View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Infoscience - École polytechnique fédérale de Lausanne

Virtual Humans in Virtual Environments: A New View of Multimedia Applications*

Daniel Thalmann

Computer Graphics Lab (LIG), Swiss Federal Institute of Technology (EPFL)
CH-1015 Lausanne, Switzerland

Telepresence is the future of multimedia systems and will allow participants to share professional and private experiences, meetings, games, parties. The sense of "presence" in the virtual environment is an important requirement for collaborative activities involving multiple remote users working with social interactions. Using virtual humans within the shared environment is a essential supporting tool for presence. We will need in the future real-time realistic 3D avatars, but also interactive perceptive actors to populate the Virtual Worlds.

1 From Multimedia to Telepresence

Traditional multimedia systems are systems that handle different forms of data, such as texts, audio, or video. Until recently, the Web has represented a typical traditional multimedia system with HTML texts, gif images and Quicktime or MPEG movies, then came VRML the Virtual Reality Modelling Language: a way of creating 3D scenes allowing the Web users to walk through 3D spaces. VRML supports the integration of Virtual Reality with the World Wide Web with the goal of broadening access to VR environments via the WWW infrastructure. More generally, Virtual Environments define a new interface for networked multimedia applications. Users will be able to move in Virtual Spaces where they can find many virtual objects including virtual humans. These virtual objects should be multimedia objects. For example, a virtual human should be a 3D graphical object, but also an object able to speak or emit sounds. A virtual dog should be able not only to move, but also to bark

Europe Union has published six reports on Visionary Research. One of the reports [1] is entitled “From Multimedia to Telepresence”. The purpose of this document is to identify visions in the area of the future multimedia systems and the technologies associated to these visions. Seven “key visions” have been identified as clusters of elementary visions. Figure 1 shows a diagram with the 7 key visions and the elementary visions. We just want to discuss a little more the second and the third key visions, thinking that their integration will be the future of multimedia systems.

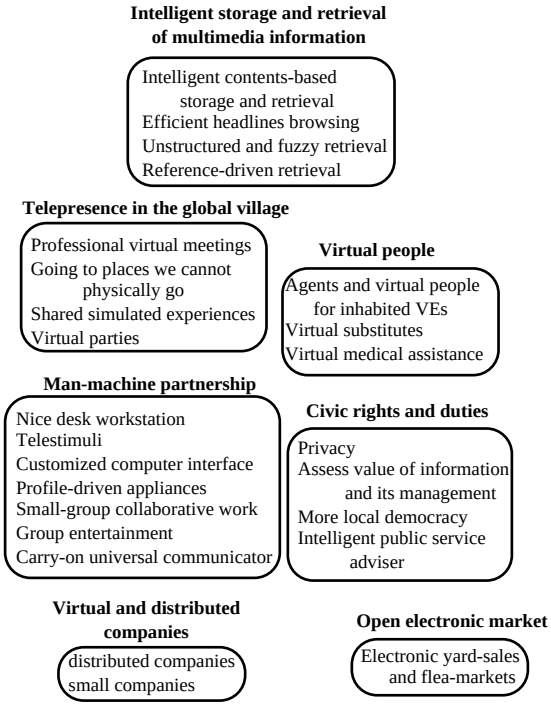

Fig. 1. Seven key visions

*Published in Proc. Intern. Workshop on Synthetic - Natural Hybrid Coding and Three Dimensional Imaging, Rhodes, 1997, pp. 3-7

1.1 Telepresence in the global village

Telepresence is the future of multimedia systems and will allow participants to share professional and private experiences, meetings, games, parties. The virtual environment should allow effective collaboration among users at different sites. This necessitates incorporating different media in the complex environment, such as image, real-time movies, 3D objects. Instead of having different applications or windows for each medium, the 3D virtual environment should be able to provide easy mechanisms to integrate these applications. There are four representative areas where telepresence is a key issue:

Professional virtual meeting with real people
During virtual meetings the participants need a better view of each other and much more freedom than they have now in the present awkward form of videoconferencing with its static positioning. As appropriate, 3D sound will be used to allow spatial localisation, and participants will be able to manipulate multimedia documents.

