, X_2, \dots, X_n) = \\ &= [P(X_1, X_2, \dots, X_n | w_i) P(w_i)] / P(X) \end{aligned} \quad (1)$$

However, as $P(X)$ is the same for all classes w_i , then it is not relevant for data classification and can be rewritten as:

$$P(X | w_i) P(w_i) = P(X_1, X_2, \dots, X_n | w_i) P(w_i) \quad (2)$$

The equation (2) is equivalent to the joint probability model:

$$P(X_1, X_2, \dots, X_n | w_i) P(w_i) = P(X_1, X_2, \dots, X_n, w_i) \quad (3)$$

Now, using successive applications of the conditional probability definition over equation (3), can be obtained:

$$\begin{aligned} P(X_1, X_2, \dots, X_n, w_i) &= P(w_i) P(X_1, X_2, \dots, X_n | w_i) = \\ &= P(w_i) P(X_1 | w_i) P(X_2, \dots, X_n | w_i, X_1) \\ &= P(w_i) P(X_1 | w_i) P(X_2 | w_i, X_1) P(X_3, \dots, X_n | w_i, X_1, X_2) \\ &\dots \\ &= P(w_i) P(X_1 | w_i) P(X_2 | w_i, X_1) \dots P(X_n | w_i, X_1, X_2, \dots, X_{n-1}) \end{aligned}$$

The Naive Bayes classifier receives this name because its naive assumption of each feature X_k is conditionally independent of every other feature X_l , for all $k \neq l \leq n$. It means that knowing the class is enough to determine the probability of a value X_k . This assumption simplifies the equation above, due to:

$$P(X_k | w_i, X_l) = P(X_k | w_i) \quad (4)$$

for each X_k and the equation (3) can be rewritten as:

$$\begin{aligned} P(X_1, X_2, \dots, X_n, w_i) &= \\ &= P(w_i) P(X_1 | w_i) P(X_2 | w_i) \dots P(X_n | w_i) \end{aligned} \quad (5)$$

unless a scale factor S , which depends on X_1, X_2, \dots, X_n . Finally, equation (1) can be expressed by:

$$P(w_i | X_1, X_2, \dots, X_n) = (1/S) P(w_i) \prod_{k=1}^n P(X_k | w_i) \quad (6)$$

Then, the classification rule for Naive Bayes is done by:

$$\begin{aligned} X \in w_i \text{ if } P(w_i | X_1, X_2, \dots, X_n) &> P(w_j | X_1, X_2, \dots, X_n) \\ \text{for all } i \neq j \text{ and } i, j \in \Omega \end{aligned} \quad (7)$$

and $P(w^* | X_1, X_2, \dots, X_n)$ with $* = \{i, j | i, j \in \Omega\}$, is done by (6).

To estimate parameters for $P(X_k | w_i)$ for each class i , it was used a maximum likelihood estimator, named P_e :

$$P_e(X_k | w_i) = \#(X_k, w_i) / \#(w_i) \quad (8)$$

where $\#(X_k, w_i)$ is the number of sample cases belonging to class w_i in all sample data D and having the value X_k , $\#(w_i)$ is the number of sample cases that belong to the class w_i in all sample data D .

Fuzzy Naive Bayes Method

Several authors proposed versions of Fuzzy Naive Bayes method. Tang et al. [26] proposed a Fuzzy Naive Bayes method with two stages: in the first one, a fuzzy clustering algorithm determine partitions in space of decision; in the second stage, the partitions obtained in the first stage are used to estimate the parameters for linguistic variables. With that methodology is possible to use continuous variables and to decrease the learning complexity of the Naive Bayes method. In another paper, Tang et al. [27] analysed the model identification using weighted fuzzy production rules. After that, the accuracy of fuzzy production rules was investigated using genetic algorithms [28].

Other approaches were used for fuzzy models of Naive Bayes method. Borgelt [4] extended the Naive Bayes method to manipulate some kinds of fuzzy information. Nurnberger [22] made mappings from Naive Bayes method to a NEFCLASS modified algorithm [21] to improve the first one. Borgelt [5] used Fuzzy Maximum Likelihood Estimation [6] to determine fuzzy partitions.

