HUMAN IMMERSION INTO THE DIS BATTLEFIELD

Robert Anschuetz (Veda), Charlie Jones (Veda), Patrick Kelly (Veda), James Molnar (Veda), and Traci Jones (STRICOM)

Veda Incorporated
3452 Lake Lynda Drive, Suite 215
Orlando, FL 32817
407/658-0044
{anschuer|jonesc|kellyp|molnarj}@orlando.veda.com

STRICOM
12350 Research Parkway
Orlando, FL 32826-3275
407/380-4340
jonest@stricom.army.mil

KEYWORDS

Virtual Reality, Real-time Motion Capture, Dismounted Infantry, Distributed Interactive Simulation, Virtual Battlefield, Synthetic Environment, Head Mounted Display

Abstract

Recent advances in human motion capture and head mounted display technologies, coupled with Distributed Interactive Simulation (DIS) capabilities, now allow for the implementation of an untethered, fully-immersable, DIS-compliant, real-time Dismounted Soldier Simulation (DSS) System. The untethered soldier, outfitted with a set of optical markers and a wireless helmet-mounted display, can move about freely within a real-world motion capture area, while position and orientation data are gathered and sent onto a DIS network via tracking cameras and image processing computers. The soldier's interaction in the virtual environment includes the ability to move within the battlefield unencumbered by wires or other peripheral devices, fire an M16A2 rifle, hear DIS battlefield audio, and communicate with other entities via a DIS radio simulator. Fully articulated human motion rotations and translations are sent onto the DIS network using Entity State and Data PDUs. Along with position and orientation information, the dismounted soldier's discrete state is transmitted in the Entity State PDU appearance field so that all receiving entities know *what* the virtual soldier is doing (i.e., running, walking, or crawling). Data PDUs are sent out with the real-time motion information so that simulations interested in displaying an articulated human figure know *how* the figure is moving. The Data PDUs are 288 bytes in length and are sent out at a frequency of 1 to 30 per second. When the soldier pulls the trigger on the M16A2, a wireless signal is sent to the host computer, which generates Fire and Detonation PDUs. Data has been captured and analyzed in the following areas: motion capture accuracy, transport delay, latency, image refresh rate, bandwidth usage, firing accuracy, and simulator sickness.

HUMAN IMMERSION INTO THE DIS BATTLEFIELD

1. INTRODUCTION

1.1. Background

Immersing the soldier into the synthetic environment is vital to training operations such as search and rescue, hostage extraction, and warfighting scenarios. A simulator that can place an individual soldier in a virtual environment can be used to support decisions regarding material, doctrine, tactics, techniques, procedures and force structure. In order to carry out realistic training, the soldier must be able to interact with the environment in real-time, preferably unencumbered by cables and heavy equipment.

Recent advances in motion capture technology and display systems allow the immersion of the dismounted soldier into the synthetic Distributed Interactive Simulation (DIS) battlefield. Previous systems have tethered the participant to cumbersome cables and large peripheral devices for tracking purposes, hindering the participant's movement capability. Also, many previous systems have ignored leg and waist capture information, choosing instead to focus only on the upper body movement.

1.2. Dismounted Soldier Simulation

The Dismounted Soldier Simulation (DSS) System immerses the soldier into the synthetic environment through the use of an untethered helmet mounted display (HMD), an optical set of motion capture markers, a DIS software layer, and image generation software, hosted on a two processor Silicon Graphics Onyx RE2 workstation.

The virtual soldier can move about freely in the gaming area while position and orientation data are sent onto a DIS network at real-time update rates. The soldier's weapon (which need not be physically in the possession of the soldier at all times) is also immersed in the synthetic environment.

The DSS system updates the three-dimensional display at 15 Hz to 30 Hz, based upon scene content. The more complex the scene the slower the frame rate. A scene is composed of the terrain database, static models such as buildings and trees, and dynamic DIS models such as tanks and aircraft.

2. TECHNICAL DESCRIPTION

2.1. Motion Capture System (MCS)

Figure 1 depicts the DSS Motion Capture System which consists of a set of 16 optical markers placed on the soldier's body, 3 placed on an M16A2, and a wireless helmet mounted display. The markers are tracked in a 12' x 12' gaming area by a set of 4 cameras that transmit the imagery to image processing (IP) hardware. The IP hardware digitizes the marker information and transmits this information to the host computer / image generator for conversion to rotation and translation values for the three-dimensional model rendering.

