M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 390
Submitted to: ACM Siggraph Symposium on Interactive 3D Graphics, Providence, Rhode Island, April 1997
Obsolete by: TR426

Using A Virtual Environment to Teach Cancer Patients T'ai Chi, Relaxation and Self-Imagery

David A. Becker   Alex Pentland

Perceptual Computing Group – MIT Media Laboratory
20 Ames St. – Cambridge, MA 02139
dbecker | sandy@media.mit.edu

Abstract

We present a computer vision and virtual reality application for cancer patients. Currently, visualization and imagery techniques are becoming an accepted adjunct in the treatment of cancer. Mounting evidence shows the effectiveness of using self-imagery to encourage the immune system to boost its effort to defend the body from disease. Furthermore, while stress has been shown to retard the immune system, relaxation has the opposite effect. To that end, we are developing a virtual reality visualization tool for cancer patients called *Staying Alive*. The system provides a means for users to relax while directly visualizing their immune systems fighting off their diseases. White and red blood cells and malignant cells populate the virtual environment. The user navigates a white blood cell through the blood stream and "digests" malignant cells found along the way. To control the virtual environment, the user, free of wires and other such encumberments, engages the system simply by sitting in a room, practicing *T'ai Chi* gestures. Computer vision techniques are employed to track the user's head and hands in 3-D and a Hidden Markov Model framework is used to recognize the gestures in real time.

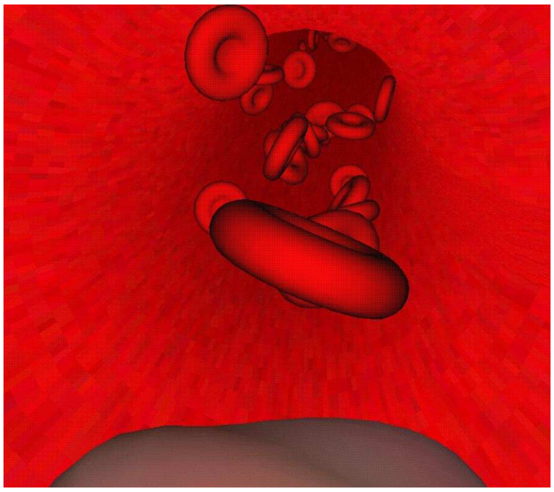

Figure 1: The virtual blood stream environment

1 Introduction

The benefits of visualization techniques have been noticed by such disparate groups as athletes and cancer patients. Athletes are told that they must picture themselves hitting the ball, making the catch, scoring the goal to be able to do so. Similarly, cancer patients around the nation can visit almost any cancer center and enroll in programs to learn how to visualize their own immune system fighting off their diseases. Though the full extent of the mind-body connection is unclear, there is a growing belief that visualizing something in the mind encourages the body to make it reality [8]. In fact, Little claims that he sees a fifty-six percent increase in long-term survival rates for cancer patients who participate in his program of visualization, relaxation and positive thinking [5].

Similarly, relaxation appears to have a beneficial effect on the immune system. Numerous studies have shown that inducing the "relaxation response" enhances the effectiveness of the immune system while stress retards it [9]. Furthermore, imagery and relaxation have been shown to greatly reduce the stress and anxiety that cancer treatments can cause [4, 2].

Because visualization and relaxation promise so much benefit to both those with and without cancer,

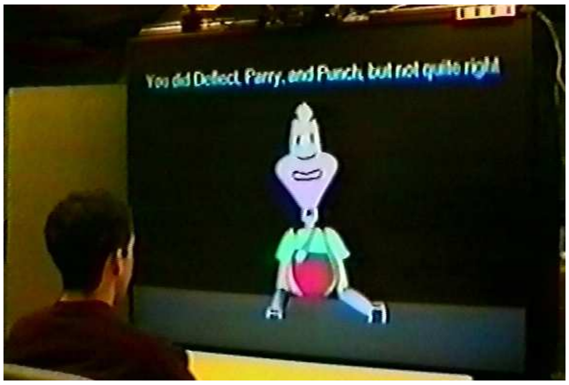

Figure 2: User Interacting with Virtual Teacher

we are developing a virtual reality visualization tool called *Staying Alive*. If simply picturing something in the mind's eye is effective, experiencing the same thing in a virtual reality environment might be even more beneficial. Not only does a virtual reality system allow users to directly visualize the virtual world, but it also allows users to take an active role in influencing that world.

