1995

Real-time Upper Body Articulation of Humans in a Networked Interactive Virtual Environment

Waldrop, Marianne S.

Real-time Upper Body Articulation of Humans in a Networked Interactive Virtual Environment

Marianne S. Waldrop[1], Shirley M. Pratt, David R. Pratt, Robert B. McGhee, John S. Falby, Michael J. Zyda
Department of Computer Science
Naval Postgraduate School
Monterey, CA 93943

ABSTRACT

With the recent introduction of realistic human entities into large-scale networked virtual environments, there exists a requirement for dynamic, real-time human motion. The future for distributed virtual environments will include upwards of 100,000 participants capable of participating from anywhere in the world, interacting with any entity that exists in the environment. In order to attain a high level of realism in the virtual world, the user must be able to dynamically interact with his environment. For the lower body, we find that scripted locomotive motion is adequate. However, the same is not true for upper body motion because humans by their nature interact with their environment largely with their hands. The focus of this research is the development of an interactive interface which will achieve dynamic upper body motion currently not possible with scripted systems. We present the basics of ongoing work involving the representation of realistic, real-time upper body motion of a virtual human in a networked virtual environment using magnetic sensors attached to the user.

1. Contact Author: waldrop@cs.nps.navy.mil.

Introduction

Recently, the computer power available to the Virtual Environments (VE) researcher has progressed to a point where realistic environments and dynamic entities can be portrayed in real-time. This has been particularly important to the pursuit of one of the grandest goals in VE research, insertion of humans into large-scale realistic interactive worlds. While we are still far from achieving all the goals of virtual human interactions, major strides have been taken.

We have focused our initial efforts on representing real-time arm motions of humans. This significantly reduces the number of degrees of freedom in the problem. We utilize scripted motions to drive leg movements of our human icon. Focusing on interactive arm motion representation does not result in a significant loss of realism because interactions with the virtual world are largely carried out with the hands.

Previous Work

The Individual Portal (I-PORT) developed by SARCOS Inc., Army Research Laboratory - Human Research and Engineering Directorate (ARL-HRED), University of Pennsylvania, and the Naval Postgraduate School (NPS) is currently one of the few force feedback motion devices able to insert a human into a large scale networked VE[1] [2]. Interactivity across the network is

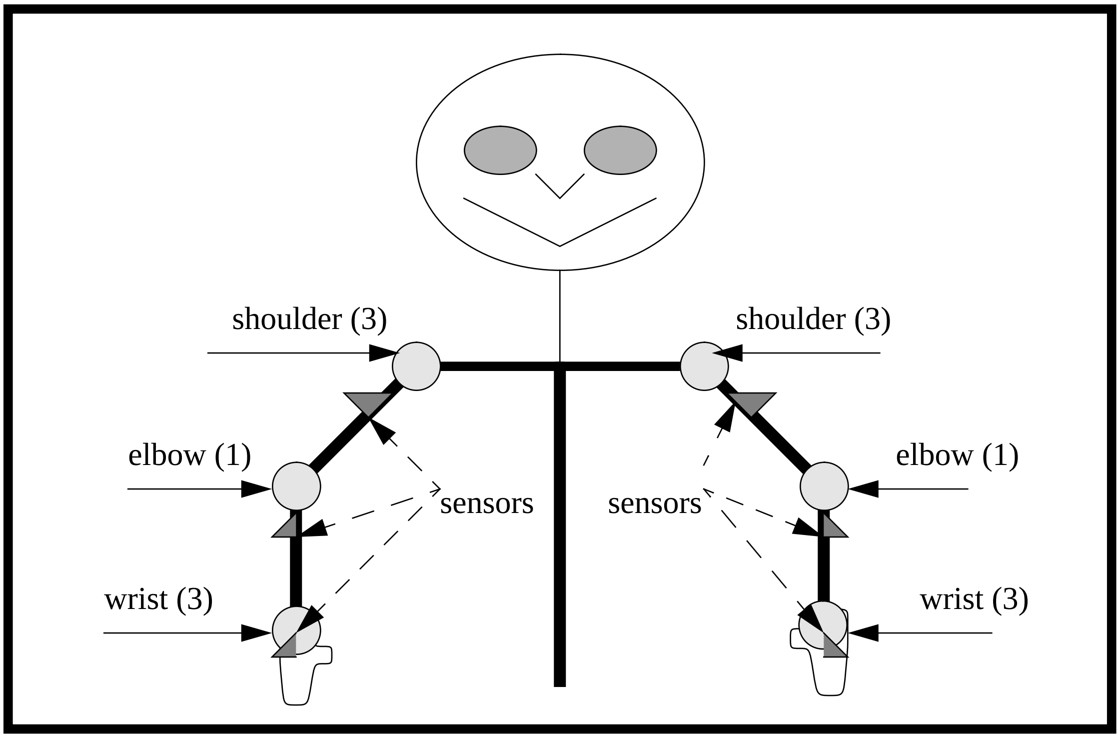

FIGURE 1. Upper body joints with associated number of DoFs for each joint

achieved by means of Distributed Interactive Simulation (DIS) Protocols [3] and multicast networking. While the lower body locomotion of the virtual individual is adequately represented using scripted motions, the upper body motion still remains open to improved means of specification. To solve this problem, there are four basic approaches: University of Pennsylvania's Jack or Jack Motion Library (JackML), the SARCOS Sensor Suit, hands constrained to remain locked on the rifle, and sensors on the arms. In the course of this project, the first three systems have been tried with varying degrees of success. The fourth method has been accomplished by the University of Pennsylvania with different objectives achieved.