Battlefield audio and radio capabilities are generated by off-the-shelf solutions which are integrated with the DSS System to provide a realistic training environment. Figure 1 depicts the integrated DSS system as a block diagram.

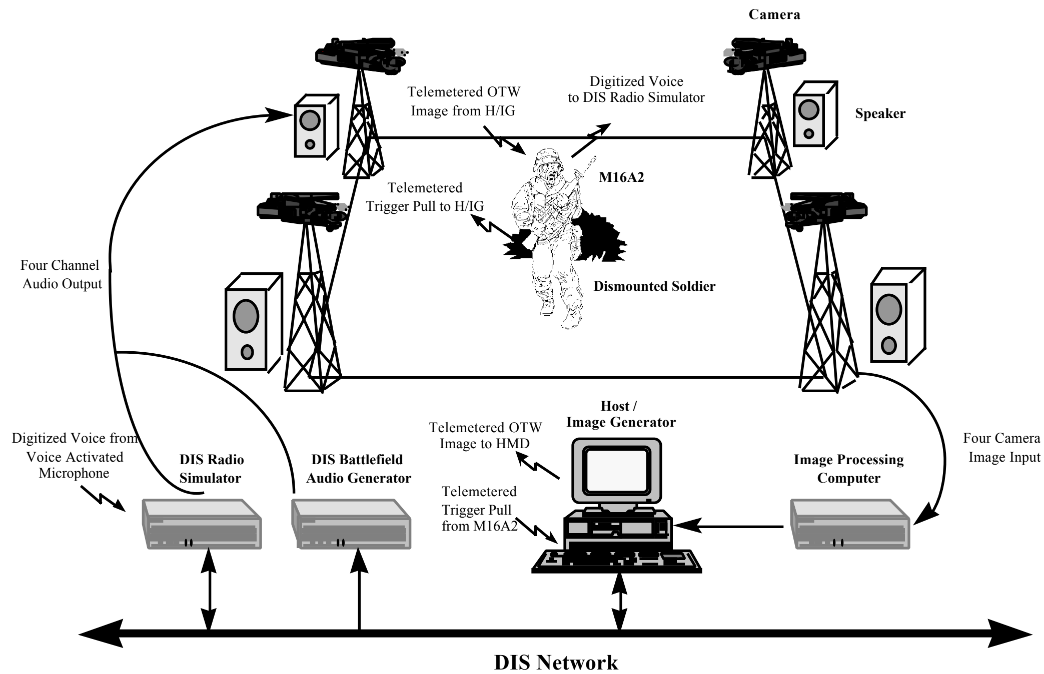

Figure 1. The Dismounted Soldier Simulation (DSS) System.

2.2. Real-time Protocol Data Unit Generation over the Distributed Interactive Simulation network

One of the most significant challenges to the DSS project was the delivery of real-time motion information to other DIS participants. The DSS sends motion information over the DIS network via both Entity State and Data Protocol Data Units (PDUs).

The Entity State PDUs allow other participants in the DIS simulation to receive the soldier's position, orientation data, and discrete state, which is broadcast via pre-defined enumerations in the appearance field. The following discrete states are supported in the DSS simulation: standing, prone, running, walking, jumping, kneeling, rolling, and crawling. The discrete states are generated based on body position as well as positional and angular velocities. Thresholds were fine-tuned as a result of research performance.

Fully articulated motion data was not included in the Entity State PDU since articulation parameters have not yet been defined for the PDU and the size of the PDU would have been close to 1000 bytes. Since no articulation parameters are used in the DSS Entity State PDUs, the PDUs are 144 bytes in length. In the event that the Dismounted Soldier is not moving, an Entity State PDU is transmitted once every five seconds, otherwise they are transmitted based on dead reckoning of the waist position and orientation data only. Entity State PDUs are also transmitted based on appearance changes.