To help make *Staying Alive* a relaxing application, we have chosen to make a user's interaction with the system be through the practice of *T'ai Chi* (*taijiquan*). *T'ai Chi* is among the more popular forms of martial arts in China and the world. Though like all martial arts, this exercise includes movements for attack and defense, there are several forms of *T'ai Chi* which are slow and gentle, such as the Yang, Wu or Sun styles [14]. The gestures for this application were chosen from the Yang style, which is meant more for relaxation and wholeness than for fighting.

2 The application: Staying Alive

The test-bed in which this project is being developed is the *Interactive Virtual Environment (IVE)* space. The *IVE* space was developed based on the idea that a user should be able to interact with a virtual world in a natural way, free from goggles, wires, data gloves, or body suits. In an *IVE* space, a user is in front of some large display (such as a large projection screen). One or more cameras are mounted in the room, and the video from the cameras can be blended with output from an SGI workstation and displayed on the screen. Users can see themselves "in" the virtual world on the screen. In addition, the input from the cameras is digitized and processed on one or more separate workstations.

The paradigm for *Staying Alive* is that a user enters into an *IVE* space and faces the screen. The user sees a virtual world of her bloodstream, with herself as a white blood cell in the blood stream, as seen in Figure 1. Red blood cells zip by as the user moves along until a malignant cell is found. The task of the user is now to navigate the white blood cell to the malignant cell and gently destroy it by engulfing it or calling other cells in to contain it.

To control the white blood cell, the user performs *T'ai Chi* gestures. The gestures are recognized, and different gestures are mapped to different actions of the white blood cell (movement, various means for removing malignant cells). The user sits in the *IVE* space and practices *T'ai Chi*, thereby having the virtual immune system fight off the disease. The whole experience is a relaxing one for the user. It is not a difficult task to control the disease cells, so that the user is not stressed by missing any malignant cells.

Since not everyone is familiar with *taijiquan*, another aspect of the project is a *T'ai Chi* "teacher." If a user performs a gesture incorrectly, a virtual teacher appears on the the screen and shows the user what he did and where his form was not correct or could use improvement. Then the teacher acts out the gesture in the correct form. Figure 2 shows a user interacting with the teacher.

3 Implementation

The main tools for implementing this project are virtual reality, gesture understanding, and user training and teaching.

3.1 Virtual Reality

The virtual bloodstream software was developed with *Inventor* and is run on an SGI Onyx. The environment is a 3-D world with simple physics governing the behavior of the blood cells.

3.2 Gesture Understanding

In order for each *T'ai Chi* gesture to convey a different meaning in the virtual world, the system must be able to recognize when the user is performing a particular movement. This interpretation of human action is a goal shared by a large number of applications in the *IVE* space. Among them is the wireless virtual reality system, *ALIVE* [6]. In order for the "dog" in this system, Silas T. Dog, to understand the intentions of the human user, a real-time system called pfinder (person-finder) was developed [12] that tracks the user's head, hands and body. While the user's distance from the camera and x-y coordinates are determined by pfinder, there is no true, three-dimensional representation of the hands (only a two and a half dimensional representation is inferred). To understand what the user is trying