Jack, a program developed by University of Pennsylvania's Center for Human Modeling and Simulation to manipulate articulated figures [4], was used to provide body joint angles in the first demonstration of the I-PORT at Ft. Benning, GA, during INCOMSS 94 in February 1994 [1][4]. While this system proved successful, there was significant system overhead involved communicating with Jack. Joint information was transferred to a separate machine via a specified port, strongly impacting performance. Also at the same demonstration, the SARCOS sensor suit was used. It allowed the user to control the upper body of the icon in real-time by overriding the Jack joint angles. However, the suit was cumbersome, difficult to adjust (the suit had to be recalibrated for each individual), and measurements were often accompanied by high noise levels, resulting in jerky motion.

During May and October 1994, the I-PORT was shown again. The May 1994 demonstration utilized Jack as before. However, the October 1994 demonstration used JackML, a library of joint angles linked to the main application, which was able to be queried in real-time on a single machine. Due to the volume of people at the demonstration, it was impractical to fit them with sensor suits. To solve the problem of specifying upper body joint angles, the icon kept his hands locked on the rifle and inverse kinematics was used to compute joint angles for the arms. While this was visually acceptable when the soldier was engaged in combat, he was never able to put down his rifle or to hold it with one hand.

Efforts at the University of Pennsylvania have been made to realistically recreate a human posture and position with minimal sensors for a less encumbered operator. They accomplished this by using four 6 degrees of freedom (DOF) magnetic sensors: one on each palm, on the waist, and on the head. Clearly, the objectives are different in that their paradigm requires more intensive mathematical calculations than those desired to meet the demands of a real-time networked environment.[5]

This led us to seek a solution similar to the Reality Built for Two (RB2) system [6]. In it, VPL used fiber optic cables to measure joint angles. The result was a fully specified set of angles[1]. Rather then using a suit containing the fiber optics, we propose to use a series of Polhemus sensors, which can be configured to be less cumbersome and more durable than the body suits described earlier.

Implementation

In order to achieve a robust and intuitive human interface, the human operator should be equipped with sensors. This type of interface will result in accurate, real-time simulation of the upper body based on actual motions of the user. A variety of Polhemus sensor configurations and associated mathematical solutions are explored in order to achieve this real-time human upper body articulation in an interactive networked virtual environment. As seen in Table 1, there is some flexibility in choosing the number of sensors to be used, resulting in potential speed and motion resolution trade-offs.

Using Jack software, users are able to manipulate a high resolution model's joint angles to achieve some motion goals in the virtual world. The resulting joint motions are then further constrained to similarly manipulate a medium resolution human model. Lastly, they are recorded for fast real-time replay via JackML. Although the medium resolution human model of the JackML software is fully articulated, having a total of 39 DOF in 17 separate joints, we are only concerned with the model's upper body structure, and in particular, the arms. The arms of the medium resolution model consist of a total of 8 joints, but the two clavicles will not be articulated in this application. This simplification still produces visually appealing results since clavicle joint values have been observed to remain nearly constant for the majority of the arm movements we are interested in rendering. Articulating the remaining 6 joints of the upper body still results in realistic arm movements.

Each joint has one or more rotational DOFs, as noted by the parenthesized numbers in Figure 1. In our current implementation, JackML is able to provide realistic lower body motions. However, interactive upper body motions need to be specified based on the user's desired real-time interactions with his environment by overriding the upperbody values from JackML.

In order to allow users to naturally manipulate the virtual human in real-time, the Polhemus *3Space* Fastrak sensor system will be used. The small body sensor receivers in conjuction with the "Long Ranger", a ceiling-mounted transmitter with a range of 15 feet, will be used to sense the human upper body motion. Based on the structure of the human model and analysis of information shown in Table 1, the path we are choosing consists of one sensor placed on the head-mounted display (HMD) and three sensors placed on each arm. The HMD sensor will sense

1. For the purpose of this paper, a system is "Fully Specified" when all joint angles are able to be directly measured. When all the joint angles can be uniquely determined, the system is said to be "Completely Specified."

Joint	Arm Segment	Joint DOF	Cumulative DOF	Number of Sensors for Complete Specification	Number of Sensors for Full Specification
Shoulder	Upper Arm	3	3	1	1
Elbow	Forearm	1	4	1	2
Wrist	Hand	3	7	2	3

TABLE 1. Joint degrees of freedom and sensor requirements

head motions and act as the reference location for the human body. Figure 1 shows the proposed placement of arm sensors. They are placed on the upper arm, on the forearm, and on the back of the hand, outboard of the wrist. Due to the anticipated types of motions in our environment, it is not critical to articulate the clavicles.