Going to places we cannot physically go
For investigation and, following that, for adventure, real places and environments that are only attainable with difficulty can be opened up using telepresence technologies.

Shared simulated experiences
Individual activity like simulated driving, flying, boating etc. can be enhanced by sharing with others, occupying the same or contiguous virtual space. The objectives of people taking part in these shared experiences can be to learn as well as to be entertained.

Virtual parties
Today there are many chat lines and also a few virtual chat rooms on the Internet (e.g. Palace from The Palace Corporation) where people can meet represented either symbolic to respect their anonymity or more realistic as an animated figure with the user's face projected onto. Offering the possibility for people to stroll around in a realistic looking 3D environment listening to other people with all kinds of multimedia services is a direct extension of the textual or 2D virtual chatting rooms.

1.2 Virtual people

The sense of "presence" in the virtual environment is an important requirement for collaborative activities involving multiple remote users working with social interactions. Using agents and virtual actors within the shared environment is a essential supporting tool for presence. Their role is very important in virtual environments with many people like virtual airports or even virtual cities.

Three representative areas were identified:

Agents and virtual people for Inhabited Virtual Environments
Two types of real-time virtual humans may coexist in the same shared virtual environment: the guided ones and autonomous ones. Guided humans or avatars are useful to represent ourselves, the users or participants. Autonomous humans are able to act on their own. Based on the perceived information, the virtual human's behavioral mechanism will determine the actions it will perform. The implementation of such types of virtual humans is explained in more details in the next section.

Virtual substitutes
A virtual substitute is an intelligent computer-generated agent able to act instead of the real person and on behalf of this person on the network. The virtual substitute has the voice of the real person and his or her appearance. He will appear on the screen of the workstation/TV, communicate with people, and have predefined behaviors planned by the owner to answer to the requests of the people.

Virtual medical assistance
Nowadays, it seems very difficult to imagine an effective solution for chronic care without including the remote care of patients at home by a kind of Virtual Medical Doctor. The modelling of virtual patient with correspondence to medical images is also a key issue and a basis for telesurgery.

2 Meeting virtual people in the global village: some research developments

Based on the EU report and some current work, we may now state that the multimedia system of the future will be a system allowing the users to share a Virtual World with 3D objects they can manipulate and transform. These objects should be able to react, to move autonomously, and to have properties to recognize and generate different media: sound, speech, images, gestures, etc. The most advanced object is

surely the virtual human who will play a key role in the new multimedia systems. As already discussed in the previous section, two types of real-time virtual humans may coexist in the same shared virtual environment: the guided ones and autonomous ones. But, we may be more specific and distinguish four types of Virtual Humans:

- participants or clones
- guided actors or avatars
- autonomous actors
- interactive perceptive actors

In the next Sections, we will discuss these four types of actors and show a few models and applications we have developed in our laboratory.

2.1 Participants or clones

The virtual actor is required to have a natural-looking body and be animated correlated to the actual body. The technique, may be called *real-time rotoscopy method* [2] and consists of recording input data from a VR device in real-time allowing to apply at the same time the same data to the virtual actor on the screen. A popular way of animating such an actor is the use of sensors like the Flock of Birds.

2.2 Guided Actors or Avatars

Guided actors are actors which are driven by the user but which do not correspond directly to the user motion. They are based on the concept of *real-time direct metaphor* [2]. The participant uses input devices to update the position of the virtual actor. This local control is used by computing the incremental change in the actor position, and for example, estimating the rotation and velocity of the center of body.