For those entities which can understand and use real-time motion information, real-time data is transmitted in Data PDUs. Data PDUs are transmitted at a frequency set by the DSS operator. Valid frequency ranges from 1 Hz to 30 Hz. The rate can be modified in order to lessen bandwidth impact at the cost of smoothness of motion. Each Data PDU contains a fixed datum motion dataset which completely describes the position and orientation data of all of the human model's articulated parts in a semi-compressed format.

The units of the rotation angles are radians and the units of the translation values are meters. Also present in the Data PDU is the discrete state as described in the Entity State PDU. Table 1 describes the experimental datum IDs defined for the Data PDU as used for the DSS system. There are a total of 31 fixed datum fields and no variable datum fields which yield a Data PDU size of 288 bytes.

ID	Name	Type	Length (bits)	Units
80000	Dismounted Infantry			
80100	Human X offset	integer	32 bits	meters
80110	Human Y offset	integer	32 bits	meters
80120	Human Z offset	integer	32 bits	meters
80130	Weapon X offset	integer	32 bits	meters
80140	Weapon Y offset	integer	32 bits	meters
80150	Weapon Z offset	integer	32 bits	meters
80160	Rotation Constant/ Translation Constant	unsigned integer/ unsigned integer	16 bits/ 16 bits	
80170	Discrete State	unsigned integer	32 bits	
80180	Back X Rotation/ Back Y Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80190	Back Z Rotation/ Left Upper Leg X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80200	Left Upper Leg Y Rotation/ Left Upper Leg Z Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80210	Left Lower Leg X Rotation/ Left Foot X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80220	Left Foot Y Rotation/ Left Foot Z Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80230	Right Upper Leg X Rotation/ Right Upper Leg Y Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80240	Right Upper Leg Z Rotation/ Right Lower Leg X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80250	Right Foot X Rotation/ Right Foot Y Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80260	Right Foot Z Rotation/ Head X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80270	Head Y Rotation/ Head Z Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80280	Left Shoulder X Rotation/ Left Shoulder Y Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80290	Left Shoulder Z Rotation/ Left Upper Arm X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80300	Left Upper Arm Y Rotation/ Left Upper Arm Z Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80310	Left Lower Arm X Rotation/ Right Shoulder X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80320	Right Shoulder Y Rotation/ Right Shoulder Z Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80330	Right Upper Arm X Rotation/ Right Upper Arm Y Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80340	Right Upper Arm Z Rotation/ Right Lower Arm X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80350	Weapon X Rotation/ Weapon Y Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80360	Weapon Z Rotation/ Waist X Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80370	Waist Y Rotation/ Waist Z Rotation	unsigned integer/ unsigned integer	16 bits/ 16 bits	radians/ radians
80380	Weapon X Translation/ Weapon Y Translation	unsigned integer/ unsigned integer	16 bits/ 16 bits	meters/ meters
80390	Weapon Z Translation/ Waist X Translation	unsigned integer/ unsigned integer	16 bits/ 16 bits	meters/ meters
80400	Waist Y Translation/ Waist Z Translation	unsigned integer/ unsigned integer	16 bits/ 16 bits	meters/ meters

Table 1. DSS Data PDU Datum Specification Record

2.3. Image Generation

The Silicon Graphics Onyx Reality Engine 2 host / image generator is responsible for the accurate rendering of the three-dimensional human model inside the terrain database. The terrain database used in the DSS program is the Quantico Village database developed by Paradigm, Inc. for the Team Tactical Engagement System for NAWCTSD. This database was overlaid onto a 200 meter by 200 meter area of the Hunter Liggett database in order to provide interaction with other DIS simulations, including ModSAF and the IST CGF.

In order to realistically immerse the soldier into a virtual environment, a helmet mounted display is used. Head tracking is provided by the four camera motion capture system. The DSS generates a viewport corresponding to the head position and orientation of the virtual soldier. This viewport is sent to the HMD of the "real" soldier so that he can see the entities and terrain in which he is interacting. The soldier can see himself in his view by either looking down at his legs or by holding his hands in front of his face. The Silicon Graphics is not configured for stereo visual output, so the same image is sent to both channels in the HMD.

The DSS system also supports a stealth viewport for use by an instructor or another simulator. The stealth view depicts the 1600 polygon, fully articulated human movement of the dismounted infantryman. This stealth view permits the instructor to fly around the database to watch and critique the behavior of the soldier as he interacts with other entities.