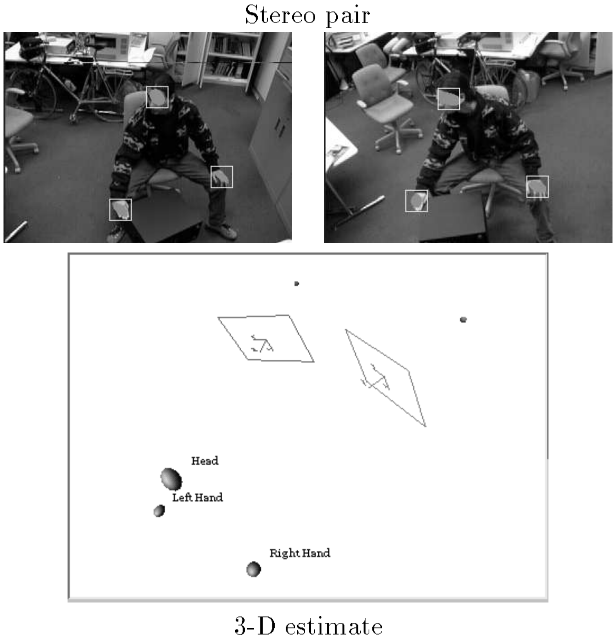

Figure 3: Real-time estimation of the position of moving human head and hands. The top images are the video and tracking from the two cameras and the bottom image is the result of triangulating

to convey, a simple gesture understanding module takes, among other things, the locations of the hands and head as given by pfinder.

A richer representation of the hands, one that includes a true third dimension, allows for the recognition of more sophisticated gestures, such as *T'ai Chi* gestures. Azarbayejani and Pentland have developed *STIVE*, a system similar to *ALIVE* but with two wide baseline cameras, allowing for less noisy depth determination [1]. *STIVE* detects human skin by chrominance value (unaffected by variations in melanin density) and can output 3D head and hand positions to an accuracy of approximately two centimeters at up to 30 frames/sec (see Figure 3).

In *Staying Alive*, the *STIVE* system is used to generate the three-dimensional locations of the head and hands. A Hidden Markov Model framework is used to interpret the gestures. Though HMM's have predominately been used by the speech and handwriting recognition communities, they are currently finding acceptance in computer vision research. Especially for work in which the recovery of the complete three-dimensional structure of a scene is not deemed necessary, HMM's are increasingly being used. Significant successes have been obtained in gesture recognition with a more simple view-based approach for which HMM's are appropriate [10, 11, 13]. Starner achieved a recognition rate of ninety-seven percent (on a forty word lexicon) using a two-dimensional representation of the hands.

In *Staying Alive*, eighteen different *T'ai Chi* movements have meaning in the virtual world. A pictorial description of some of the gestures can be found in Figure 4.

Five state, forward chaining HMM's (with skip states) are used to model the gestures. As can be seen from Figure 5, the HMM can model gestures with 3, 4, or 5 states. In addition, a two state HMM is employed to model "silence," which is used to reset the system if desired. A single user performed each gesture eighteen times to provide the training. For testing, six examples of each gesture were recorded in the same location as the training, in a translated location, and in a translated and rotated location.

It is not in the scope of this paper to fully discuss the selection of feature vectors (see [3] for such analysis). However, using the (dx,dy,dz) (delta between frames) locations of the hands and head provides a ninety-eight percent recognition level when the user practices *T'ai Chi* in the same location as where the training took place. A body-centered polar coordinate system is used to allow for invariance to translations and rotations, with accuracy of ninety-three percent on such data. The user can generally be assumed to be facing the screen, making rotations unlikely. Currently, a recognition rate of ninety-six percent is achieved on translated *T'ai Chi* using the polar coordinates. The Viterbi algorithm is used to perform the recognition, which occurs in real-time.

3.3 *T'ai Chi* Teacher

The goal of the *T'ai Chi* teacher is to evaluate the user's performance of *T'ai Chi* and provide constructive feedback in order to teach the user better form.

The first part of this task is accomplished with the results from the Viterbi recognition. The Viterbi algorithm returns both the most likely gesture performed and also a log likelihood for that gesture. This probability is compared with two thresholds that each gesture is assigned: a threshold above which the system assumes that the user is performing the gesture accurately and doesn't need feedback and another threshold below which the system cannot confidently recognize the gesture. If the recognition probability falls between these two thresholds, the system assumes the user is attempting to perform the given gesture, but that the user is inaccurate enough in the motion that she needs feedback.

