A Virtual Environment Using Virtual Reality and Artificial Neural Network

Abdul Rahaman Wahab Sait

Lecturer, Dept. of Computer Science
Alquwaya Community College, Shaqra University
Alquwaya, Kingdom of Saudi Arabia

Mohammad Nazim Raza

Lecturer, Dept. of Computer Science
Alquwaya College of Science, Shaqra University
Alquwaya, Kingdom of Saudi Arabia

Abstract—In this paper we describe a model, which gives a virtual environment to a group of people who uses it. The model is integrated with an Immersible Virtual Reality (IVR) design with an Artificial Neural Network (ANN) interface which runs on internet. A user who wants to participate in the virtual environment should have the hybrid IVR and ANN model with internet connection. IVR is the advanced technology used in the model to give an experience to the people to feel a virtual environment as a real one and ANN used to give a shape for the characters in the virtual environment (VE). This model actually gives an illusion to the user that as if they are in the real communication environment.

Keywords- component; Model; Virtual environment; Immersible virtual reality; Internet ; Artificial neural networks.

I. INTRODUCTION

The aim of this work is to develop a hybrid IVR and ANN model which will create a VE where a group of people chat together and at the same time they will feel as if they are in a real environment. We have used an IVR and ANN interface in this model to take the people into the VE. The model uses internet as a medium to connect the people at various regions. ANN plays a vital role here to give the presence of the image more real and make people be immersed into the environment [8][16]. Precisely the model will make a dream world to the person who uses it.

II. RELATED WORK

Manuel oliveria et al[2]. analyses a model for shared VE. They analyzed the technology uses for shared environment. They presented a model for the users to participate in a shared online game and social 3D environment. They have used virtual reality markup language to create the interface. In the environment the user will be represented as an avatar, which is an object in the virtual world. The user will be partially immersed in the environment. They have analyzed the issues regarding the VR architecture and did not give much detail about the design. We have taken ANN to synthesize the image and voice in the VE.

Dominic W.Massaro et al [3]. presents an idea on a system that synthesizes visual speech directly from the acoustic waveform. They have trained ANN to map the cepstral co-efficients of an individual's natural speech to the control parameters of an animated synthetic talking head. With this technique, the animated talking head is generated from and aligned with the original speech of the talker.

Pengyu Hong et al [4]. gave an advanced approach on a real time speech – driven synthetic talking face provides an effective multimodal communication interface in distributed collaboration environments. They have introduced an algorithm for motion unit based facial motion tracking system. The algorithm achieves more robust results by using some high level knowledge models. They have used vector quantization, which is a classification based audio to visual conversion approach. They have trained ANN with Audio – visual training data for the system to produce the audio according the face animation. They have taken the use of the audio – visual database to obtain the visual information.

Yigiang Chen et al [5]. presents a way using the combination of clustering and machine learning methods to learn the correspondence between speech acoustic and MPEG-4 based face animation parameters. They have trained ANN to map the linear predictive co-efficients and some features of an individual's natural speech to face animation parameter. They have calculated linear predictive co-efficients and some parameters to obtain a useful vocal representation.

George votsis et al [1]. analysed the method for the recognition of user's emotional state of a PC user. They have conducted some experiment with the audio – visual databases with the ANN.

In this paper we describe a chosen VE for the user to communicate with other user and at the same time they will feel the environment as a real one by using the model. To make the environment more real we have taken ANN as a processing tool to present the user expressions in the VE.

III. PROPOSED MODEL OVERVIEW

A. Immersible Virtual Reality Design

IVR is the future technology which has the ability to make people to be immersed in the artificial environment. It acts as a vehicle for the user to move into their dream world.

As our requirements analysis proved the necessity to provide a system design for realistic shape visualization. It was obvious to use IVR techniques for state of the art 3D graphics display [2][15][16]. As IVR technology has reached a very mature level, it is very well suited for real industrial

applications which require a high level of robustness. The sense of touch can greatly enhance the users' sense of occurrence within VE which may effect in increased task performance. A successful approach has been passive haptics, which consists of providing physical artifacts for the users to physically touch whilst seeing their virtual representation. To benefit fully from the advantages of touch and keep the general purpose of VEs it is necessary to use active haptics or force feedback haptics[2].