2.4. Network Bandwidth Limitations

Real-time motion information can be transmitted to other DIS entities in two ways: 1) discrete state information or 2) continuous motion information. The first method places the burden of rendering the 3D figure's real-time motion on the receiving simulator. Transitional "morphing" between states (i.e., from prone to standing to walking) must be performed by the receiving simulator. The second method places a burden on the sending simulator and also on the network which must support high bandwidth rates for real-time motion of a complex articulated model. An Ethernet network, with a bandwidth of 10 megabits per second (Mbps), theoretically could support approximately 150 DSS systems, each generating 30 Hz real-time motion data. However, many studies have shown that the realistic Ethernet bandwidth limitation is approximately 4 Mbps , therefore allowing approximately 60 DSS systems running simultaneously with no other entities on the network. An FDDI network, with a bandwidth of 100 Mbps, theoretically could support approximately 1500 DSS systems, each generating 30 Hz real-time motion data, before reaching its capacity.

2.5. Movement within the Database

With the gaming area limited to 12'x12', extensive movement in the simulated battlefield over wide areas was not possible without an alternative means of traversing the virtual environment. The DSS currently utilizes a "forced walk/run" technique based on the soldier's position within the gaming area.

Concentric circles of diameters 8 and 10 feet are drawn on the gaming area which provides 3 distinct areas of movement: 1) within the 8 foot diameter circle, the "real" soldier's movements within the gaming area are transferred directly to the simulated battlefield, 2) between the 8 foot and 10 foot radius circles, the soldier is placed in a "forced walk" condition at 1.5 meters per second, and 3) outside the 10 foot radius circle, the soldier is placed in a "forced run" condition at 3.0 meters per second. The direction the soldier moves is the direction of the vector from the center of the circle to the soldier's position. This can be visualized by imagining the circle as a large compass where the soldier determines the point of the arrow.

2.6. Radio Communication

Command and control communications between the DSS and other DIS entities has been modeled via DIS Signal, Transmitter, and Receiver PDUs through the use of an off-the-shelf radio simulator. A voice activated microphone permits transmission of voice to the radio simulator, which uses DIS Signal and Transmitter PDUs to transmit the voice data to other simulators on the DIS network. The radio simulator receives Signal and Transmitter PDUs from the DIS network and responds with Receiver PDUs. The decoded voice data is sent to the dismounted combatant through large speakers placed on the periphery of the gaming area. The audio can alternatively be sent to wireless headphones if desired.

2.7. Battlefield Audio

Battlefield audio has been captured off the DIS network and transferred to the live soldier in the DSS system through the use of an off-the-shelf audio generation tool. As DIS entities move, fire, and detonate around the soldier, he can hear the audio generated by the entities in realistic stereo as if he were immersed in the battlefield.

2.8. M16A2 Rifle

The DSS currently supports motion capture of a weapon through 3 markers placed on a demilitarized M16A2 rifle. The soldier can pick up the rifle and position it within the real-world gaming area, and his movements are captured in the "virtual" world exactly as they are acted out. The trigger pull is captured through a wireless system which provides a signal to the host computer. Fire and Detonation PDUs are generated and sent out over the DIS network based on the position and orientation of the M16A2. Preliminary tests indicate that the firing accuracy is very high. The soldier is able to discern and hit a man-sized target at a range of approximately 100 meters.

3. EXPERIMENT RESULTS

3.1. Gaming Area Size

The motion capture system for the DSS system initially was configured for an effective gaming area of 6' x 6' (error ±0.1 inches). This area proved too small to realistically depict the motions of a dismounted soldier. By changing the lenses to wider angle lenses, increasing the lights from 750 watts to 1000 watts, and making the motion capture markers larger, the gaming area was increased to 12' x 12' (error ±0.2 inches).

As noted, as the gaming area is increased, tracking errors and inaccuracies increase as well. These errors are caused by the non-linearity of the camera lenses (i.e., the wider angle lenses used the greater the distortion from the edges of the lenses) and the introduction of false markers into the gaming area (i.e., external reflective objects and the lights themselves).