There is a third approach, in which fuzzy discretization methods [7,30] were used in the first stage to allow the use of the Naive Bayes method after. However, this approach can affect classification bias and variance of the Naive Bayes method.

In spite of the qualities of all approaches found in the literature, the probability distributions of random variables from several kinds of applications, do not follow well defined statistical patterns [13,14]. This way, the fuzzy approach by linguistic variables can improve some assessment system previously proposed and allowing use of continuous variables. For those reasons, the approach proposed by [26,28] is followed in this paper.

Based on the same space of decision with M classes, a Fuzzy Naive Bayes method computes conditional class probabilities and then predicts the most probable class of a vector of training data $X = \{X_1, X_2, \dots, X_n\}$, according to sample data D . In this case, it is assumed that each $X_k, k=1, \dots, n$, is a linguistic variable expressed by linguistic

values. The parameters of Fuzzy Naive Bayes method are learning from data using two main steps. In the first one, an unsupervised fuzzy clustering method estimates partitions in feature space. After that, conditional probabilities are estimated for Fuzzy Naive Bayes method.

The first step in that methodology is to obtain fuzzy partition from data. In particular, it is necessary to estimate the centroids of possible clusters in training data. According [26], the number of centroids is independent of the number of classes. Due to the uncertainty of features measures, it is appropriate to use fuzzy clustering. The Fuzzy c-means (FCM) algorithm [1] or its modifications can be applied in this case.

The FCM is an interactive algorithm [26] to find out a fuzzy partition matrix U over sample data D and each cluster centroid, denoted by V_i , $i \in \Omega$. Before apply FCM algorithm, some parameters must be defined: the number M of clusters, the α fuzziness parameter and the ϵ tolerance. The results of FCM are the cluster centroids V_i and they are necessary to estimate the membership functions μ_{Xki} of linguistic variables X_k . Several different kinds of membership functions can be defined, as well as different shapes too [26].

To estimate parameters in Naive Bayes method for $P(X_k \setminus w_i)$ for each class w_i , the equation (8) is used. In Fuzzy Naive Bayes method, the first stage obtained m fuzzy partitions for the sample data D , denoted by D_i , $i=1, \dots, m$, and the membership functions μ_{Aki} for each fuzzy partition A_{ki} , in the domain of the variable X_k . The conditional probabilities for X_k for the class w_i , from sample data D can be estimated as:

$$P(X_k = A_{ki} \setminus w_i) = [\sum_{X \in D_i} A_{ki}(x_k)] / [\sum_{i=1}^m \sum_{X \in D_i} A_{ki}(x_k)] \quad (9)$$

where x_k is the k -th component of the vector X , with $X \in D_i$. It is important to note that equation (9) is a generalization of the equation (8).

According to Bayes rule:

$$P(w_i \mid X_1 = A_{1i}, \dots, X_n = A_{ni}) = \prod_{k=1}^n P(X_k \setminus w_i) P(w_i). \quad (10)$$

Then, the classification rule for Fuzzy Naive Bayes is done by:

$$X \in w_i \text{ if: } P(w_i \mid X_1 = A_{1i}, \dots, X_n = A_{ni}) > P(w_j \mid X_1 = A_{1i}, \dots, X_n = A_{ni}) \text{ for all } i \neq j \text{ and } i, j \in \Omega \quad (11)$$

THE ASSESSMENT TOOL

The assessment tool proposed should supervise the user's movements and other parameters associated to them. The system must collect information about positions in the space, forces, torque, resistance, speeds, accelerations, temperatures, visualization position and/or visualization

angle, sounds, smells and etc. The virtual reality simulator and the assessment tool are independent systems, however they act simultaneously. The user's interactions with the simulator are monitored and the information is sent to the assessment tool that analyzes the data and emits a report on the user's performance at the end of the training. Depending on the application, all those variables or some of them will be monitored (according to their relevance to the training).

The virtual reality system used for the tests is a bone marrow harvest simulator [9]. In a first movement on the real procedure, the trainee must feel the skin of the human pelvic area to find the best place to insert the needle used for the harvest. After, he must feel the tissue layers (epidermis, dermis, subcutaneous, periosteum and compact bone) trespassed by the needle and stop at the correct position to do the bone marrow extraction. In our VR simulator the trainee interacts with a robotic arm and his/her movements are monitored in the system by some variables [9]. For reasons of general performance of the VR simulator, were chosen to be monitored the following variables: spatial position, velocities, forces and time on each layer. Previously, the system was calibrated by an expert, according M classes of performance defined by him. 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 was defined as $M=3$: 1) correct procedures, 2) acceptable procedures, 3) badly executed procedures. So, the classes of performance for a trainee could be: "you are well qualified", "you need some training yet", "you need more training".