In order to improve the level of perception our requirements analysis identified the additional need to utilize haptic force feedback. This gives intuitive and accurate interaction control to the system.

B. Artificial Neural Networks Design

An ANN is an information processing model that is similar to the way the biological nervous systems process information. It has the ability to formulate meaning from complicated or uncertain data. It exhibits the ability of adaptive learning [10]. It process information in a similar way the human brain does. It is composed of a large number of highly interconnected processing elements called neurons, working in parallel to solve a specific problem [11].

The fundamental form of ANN consists of three layers. A layer of input unit is connected to an intermediate layer of hidden units, which is connected to a layer of output unit. The behavior of the output unit depends on the activity of the hidden and input unit [9].

ANN architecture for our design is implemented with back propagation algorithm. Through back propagated networks, learning occurs during the training phase in which each input sample in the training set is applied to the input units and propagated forward[13]. The hidden layer in the middle is the heart of the process. Increase in the hidden layer gives more accurate result. The output unit will do certain matching work with the hidden layer to give the result. After a back propagated network has learned the correct classification for a set of inputs, it can be tested on a second set of inputs to see how well it classifies untrained samples. In our model, we have used this architecture to simulate the image and express their emotion of the user in the VE [14]. Users' voice, video and images are taken as input to the train the ANN [9][10[11][12]. Later stage

those data can be used for the security purpose to check the users' identity. In Fig 1 we have shown the simple IVR and ANN model.

C. Internet and Database

Internet is used as a pathway for the modern software to connect the people from all over the world together at one point of contact on the web. We require a high speed internet connection to run the model without any problem for the end user. We have to train ANN with the users' sample on the internet. As the interface runs on the internet it requires more efficiency to carry multimedia data without having any problem [13]. Another concern is the scalability of the system for a large number of users and a large amount of multimedia files. Ideally the multicast model is suited for group communication by using the same logical address.

Database is the heart of all software. As our requirement investigation we require the database with prosodic feature. The database should have the capability to contain the high quality video samples. Ideally it should have only genuine expressions of emotions. A relatively large number of facial expression recognition databases are widely employed to store and retrieve the audio and visual data [6][7]. These kinds of databases are very much useful to interact with ANN. In Fig 2 we have shown the process overview of the proposed model.

Security is one of the most concerning issues in any system [2]. The most prominent security mechanism is a user validation process during a log – on phase. In our model all users those who are going to share the environment should record their voice, video and images with the use of interface. By taking those data, the interface will check their identity in the future. So there is a less chance of security breach in this model.

IV. MODEL - WORKING METHODOLOGY

We have designed a prototype which will be used to take the people into the VE, ANN is used as a processing tool to give the virtual presence of the people by taking inputs like their image and video. Internet used as a medium to connect the IVR model. In the Fig 3, we gave a pictorial representation of functioning of the model.
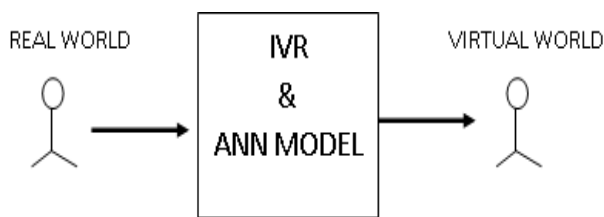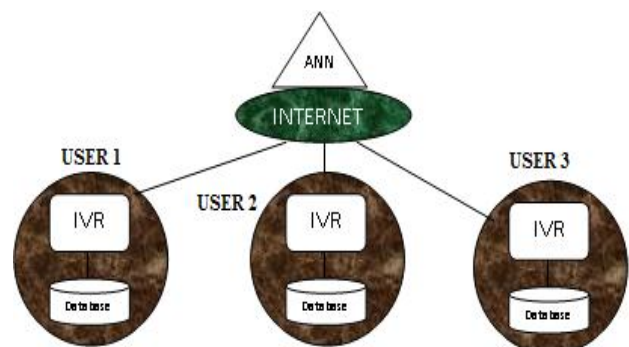