If a requirement exists for the gaming area to be any larger than the current dimensions, multiple sets of camera/light pairs can be configured in place of one set of four camera/light pairs. Using neighboring multiple sets, the dismounted soldier could walk from one gaming area into another, thereby increasing the gaming area considerably. Also, the addition of one more camera/light pair (with minor software changes) would increase the gaming area somewhat and also decrease the point error.

3.2. Accuracy

The accuracy of the DSS system is represented by two metrics: the transport delay (TD), or latency, and the motion capture mean error (MCME). The TD is the time between the live dismounted soldier moving and the time that the corresponding computer generated model makes the same movement. The average TD measured is approximately 1/15 second. The TD is measured by comparing the time when the "real" soldier moved to the time when the "virtual" soldier moved. Included in this delay is the approximate 1/30 second screen refresh rate. The MCME is the difference in position that the four cameras measure the markers. The MCME is currently 0.2 inches. The TD can vary slightly based on where the "virtual" soldier is positioned on the database. The MCME can vary slightly based on where the "real" soldier is positioned in the gaming area. If the TD becomes too large, the animation looks choppy. If the MCME becomes too large, visible errors in computer image generation occur and the system needs to be restarted.

3.3. Calibration

The DSS system requires two forms of calibration. The first calibration is performed when the system is set up (i.e., cameras mounted, lights positioned) and when the system goes out of calibration (typically once per month). This requires precise measurement of a set of 32 markers in known positions which takes approximately four hours to complete.

The second calibration is performed when the dismounted soldier is ready to begin moving in the virtual environment at the start of an exercise. This requires the soldier to assume a calibration position while the computer "memorizes" which markers are which. This takes approximately 5 seconds. This calibration also needs to be performed whenever the soldier goes out of the gaming area or whenever an error in motion capture occurs.

3.4. Discrete States

The discrete states which the DSS system recognizes are: Running, Walking, Crawling, Rolling, Jumping, Kneeling, Prone, and Standing. The states for the weapon are: Held, Fired, and Dropped. States are determined by an algorithm which compares specific human parameters such as waist height, velocity, and angular velocity to empirically derived values for the particular state.

3.5. Dead Reckoning

Real entity approximation, or dead reckoning, is performed on the soldier's positional data as sent across in the Entity State PDU. Dead reckoning on the individual human joint angles was not performed as it represents too large an effort for the initial phase and its benefits are unclear due to the fact that human motion is erratic and therefore unpredictable. Moshell and Cortes [MOSH 95] also confirm that dead reckoning for human figure motion does not alleviate network loading.

The Entity State PDUs transmitted by the DSS are dead reckoned in accordance with the DIS algorithms

1-9. The dead reckoning algorithm to be used can be selected by the user in the graphical user interface. The DIS PDUs received by DSS are dead reckoned in accordance with the DIS algorithm #2 (1^{st} order in position).

3.6. Transmit List

In order to minimize bandwidth utilization, the operator of the DSS system is provided with the option to only send portions of the motion data over the DIS network. For instance, the operator can select to only send upper body rotations in the Data PDU if lower body movement is not required. Experiments can also be performed to simulate body symmetry by using only the left half data and reflecting this to the right side.

3.7. Tolerance Thresholds

Tolerance thresholds were implemented as a form of network optimizing. Thresholds, ε, can be assigned to the dismounted soldier's joint data. If the change in joint angle is less than ε, the human figure's model is not updated. Values of $0.1° \leq \varepsilon \leq 5.0°$ were tested, giving various levels of smoothness. With this option, if the soldier is not moving, no Data PDUs are sent out over the DIS network. Only when the soldier moves his arms or legs greater than a certain threshold will these values be sent over the network. If bandwidth problems occur, this method works well.

3.8. Simulator Sickness

No significant levels of simulator sickness have been indicated by helmet mounted display usage despite the 1/15 second transport delay from motion to viewport. Most of the time, the soldier is looking straight ahead, where the lag is not distinguishable. The longest that a subject has worn the helmet mounted display is approximately four hours without any noticeable sickness.