The information of variability about these procedures is acquired using Fuzzy Naive Bayes method. In our case, we assume that the font of information for w_i classes is the vector of the sample data D . The user makes his/her training in virtual reality simulator and the Assessment Tool based on Fuzzy Naive Bayes collects the data from his/her manipulation. All probabilities of data for each class of performance are calculated by (9) and at the end the user is assigned to a w_i class of performance by (11). So, when a trainee uses the system, his performance is compared with each expert's class of performance and the Assessment Tool based on Fuzzy Naive Bayes assigns the better class, according to the trainee's performance. At the end of the training, the assessment system reports the classification to the trainee.

CONCLUSIONS AND FUTURE WORKS

In this paper we presented a new approach to online training assessment in virtual reality simulators. This approach uses an Assessment Tool based on Fuzzy Naive Bayes and solves the main problems in assessment procedures: use of continuous variables, low complexity and high accuracy. Systems based on this approach can be applied in virtual reality simulators for several areas and can be used to assess

a trainee into classes of learning giving him a status about his performance.

As future work, we intend to test and to make a statistical comparison between others methodologies and the methodology proposed in this paper.

ACKNOWLEDGMENT

This work is partially supported by Brazilian Council for Scientific and Technological Development, CNPq (Process 303444/2006-1 and Process CT-INFO-CNPq 506480/2004-6) and Brazilian Research and Projects Financing, FINEP (Grant 01-04-1054-000).