Polhemus sensors are able to return 6 DOFs: the position, X, Y and Z, and orientation, heading, pitch and roll (H, P and R, respectively). Because six sensors, located on the arms, are returning six parameters each, we have a total of 36 known parameters, and from Table 1, a total of 14 upper body joints which need to be calculated. This is what we refer to as a “fully specified” solution. This is desirable in order to alleviate problems which can occur when constraining a highly redundant problem where goal-positioning tasks have an infinite number of solutions [7].

Paradigm For Implementation

NPSNET, a large-scale, DIS-compliant, 3D networked virtual world, is managed and developed by the students, staff, and faculty at the Naval Postgraduate School (NPS). It is a virtual combat environment inhabited by dismounted infantry and military vehicles which can move, but have limited interaction amongst themselves [1]. Currently, system users interact with the environment through conventional input devices such as keyboard, flight control sticks (FCS) and SpaceBall. The FCS maintain control of speed and direction of travel of the driven human entity, weapon firing capability, and limited posture control. When FCS are used to control a virtual human, location, speed and posture of the infantryman can be interactively changed. Limited static and dynamic human postures along with intermediate posture changes are available as interactive choices for dismounted infantrymen. Additionally, users can modify head motions of the simulated humans in real-time through a VIM HMD [8]. However, no interface has yet been incorporated in NPSNET which allows real-time manipulation of the upper body. Keyboard commands exist which can trigger the virtual human to perform scripted upper body motions (i.e. hand signals), but no “on-the-fly” manipulation is yet possible.

NPSNET uses hierarchical networking techniques to represent humans on remote machines sharing a virtual environment. Information regarding the position, orientation and an enumerated posture (i.e. upright still, walking, running, kneeling, prone, crawling, etc) of a human entity is communicated over the network at five second intervals in DIS entity state protocol data units (EPDUs). More frequent updates are sent if the parameters exceed arbitrarily chosen threshold values for using simple dead reckoning algorithms to interpolate the entity’s position between DIS EPDU updates. The use of DIS articulated parameter records to append values for numerous joint angles to EPDUs is very inefficient requiring four times the amount of data necessary to describe joint angle data [1]. Instead, NPSNET communicates joint angle information in parallel using efficiently packed multicast network packets. Thus, DIS EPDUs broadcast basic low resolution information about a human entity to all networked machines. Multicast network packets are used to disseminate high resolution information about human joint motions which can be selectively received by interested remote machines subscribing to multicast network groups.

Conclusions

Research conducted in preparation for implementing the upper body sensor interface indicates that the fully specified solution can achieve fast joint angle calculations to perform real-time upper body motion in the virtual networked environment. Additionally, using three sensors per arm provides realistic movements in light of the resulting tightly constrained solution that we have created through the number and placement of sensors in this interface structure. However, the complete solution is a viable option if this number of sensors is not available, albeit at the expense of speed and/or realistic arm movements.

As the number of calculations to update the movement of the human entity increases, so does the lag or latency between movement updates, thus degrading the quality and realism of the networked environment. Furthermore, there is more risk of producing less natural movement when redundancies have to be constrained. Ensuring natural movement and speed of the virtual human and real-time performance are of primary importance in achieving an effective human interface in a large-scale networked virtual environment.

Acknowledgments

We would like to thank members of the University of Pennsylvania Center for Human Motion Studies and SARCOS Research, Inc for their collaborative efforts on this project. This work would not have been possible without the support of our sponsors at Advanced Research Project Agency (ARPA) and the Army Research Laboratory (ARL).

References

[1] Pratt, D.R., et al., “Insertion of an Articulated Human into a Networked Virtual Environment,” *Fifth Annual Conference on AI, Simulation and Planning in High Autonomy Systems*, Dec 7-9, 1994.

[2] Granieri, J. P., Badler, N. I., “Simulating Humans in VR,” *Applications of Virtual Reality*, Academic Press, 1995. To appear.

[3] Institute for Simulation and Training, Orlando, FL. Standard for Distributed Interactive Simulation -- *Application Protocols* (v 2.0, 4th draft, revised), 1993.

[4] Badler, N.I., Phillips, C.B., Webber, B.L., *Simulating Humans: Computer Graphics Animation and Control*, Oxford University Press, New York, 1993.

[5] Badler, N.I., Hollick, M.J., and Granieri, J.P., “Real-Time Control of a Virtual Human Using Minimal Sensors,” *Presence: Teleoperators and Virtual Environments*, Winter 1993 , Volume 2, Number 1, MIT, pp. 82-86.

[6] Blanchard, C., et. al., “Reality Built for Two: A Virtual Reality Tool,” *Proceedings of the 1990 Symposium on Interactive 3D Graphics*, Vol. 24, No. 2, March 1990.

[7] Phillips, C.B., Zhao, J., Badler, N. I., “Interactive Real-time Articulated Figure Manipulation Using Kinematic Constraints,” *Proceedings of the 1990 Symposium on Interactive 3D Graphics*, Vol. 24, No. 2, March 1990, pp. 245-249.

[8] Pratt, S.M., et al., “Humans in Large-Scale, Real-time, Networked Virtual Environments,” submitted to *Presence*.