3.9. Statistics

Network statistics are available from the DSS system for both outgoing and incoming PDU traffic. The DSS system records the current rate at which outgoing Entity State, Data, Fire, and Detonation PDUs are transmitted. The DSS system also records the current rate at which incoming Entity State and Data PDUs are received from the DIS network. In the DSS lab environment with one DSS system, and a ModSAF station and an IST CGF station generating up to ten entities, no network traffic load problems were encountered.

4. FUTURE PLANS

1. Sensors

   Sensors such as night vision goggles, wrist watch, compass, and global positioning satellite (GPS) receiver will be modeled. These will enable the DSS to simulate night operations and enable the real soldier to navigate over the synthetic environment.

2. Alternative Tracking Methods

   Currently, the DSS utilized CCD camera motion capturing technology. High intensity lights are required to illuminate the reflective markers on the soldier to pick up position information. These lights can be bothersome in a simulated environment which requires night-training. Alternative tracking methods such as infrared capture technology are being considered.

3. Additional Soldiers

   Currently there is support for a single soldier for each DSS. The motion capturing of the M16A2 weapon already demonstrates the capability to track two independent motion bases in the gaming area. With additional computing and graphics power, one or more additional soldiers could be added to the system quite easily.

4. Gesture Recognition

   Currently the DSS recognizes the soldier's body movements (walking, crawling, running, etc.). However, if the system could recognize arm movements through gesture recognition, the soldier could react to signals from troop leaders. Or conversely, the soldier could give signals to computer generated DI's which would react to the signals through some means of artificial intelligence. The soldier's competency in giving or following gestures could also be measured.

5. Increased Gaming Area

   The size of the gaming area is currently limited to 12'x12'. This is a function of the size of the room in which the DSS system is placed. The theoretical size of the gaming area could be increased until the accuracy was such that the motion capture was no longer consistently accurate, say to a size of 20'x20'. If even a larger gaming area was desired, multiple rooms could be outfitted with DSS capture systems. As the soldier transitions from one room to the next, one set of cameras would hand over the tracking to another set of cameras. In this way, an entire building or large field could be modeled without the need to use "artificial" movement techniques such as treadmills, unicycles, foot pedals, or forced walk/run techniques.

6. Additional Weapons

   Light weapons such as the AT4 and JAVELIN anti-tank weapon will be modeled with the appropriate ballistics and munitions fly-out data. Of primary concern is the study of the loss in accuracy of target acquisition due to DSS transport delay and marker accuracy as this has a direct effect on negative training.

7. Peripheral Devices such as Omni-Directional Treadmill

   The DSS currently offers no way for the real soldier to move over large distances by foot with accurate resistance felt from ground topography. Devices such as an omni-directional treadmill may provide a solution to the problem.

5. BIBLIOGRAPHY

Enumeration and Bit-encoded Values for Use with IEEE 1278.1-1995, Standard for Distributed Interactive Simulation - Application Protocols, IST-CR-95-05.

Jones, Traci, Dismounted Infantry (DI) Technology Transfer Plan, U.S. Army STRICOM, 28 November 1995.

Klasky, R., Anschuetz, R., Molnar, J., Jones, T., Dismounted Soldier System: The First Virtual Soldier Simulator, 14th DIS Workshop, vol. 1 - Position Papers, 11-15 March 1996, pp. 43+.

Moshell, JM, and Cortes, A, Human Figures and the DIS Protocol: An Analysis, 13th DIS Workshop, vol. 1 - Position Papers, 18-22 Sept 1995, pp. 763+.

Moshell, JM, and Cortes, A, Individual Combat Simulation Systems and the DIS Protocol: An Analysis, JMM95.30, 22 December 1995.

Nemzow, M, The Ethernet Management Guide, McGraw-Hill, New York, 1995, p. 66.

Porter, CD, and O'Keefe, JA, Development of a Dismounted Infantry Human PDU, INCOMMSS-94.

Standard for Distributed Interactive Simulation - Application Protocols, Version 2.0, 4th Draft, Institute for Simulation and Training, 16 March 1994.

Zyda, MJ, Pratt, DR, Pratt, S, Barham, P, and Falby, JS, NPSNET-Human: Inserting the Human into the Networked Synthetic Environment, 13th DIS Workshop, vol. 1 - Position Papers, 18-22 Sept. 1995, pp. 103+.