2.3 Autonomous Actors

Autonomous actors are able to have a behavior, which means they must have a manner of conducting themselves. Typically, the actor should perceive the objects and the other actors in the environment through virtual sensors [3]: visual, tactile and auditory sensors. Based on the perceived information, the actor's behavioral mechanism will determine the actions he will perform. An actor may simply evolve in his environment or he may interact with this environment or even communicate with other actors. In this latter case, we will consider the actor as a interactive perceptive actor.

The concept of **virtual vision** was first introduced by Renault et al. [4] as a main information channel between the environment and the virtual actor. The synthetic actor perceives his environment from a small window in which the environment is rendered from his point of view. As he can access z-buffer values of the pixels, the color of the pixels and his own position, he can locate visible objects in his 3D environment. To recreate the a **virtual audition** [5], in a first step, we had to model a sound environment where the synthetic actor can directly access to positional and semantic sound source information of a audible sound event. For **virtual tactile sensors,** our approach [6] is based on spherical multi-sensors attached to the articulated figure. A sensor is activated for any collision with other objects. These sensors have been integrated in a general methodology for automatic grasping (see Figure 2).

Fig. 2. Examples of grasping different objects

2.4 Interactive Perceptive Actors

We define an interactive perceptive synthetic actor as an actor aware of other actors and real people. Such an actor is also assumed to be autonomous of course. Moreover, he is able to communicate interactively with the other actors whatever their type and the real people.

2.5 Intercommunication between Synthetic Actors

Behaviors may be also dependent on the emotional state of the actor. A nonverbal communication [7] is concerned with postures and their indications on what people are feeling. Postures are the means to communicate and are defined by a specific position of the arms and legs and angles of the body. These nonverbal communication is essential to drive the interaction between people without contact or with contact (Figure 3).

Fig. 3. Nonverbal intercommunication

2.6 Sensing the Real World for a Synthetic Actor

The real people are of course easily aware of the actions of the synthetic actors through VR tools like Head-mounted displays, but one major problem to solve is to make the virtual actors conscious of the behavior of the real people. Virtual actors should sense the participants through their virtual sensors. Such a perceptive actor would be independent of each VR representation and he could in the same manner communicate with participants and other perceptive actors. Perceptive actors and participants may easily be exchanged as demonstrated with the interactive game facility. For virtual audition, we encounter the same problem as in virtual vision. The real time constraints in VR demand fast reaction to sound signals and fast recognition of the semantic it carries. Concerning the tactile sensor, we may consider the following example: the participant places an object into the Virtual Space using a CyberGlove and the autonomous virtual actor will try to grasp it and put it on a virtual table for example. The actor interacts with the environment by grasping the object and moving it.

As an example, we have produced a fighting between a real person and an autonomous actor. The motion of the real person is captured using a Flock of Birds. The gestures are recognized by the system and the information is transmitted to the virtual actor who is able to react to the gestures and decide which attitude to do. Figure 4 shows an example.

Fig. 4. Fight between a participant and an interactive perceptive actor

2.7 Real-Time Synthetic Actors in Distributed Virtual Environments

The VLNET [8] (Virtual Life NETwork) system supports a networked shared virtual environment that allows multiple users to interact with each other and their surrounding in real time. The users are represented by 3D virtual human actors, which serve as agents to interact with the environment and other agents. The agents have similar appearance and behaviors with the real humans, to support the sense of presence of the users in the environment. In addition to guided actors, the environment can also include autonomous and interactive perceptive actors used as a friendly user interface to different services. Virtual humans can also be used in order to represent the currently unavailable partners, allowing asynchronous cooperation between distant partners.

As application, we selected a virtual tennis game with a guided actor, the interactive perceptive actress Marilyn, and an interactive perceptive referee, both with a virtual vision. Figure 5 shows an interactive session. The geometric characteristics of the tennis court make part of the players knowledge. For the dynamics simulation of the ball, gravity, net, ground and the racquet we use physics-based calculations. The tracking of the ball by the vision system is controlled by a special automata that tracks the ball, estimates the collision time and collision point of ball and racquet and performs successfully a hit with given force and a given resulting ball direction. The referee judges the game by following the ball with his vision system. Marilyn can also hear sound events and obeys the decisions of the referee.

