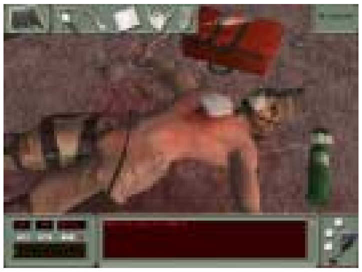
The Technology Behind Full Body 3D Patients

James Zimmer, Paul Kizakevich, Jerry Heneghan, Henry Schwetzke, Stephen Duncan
*RTI International, Research Triangle Park, NC
www.rti.org/vr

Introduction:



Over the last decade, there has been substantial emphasis in creating interactive, virtual medical training applications for surgery and other medical procedures. However, other than instrumented manikins, there has been less emphasis on simulation of the patient as a whole person.

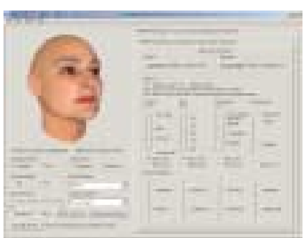
Medical professionals still train primarily through the apprenticeship model of transferring knowledge from experienced professionals to young medical students. Diagnostic, social, and procedural skills are acquired primarily through practical experience with actual patients or live patient actors. This learning model could be accelerated and modernized through the introduction of 3D virtual patients on standard PC computers.

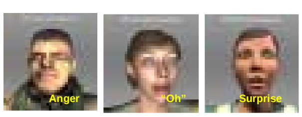
The goal was to develop virtual patients with medical, physical, and psychological disorders as an effective alternative for practicing assessment, diagnosis, treatment, and interpersonal skills. With the introduction of 3D interactive virtual patients, medical students can increase the amount of practical case experience they receive before working with actual patients. Likewise, practicing professionals could use virtual patients in their continuing education requirements and to demonstrate competency in critical areas of patient care. Users can develop their skills, sharpen their decision-making, and increase their exposure to a greater range of medical scenarios. In this presentation we describe RTI's Sim-Patient™ technology of full body 3D patients for interactive medical training.

Enabling Technologies:

Modeling Character Bodies

Visually realistic human characters depend on well designed body meshes and face meshes. Character bodies are modeled using multiple meshes representing the nude form and combinations of clothed and unclothed anatomical segments. Edge-loop techniques are employed which allow for clean deformation of each mesh as the character animates. Clothing layers can be removed to reveal the underlying body segment for local examination of abnormalities such as wounds, swelling, rashes, and pallor.





Head Modeling and Lip Syncing

Character heads are designed using parametric head generation tools, then fine-tuned and enhanced via artistic methods. Each head contains a tongue and a complete set of teeth that allow for a full range of expressions and mouth shapes. For each expression and mouth shape (i.e., phoneme), a new version of the head is created and tagged.

Speech is recorded using professional voice talent. The resulting sound files are then analyzed with signal processing software, and a file is created that identifies which phonemes the character should be exhibiting at any point in time. If recorded speech is unavailable for a given phrase, a text-to-speech engine is used to generate speech on the fly.

At run time, custom rendering software takes these various versions of the head and blends them together to form smooth facial expressions as the character speaks.

Commercial 3D Game Engines

The video game industry has accelerated many of the technologies that make 3D training on a PC possible. In the competitive market of video games, game developers have learned to optimize the 3D performance of their games while maximizing the visual experience of the user. Many of these companies derive secondary revenue by licensing their base technology to third Parties.

To support 3D patient simulation, we employed a commercially-available 3D game engine (Renderware®, Criterion Ltd.) as our graphics foundation. Like many engines this package has support for popular 3D modeling packages like 3D Studio Max, character animation capabilities, and many other features useful in depicting realistic virtual humans. A meta layer was developed to match the intrinsic features of the game engine with our 3D graphics requirements and enhance its functionality for virtual patient simulation.

Image Based Region Selection

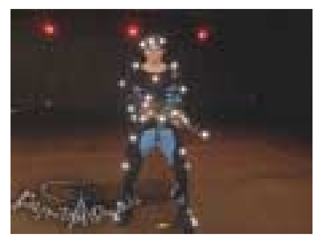
Full-body character meshes are often designed to facilitate smooth motion playback and efficient rendering speed. Therefore, the polygons that make up whole-body meshes do not always coincide with a unique region of human anatomy. A unique approach needed to be devised to support selection of macro anatomy (e.g., forearm) as well as smaller specific anatomy (i.e., antecubital vein) and regions (e.g., cardiac auscultation sites).

Our solution employs a set of color texture images which represent various anatomical and medical regions. These texture images are created using a 3D paint application. Each uniquely colored region represents a specific area of the body. By using multiple maps, the system can use one map to represent gross regions such as individual finger joints, and another map to represent stethoscope auscultation sites. The rendering system never displays these maps on the screen. Instead the maps are only used internally by our simulation code.

Motion Capture

Body language and motion are strong visual behavioral attributes of the real human. For whole-body patient simulation, realistic visualization of voluntary and involuntary motion is necessary to portray clinical signs such as vomiting, convulsions, and distress. To simulate such motion, both hand-created and motion data "captured" directly from a live human actor are used. Motion capture techniques take advantage of a range of sensor, optical, and computational technologies.

For this work, data was acquired using an optical system which tracks visible patches on the human actor's body. A professional motion-capture studio was employed to acquire and pre-process the raw motion data. Raw data often contains anomalies, like feet going through floors, so having the studio edit the data produced clean data files and expedited our production.

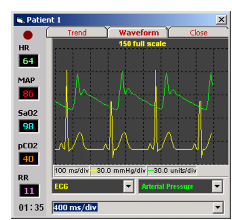


Natural Language Processing

To converse with the virtual patient, a proprietary Natural Language Processor, (RTI's AVATALK™) is used to acquire and analyze free-speech input for contextual meaning. Off-the-shelf voice recognition is used to receive a string of words, then performs a contextual grammar analysis to assign meaning to the phrase. Such speech input is generally used to elicit patient information (e.g., symptoms or history), direct patient activity (e.g., "raise your arm"), or inform the patient to affect compliance (i.e., "let me look in your ear").

Physiology Engine

Many of our virtual patients use the a real-time physiology engine with multiple modeling and transport modeling to simulate gases, drugs, and other substances circulated throughout the body. The engine integrates a cardiovascular model, a respiratory gas exchange model, drug effect models, pharmacokinetic models, and real-time waveforms. The physiology engine (based on the BODY model, K. Starko and Ty Smith) was extended to include trauma conditions (e.g., cardiac tamponade) and to support multiple patient simulation within a scenario.



Key Accomplishments

The following innovations have resulted from this research on full-body virtual simulated patients:

- 3D characters with medically-relevant animations
- Macro and micro body region selection
- Medical device models with functional simulation
- Methods for attaching medical devices to animated characters
- Enhanced physiology models supporting multiple patient simulation
- Synchronized human lip motion with either text to speech or recorded speech production

This work was supported in part by the U.S. Army Medical Research and Materiel Command, the Agency for Health Care Research and Quality, and the National Science Foundation. The conclusions do not necessarily reflect the position or the policy of the government,and no official endorsement should be inferred.