When this situation occurs, the *T'ai Chi* teacher is invoked. The teacher takes the motion that the user

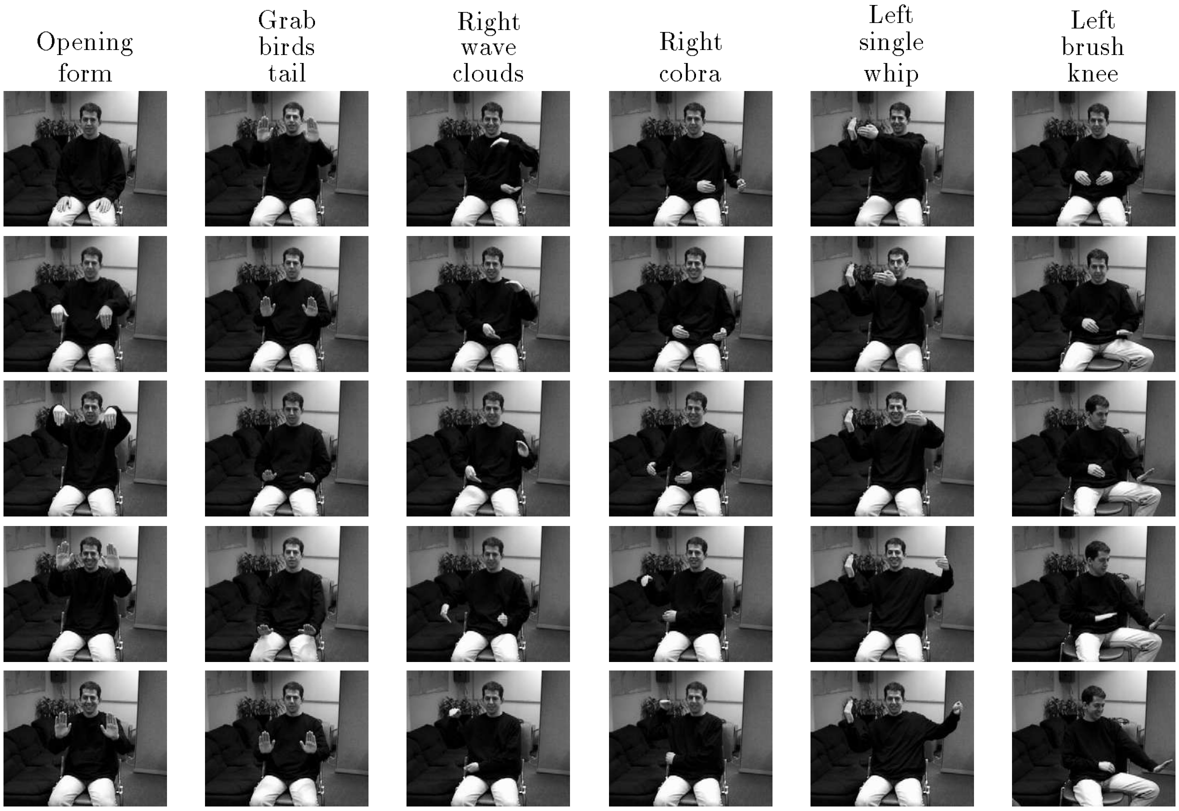

Figure 4: Six of the eighteen T'ai Chi gestures.

just performed and compares it with an idealized version of the intended gesture. The idealized version is obtained by time-warping and averaging all the training examples of the gesture. Then, the user's motion is similarly time-warped (unless the time of the gesture clearly falls outside reasonable bounds) and compared to the idealized version. The teacher then plays back what the user did and shows the user the segment during which the motion most differed from the expected motion for the gesture. After that, the teacher then acts out the idealized motion of the gesture, and the application continues.

4 Future Work

The hand-tracking system, *STIVE* is capable of capturing position, orientation and shape of the hands and head. *Staying Alive* currently only uses the position. With the orientation and shape, recognition of *T'ai Chi* gestures, in which hand orientation is important, should improve.

Eventually, an interesting experiment will be to determine the actual benefit of the system on cancer patients. Patients using the system could be compared to those doing more standard visualization techniques to see if either the quality of life improved, presumably from reduced stress and anxiety, or if the length of remission increased.