REFERENCES

- [1] Bezdek, J. C.; *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.
- [2] Burdea, G., Patounakis, G., Popescu, V. and Weiss, R.E. "Virtual Reality Training for the Diagnosis of Prostate Cancer". *Proceedings of IEEE Virtual Reality Annual Int. Symposium*, pp. 190-197, 1998.
- [3] Burdea, G.; Coiffet, P. *Virtual Reality Technology*. 2nd ed., Wiley Interscience, 2003.
- [4] Borgelt, C.; Gebhardt, J.; A Naive Bayes Style Possibilistic Classifier. *Proceedings of 7th European Congress on Intelligent Techniques and Soft Computing (EUFIT'99)*, 1999. [cdrom].
- [5] Borgelt, C.; Timm, H.; Kruse, R.; Using fuzzy clustering to improve naive Bayes classifiers and probabilistic networks. *Proc. Ninth IEEE Int. Conf. on Fuzzy Systems (FUZZ IEEE 2000)*, p. 53-58, 2000.
- [6] Gath, I.; Geva, A.B.; Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.11, n. 7, p. 773-780, 1989.
- [7] Kononenko, I.; Inductive and Bayesian learning in medical diagnosis. *Applied Artificial Intelligence* v.7, n.4, p.317-337, 1993.
- [8] Machado, L. S., Moraes, R. M. and Zuffo, M. K. "Fuzzy Rule-Based Evaluation for a Haptic and Stereo Simulator for Bone Marrow Harvest for Transplant". *Proceedings of 5th Phantom Users Group Workshop*, 2000.
- [9] Machado, L. S., Mello, A. N., Lopes, R. D., Odone Filho, V.; Zuffo, M. K. A Virtual Reality Simulator for Bone Marrow Harvest for Pediatric Transplant. *Studies in Health Technology and Informatics*. Vol. 81, pp.293-297, 2001.
- [10] Machado, L. S.; Moraes, R. M. Online Training Evaluation in Virtual Reality Simulators Using Evolving Fuzzy Neural Networks. *Proceedings of the 6th Int.FLINS Conf.* Belgium., pp.314-317, 2004.
- [11] Mahoney, D.P. The Power of Touch. *Computer Graphics World*. Vol. 20, No. 8. August, pp. 41-48, 1997.
- [12] McBeth, P. B. et al. Quantitative Methodology of Evaluating Surgeon Performance in Laparoscopic Surgery". *Studies in Health Technology and Informatics*, v. 85, pp.280-286, 2002.
- [13] Melo, A. C. O.; Moraes, R. M.; Machado, L. S.; Gaussian Mixture Models for Supervised Classification of Remote Sensing Multispectral Images. *Lecture Notes in Computer Science*, v. 2905, p. 440-447, 2003.
- [14] 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, pp. 733-740, 2003.
- [15] Moraes, R. M.; Machado, L. S.; Using Fuzzy Hidden Markov Models for Online Training Evaluation and Classification in Virtual Reality Simulators. *Int. Journal of General Systems*, v. 33 n.2-3, pp. 281-288, 2004.
- [16] Moraes, R. M.; Machado, L. S. Maximum Likelihood for On-line Evaluation of Training Based on Virtual Reality. *Proceedings of Global Congress on Engineering and Technology Education (GCETE'2005)*. Março, Santos, Brasil. 2005, pp. 299-302, 2005.
- [17] Moraes, R. M.; Machado, L. S.; Continuous Evaluation in Training Systems Based on Virtual Reality. *Proc. Global Congress on Engineering and Technology Education (GCETE'2005)*. March, Santos, pp.1048-1051, 2005.
- [18] Moraes, R. M.; Machado, L. S.; Assessment Based on Naive Bayes for Training Based on Virtual Reality. *Proc. Int. Conf. on Engineering and Computer Education (ICECE'2007)*. March, Santos, pp. 269-273, 2007.
- [19] Moraes, R. M.; Machado, L. S.; Fuzzy Bayes Rule for On-Line Training Assessment in Virtual Reality Simulators. *Journal of Multiple-Valued Logic and Soft Computing* (to appear).
- [20] Morris, D.; Sewell, C.; Barbagli, F. et al; Visuohaptic simulation of bone surgery for training and evaluation. *IEEE Computer Graphics and Applications*, v.26 n. 6, 2006, pp. 48-57.
- [21] Nauck, D.; Kruse, R.; NEFCLASS – A Neuro-fuzzy Approach for the Classification of Data. *Proceedings of ACM Symposium on Applied Computing*, p. 461-465, 1995.
- [22] Nurnberger, A.; Borgelt, C.; Klose, A.; Improving naive Bayes classifiers using neuro-fuzzy learning. *Proceedings of 6th International Conference on Neural Information Processing (ICONIP '99)*, p. 154-159, 1999.
- [23] Rosen J.; Richards, C.; Hannaford, B.; Sinanan, M. Hidden Markov Models of Minimally Invasive Surgery, *Studies in Health Tech. and Informatics: Medicine Meets Virtual Reality*, v. 70, pp. 279-285, 2000.
- [24] Rosen J., Solazzo, M.; C., Hannaford, B.; Sinanan, M.; Objective Laparoscopic Skills Assessments of Surgical Residents Using Hidden Markov Models Based on Haptic Information and Tool/Tissue Interactions. *Studies in Health Technology and Informatics: Medicine Meets Virtual Reality*, v. 81, pp. 417-423. 2001.
- [25] Salisbury, K. Haptics: The Technology of Touch. *HPCWire Special*. n. 10. Nov 1995.
- [26] Tang, Y.; Pan, W.; Li, H.; Xu, Y.; Fuzzy Naive Bayes classifier based on fuzzy clustering. *Proceedings of 2002 IEEE International Conference on System, Man and Cybernetics*. October, 2002.
- [27] Tang, Y.; Pan, W.; Qiu, X.; Xu, Y.; The identification of fuzzy weighted classification system incorporated with Fuzzy Naive Bayes from data. *Proceedings of 2002 IEEE International Conference on System, Man and Cybernetics*. October, 2002.
- [28] Tang, Y.; Xu, Y.; Application of fuzzy Naive Bayes and a real-valued genetic algorithm in identification of fuzzy model. *Information Sciences*, v.169, p.205-225, 2005.
- [29] Vince, J. *Virtual Reality Systems*. Addison-Wesley, 1995.
- [30] Yang, Y.; Webb, G. I.; A Comparative Study of Discretization Methods for Naive-Bayes Classifiers. *Proceedings of 2002 Pacific Rim Knowledge Acquisition Workshop (PKAW'02)*, p. 159-173, 2002.