Figure 1- IVR and ANN Design

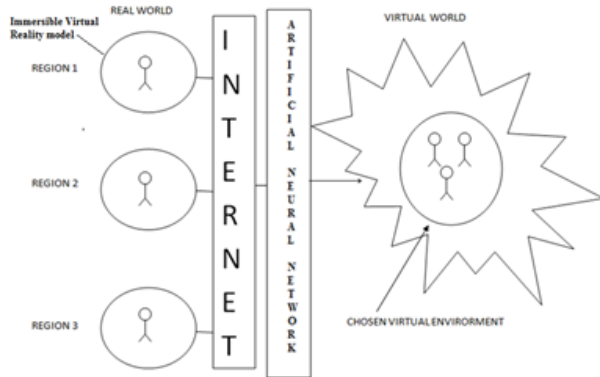Figure 2: Process overview of the model

Figure 3 - Pictorial representation of functioning of the model

- Users are the participants going to share the environment. They are the manager of the model.

- Internet is the medium which connects different users together with the use of the interface. Internet speed is one of the catalysts for the model.

- ANN is the back processing tool of the model for the simulation of the user's presence in the network.

We need IVR design and ANN implemented interface and internet connection to run the model. The model has two stages, first stage is for the interface, and user has to key in details like their name, age and the region into the model. After that they have to upload their image, voice and video. These data will be stored in the audio – visual database. Later stage user data will be given to the interface to configure the user with the model. Next time when the user enters into the system

Figure 4: Illustration - 1 of the Interface

the interface will check their identity with the data already stored into it. The ANN architecture has to be trained with the samples during the training phase [13]. Here we gave illustrations of the interface in Fig 4 and Fig 5.

Figure 5: Illustration - 2 of the Interface

The second or final stage is a VE, the IVR will give the chosen environment to the users those who have participated in the network and ANN will give simulated images of the users in the VE. The interface has to be designed with some environment with the use of 3D-graphics [2][9][16]. When users select the environment the IVR design will create the environment for them. Internet is the back bone for this model. Internet speed will also important criteria for the model. ANN has the capability to recover from the failure.

V. FUTURE SCOPE AND CONCLUSION

We have presented a prototype of future technology and its methodology. We will discuss the other side of the model like internet and various design issues in the future. This hybrid VR and ANN model gives a VE in which users can meet and chat together. Further research is being conducted in order to develop a user friendly and plausible model for the user. This is the direction that we are following in our on- going research.

REFERENCES

[1] George votsis, Nikolas D. Doulamis, Anastasios D.Doulamis, Nicolas Tsapatoulis and Stefanos D. Kollias, "A Neural – Network – Based Approach to Adaptive Human Computer Interface", ICANN 2001, LNCS 2130, PP 1054 – 1059, 2001.

[2] Manuel oliveria, Joel Jordan, Joao Pereira, Joaqium Jorge and Anthony steed, "Analysis Domain Model for Shared Virtual Environments", The International Journal of Virtual Reality, 2009,Vol. 8, No. 4, PP 1 -30.

[3] Dominic W.Massaro, Jonas Beskow, Michael M.Cohen, Christopher L.fry and Tony Rodriguez,"Picture my voice: Audio to Visual Speech Synthesis Using Artificial Neural Networks", Proceeding of AVSP'99,International Conference on Auditory – Visual Speech Processing, PP 133 – 138.

[4] Pengyu Hong, Zhen Wen and Thomas S.Huang,"Real – time Speech Driven Face Animation with Expressions using Neural Network", IEEE Transactions on Neural Networks, Vol.13, Issue.4,Jul 2002,PP 916 – 927.