In addition, another interesting extension of this work would be the integration of affective computing, "computing that relates to, arises from, or deliberately influences emotion" [7]. Specifically, Picard discusses the development of emotion recognition by computers, most simply through the use of sensors monitoring physiological responses such as heart rate or blood pressure. There are many issues which complicate the task. For example, similar physiological responses can arise through influences other than emotion (such as exercise) [7]. However, given a system which is capable of recognizing stress, a fascinating addition to the application would be the ability of the system to make the environment more soothing if the user were not becoming more relaxed.

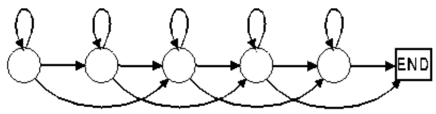

Figure 5: The five state HMM used for recognition

5 Conclusion

There has been a recent explosion of evidence supporting the idea that visualization, imagery and relaxation bolster the immune system. A virtual reality application providing the tools for practicing these activities could be a real benefit to cancer patients. Not only does it allow for direct visualization, but it also allows a user to easily control the environment at a time when patients are facing a loss of control in their lives.

We believe that this direction for virtual reality applications is a promising one. Furthermore, we believe the type of user feedback provided by the *T'ai Chi* teacher is a constructive way to help people learn certain tasks.

6 Acknowledgments

Special thanks to Dr. Gary Ratkin for valuable guidance in the medical issues. Thanks also to Lee Campbell and Ali Azarbayejani for assistance with the gesture recognizer, the figures, and the text of this paper. Marc Lebovitz was exceptional in developing the virtual bloodstream.

7 References

[1] Ali Azarbayejani and Alex Pentland. Real-time self-calibrating stereo person tracking using 3-D shape estimation from blob features. In *Proceedings of 13th ICPR*, 1996.

[2] H. C. Salzberg C. S. Feldman. The role of imagery in the hypnotic treatment of adverse reactions to cancer therapy. *J. S. C. Medical Association*, May 1990.

[3] Lee W. Campbell, David A. Becker, Ali Azarbayejani, Aaron Bobick, and Alex Pentland. Features for gesture recognition using realtime 3d blob tracking. MIT Media Lab Perceptual Computing Group Technical Report No. 379, Massachusetts Institute of Technology, 1996. Submitted to 2nd International Face and Gesture Workshop, Killington VT.

[4] T. W. Decker, J. Cline-Elsen, and M. Gallagher. Relaxation therapy as an adjunct in radiation oncology. *Journal of Clinical Psychology*, May 1992.

[5] Bill Little. *Eight Ways to Take an Active Role in Your Health.* Harold Shaw Publishing, 1994.

[6] P. Maes, B. Blumberg, T. Darrell, and A. Pentland. The ALIVE system: full-body interaction with animated autonomous agents. In *ACM Multimedia Systems*, in press.

[7] R. W. Picard. Affective computing. MIT Media Lab Perceptual Computing Group Technical Report No. 321, Massachusetts Institute of Technology, 1995.

[8] Bernie Siegal. *Love, Medicine, and Miracles: lessons learned about self-healing from a surgeon's experience with exceptional patients.* Harper and Row, 1986.

[9] S. E. Sims. Relaxation training as a technique for helping patients cope with the experience of cancer: a selectivereview of the literature. *Journal of Advanced Nursing*, Sep 1987.

[10] T. E. Starner and A. Pentland. Visual recognition of american sign language using hidden markov models. In *Proc. of the Intl. Workshop on Automatic Face- and Gesture-Recognition*, Zurich, 1995.

[11] A. D. Wilson and A. F. Bobick. Learning visual behavior for gesture analysis. In *Proc. IEEE Int'l. Symp. on Comp. Vis.*, Coral Gables, Florida, November 1995.

[12] Christopher Wren, Ali Azarbayejani, Trevor Darrell, and Alex Pentland. Pfinder: Real-time tracking of the human body. *SPIE Conference on Integration Issues in Large Commercial Media Delivery Systems*, Oct 1995.

[13] J. Yamato, J. Ohya, and K. Ishii. Recognizing human action in time-sequential images using hidden markov model. *Proc. Comp. Vis. and Pattern Rec.*, pages 379–385, 1992.

[14] Shing Yen-Ling. *T'ai Chi Ch'uan: The Basic Exercises.* Sugawara Martial Arts, 1990.