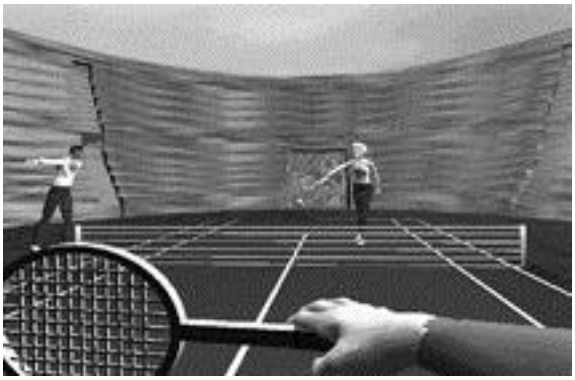

Fig. 5. Tennis game with real-time participant and interactive perceptive actors

3 Conclusion

Telepresence is the future of multimedia systems and will allow participants to share professional and private experiences, meetings, games, parties. The concepts of Distributed Virtual Environments is a key technology to implement this telepresence. Using virtual actors within the shared environment is a essential supporting tool for presence. Real-time realistic 3D avatars will be essential in the future, but we will need interactive perceptive actors to populate the Virtual Worlds. The ultimate objective in creating realistic and believable virtual actors is to build *intelligent autonomous* virtual humans with *adaptation*, *perception* and *memory*. These actors should be able to act *freely* and *emotionally*. Ideally, they should be *conscious* and *unpredictable*. But, how far are we from such a ideal situation ? Our interactive perceptive actors are able to perceive the virtual world, the people living in this world and in the real world. They may act based on their perception in an autonomous manner. Their intelligence is constrained and limited to the results obtained in the development of new methods of Artificial Intelligence. However, the representation under the form of virtual actors is a way of visually evaluating the progress. In the future, we may expect to meet intelligent actors able to learn or understand a few situations.

References

[1] D.Thalmann, L.Chiariglione, F.Fluckiger, E.H. Mamdani, M.Morganti, J.Ostermann, J.Sesena, L.Stenger, A.Stienstra, Report on Panel 6: From Multimedia to Telepresence, Expert groups in Visionary Research in Advanced Communications, ACTS, European Commission, 1997.

[2] D.Thalmann, Using Virtual Reality Techniques in the Animation Process in: Virtual Reality Systems (Earnshaw R, Gigante M, Jones H eds), Academic Press, pp.143-159.

[3] D. Thalmann, Virtual Sensors: A Key Tool for the Artificial Life of Virtual Actors, Proc. Pacific Graphics '95, Seoul, Korea, 1995, pp.22-40.

[4] O.Renault, N.Magnenat Thalmann, D.Thalmann, A Vision-based Approach to Behavioural Animation, The Journal of Visualization and Computer Animation, Vol 1, No 1, 1990, pp.18-21.

[5] H.Noser, D.Thalmann, Synthetic Vision and Audition for Digital Actors, Proc.Eurographics‘95, 1995, pp.325-336.
[6] Z.Huang, R.Boulic, N.Magnenat Thalmann, D.Thalmann, A Multi-sensor Approach for Grasping and 3D Interaction, Proc. CGI ‘95, Academic Press, 1995, pp.235-254.
[7] P.Bécheiraz, D.Thalmann, A Model of Nonverbal Communication and Interpersonal Relationship between Virtual Actors, Proc.Computer Animation ‘96, IEEE Computer Society Press, 1996.
[8] T.K.Capin, I.S.Pandzic, N.Thalmann, D.Thalmann, Virtual Human Representation and Communication in the VLNET Networked Virtual Environments, IEEE Computer Graphics and Applications, Vol.17, No2, 1997, pp.42-53.