[5] Yigiang Chen, Wen Gao, Zhaogi Wang and Lizao," Speech Driven MPEG -4 Based Face Animation via Neural Network", PCM '01 Proceedings of the second IEEE pacifici Rim Conference on Multimedia: Advances in multimedia information processing,PP 1108 – 1113.

[6] Patterson E.K., Gurburz S, Tufekci Z and Gowdy J.N., "CUAVE: A New Audio – Visual Database for Multimodal Human – Computer Research",IEEE Conference on Acoustics speech and signal processing (ICASSP) 2002, PP: II 2017 – II 2020.

[7] T.Kanade, J.F.Cohn and Y.L.Tian, " Comprehensive Database for Facial Expression Analysis", Proceeding 4th IEEE International Conference on Automatic Face & Gesture Recognition (FG '00), PP 46 – 53, 2000.

[8] Caudell, T.P. "Application of Neural Networks to Virtual Reality:, 3rd IEEE International workshop on Robot and Human Communication,1994, 18 – 20 Jul 94, PP 26 – 30.

[9] John Weissmann and Ralf salomon," Gesture Recognition for Virtual Reality Applications using Data gloves and Neural Networks", International Joint Conference on Neural Networks (1999), Vol. 1, Publisher:IEEE, PP 2043 – 2046.

[10] M.R.Everigham, B.T. Thomas and T.Troscianko, "Head – Mounted Mobility Aid for Low Vision using Scene Classification Techniques", The International Journal of Virtual Reality (1998), Vol. 3, No.4, PP 1 – 10.

[11] Ifat – Al – Baqee, A.S.M. Mohsin, Kurratul Ain, Mohammad Ashraf, Hossain Sadi and Md.Rakibul Hasan, "Developing a Neural Network – based Method of Faster Face Recognition by Training and Simulation", The International Journal of Engineering Science, Vol. 2, Iss. 11, PP 6694 – 6703.

[12] Steve Lawrence, C.Lee Giles, Ah Chung Tsoi and Andrew D.Back " Face Recognition: A convolutional Neural Network Approach",IEEE Transactions on Neural Network, Special Issue on Neural Network and Pattern Recognition, PP 1 – 24.

[13] Hagan M.T. and M.Menhaj, 1999,"Training Feed – forward Networks with the Marquardt Algorithm", IEEE Transactions on Neural Networks, Vol.5, No.6,1999, PP 989 – 993.

[14] Lu Ye, "HRTF and Neural Network Based Prediction and Simulation Method for Indoor Sports Acoustic", International Conference on Internet Technology and Applications,2010,20 – 22 Aug 2010, PP 1 – 4.

[15] Xiubo Liang, Zhen Wang, Weidong Geng and Franck Multon," A Motion – based User Interface for the control of Virtual Humans Performing Sports", The International Journal of Virtual Reality, Vol. 10 No.3, PP 1 – 8.

[16] Yang – Wai chow, " A Cost – Effective 3D interaction Approach for Immersive Virtual Reality", The International Journal of Recent Trends in Engineering, Vol. 1, No. 1, May 2009, PP 529 – 531.

AUTHORS PROFILE

ABDUL RAHAMAN WAHAB SAIT was born in Apr. 19, 1981 and he has completed his Masters in Information technology in 2003 in Madras University, India. Later he has done his Master of philosophy in computer science in 2007 at Periyar university, India. Now he is working as a Lecturer in Computer science, Alquwaya, Shaqra University,Kingdom of Saudi Arabia. He got interest in Virtual reality and Artificial Neural Networks. He has written many computer articles and presented paper in National conference in India.

MOHAMMAD NAZIM RAZA was born in jun. 20, 1983 and he has done his Master in Computer Application during the year 2007. Currently he is in Saudi Arabia and working as a Lecturer in Computer Science, Alquwaya, Shaqra University, Kingdom of Saudi Arabia. His interests are in Computer networks, Virtual reality and Artificial Neural networks.