Digital Circus: A computer-vision based interactive Virtual Studio

Flavia Sparacino, Christopher Wren, Glorianna Davenport, Alex Pentland

MIT Media Laboratory
20 Ames Street; Cambridge MA 02139 USA
{flavia, cwren, gid, sandy}@media.mit.edu

Abstract

We have built a low-cost virtual studio which does not need a blue screen to operate. Hence, such a setup can be used at home and allows to distribute the production outside of the physical TV studio. Our studio is based on Pfinder: a real-time computer-vision body tracking and gesture recognition system. We use Pfinder also to enable interaction among the composited participants, and the objects and creatures in the virtual setting. To demonstrate our system, we have built a digital circus in which multiple individual participants can connect from remote locations, see all their images composited in the virtual circus, and interact with the objects on the virtual set. Other possible applications range from full body real-time teleconferencing to remote collaborative work, networked performance, entertainment, and education.

1 Introduction

Virtual Studios are new forms of TV video productions which combine real foreground images, shot in the studio, with 3D computer-generated background scenes. Current Virtual Studio systems need: cameras shooting the foreground action against a blue background; tracking to provide camera position information; rendering to generate images with correct perspective in the virtual scene; and z-mixing to layer background and foreground correctly with respect to depth information.

Research in Virtual Studios is tightly related with work in building immersive and interactive Virtual Environments in which people can interact with virtual objects, animated 3D creatures, or among themselves. Ultimately these two venues of research converge towards the creation of low-cost, real-time, Virtual Studios in which the 3D set is an interactive virtual environment, and the compositing of multiple participants in the same set opens up new possibilities for TV production, such as game playing, collaborative storytelling, or immersive telepresence.

By using real-time computer vision techniques we have built a low-cost virtual studio which does not need a blue screen to operate. Hence such a setup can be used at home and allows to distribute the production outside of the physical TV studio. In addition, we use computer-vision based real-time gesture recognition which enables interaction among the composited participants and the objects and creatures in the virtual setting.

Our virtual studio is based on a real-time computer vision system called "Pfinder," i.e. "person finder." Pfinder is a system for body tracking and interpretation of movement of a single participant. It uses only a wide angle camera pointing towards the foreground action space—

the home virtual studio—and a standard Silicon Graphics O2 computer.

Although real-time interpretation of the participant's actions is essential for interactivity, authoring a timely, appropriate, and coordinated response by the variety of virtual objects on the set can be a complex task. In order to facilitate authoring and to ensure a timely interaction, we have endowed the virtual objects and creatures that inhabit the 3D studio set with autonomous responses and behaviors. By using a mixture of purely responsive and behavior-based AI techniques we are able to distribute the authoring complexity from a central program to the various objects/characters in the set. Objects or creatures in the set can respond appropriately according to the context of interaction without having to script in advance all the possible combinatorics of time/action sequences on the virtual set.

Finally, to demonstrate our system, we have built a digital circus as our 3D set. Multiple individual participants can connect from remote locations and see their images all composited in the virtual circus. Participants are engaged in a game of transformations which involves appearance/disappearance of objects, scaling the image of other participants very large or small, triggering events such as firing a cannon woman from a virtual cannon, or interacting with butterfly creatures which inhabit the set. All of these actions are interpreted and guided by the real-time gesture recognition and compositing features of our computer vision system as well as the responsive and behavioral authoring of the virtual set.

In the following sections we will describe in more details the elements of the home studio. We first provide an outline of how Pfinder works and what it does. We describe the motivations which led us to building a circus as our virtual set, and explain what participants can do in this new type of entertainment environment. We explain our approach in authoring the virtual studio as an interactive entertainment system. Finally we cite examples of related work and draw conclusions.

2 Pfinder

Pfinder uses a multi-class statistical model of color and shape to segment a person from a background video scene, and then to find and track people's body parts—head/hands/feet/center of body—in a wide range of viewing conditions (Figures: 1,2). It adopts a Maximum "A Posteriori" Probability approach to detection and tracking of the human body, using sets of 2D regions. It incorporates a priori knowledge about people primarily to bootstrap itself and to recover from errors. Its only constraints are a constant lighting and an unmoving background.

The process of detection and tracking of the human participant is guided by a 2D contour shape analysis which recognizes a body silhouette inside which single body parts can be reliably labeled. This silhouette is used to cut out the foreground video image of the person composited in the 3D set. This process, which achieves the compositing without blue screen, is done in two steps. First a support graphical silhouette is constructed inside the 3D environment and successively the corresponding video image is texture-mapped onto it.

Z-mixing is provided by giving Pfinder a camera model of the "home studio." The system back-projects the 2D image transformation to produce 3D position estimates using the assumption that a planar user is standing perpendicular to a planar floor. Back-projection ensures correct perspective as well as correct depth layering. Given that our virtual studio uses a fixed camera, we do not address the camera tracking issue.

Pfinder runs at a speed of 30Hz (standard video speed) and is able to track body features—head/hands/ feet/center of body—at about 20-24Hz. This allows us to recognize the participant's body postures, such as pointing towards an object in the virtual set, or sitting, as well as to feed an HMM-based gesture recognizer which can interpret more complex dynamic motions.

We will summarize here how Pfinder works. A more detailed description can be found in [1]. We can represent Pfinder's 2-D regions, also called blobs, by their low-order statistics. Clusters of 2-D points have 2-D spatial means and covariance matrices. The blob spatial statistics are described in terms of their second-order properties. The aggregate support map $s(x, y)$ over all the blob models represents the segmentation of the image into spatio-color classes.

Each blob has a spatial (x, y) and color (Y, U, V) component. Color is expressed in the YUV color space. We could additionally use motion and texture measurements as part of the blob descriptions, but current hardware has restricted us to use position and color only. Because of their different semantics, the spatial and color distributions are assumed to be independent. Each blob can also have a detailed representation of its shape and appearance, modeled as differences from the underlying blob statistics. The ability to efficiently compute compact representations of people's appearance is useful for low-bandwidth applications [2].

The statistics of each blob are recursively updated to combine information contained in the most recent image with knowledge contained in the current class statistics and the priors.

We assume that the majority of the time Pfinder will be processing a scene that consists of a relatively static situation such as a living room, and a single moving person. Consequently, it is appropriate to use different types of model for the scene and for the person.

We model the scene surrounding the human as a texture surface; each point on the texture surface is associated with a mean color value and a distribution about that mean.

One of the key outputs of Pfinder is an indication of which scene pixels are occluded by the human, and which are visible. This information is critical in low-bandwidth coding, and in the video/graphics compositing required for "augmented reality" applications.

In each frame, visible pixels have their statistics recursively updated using a simple adaptive filter. This allows us to compensate for changes in lighting and even for object movement. For instance, if a person moves a book it causes the texture map to change in both the locations where the book was, and where it now is. By tracking the person we can know that these areas, although changed, are still part of the texture model and thus update their statistics to the new value. The updating process is done recursively, and even large changes in illumination can be substantially compensated within two or three seconds.

Given a person model and a scene model, we can then acquire a new image, interpret it, and update the scene and person models. To accomplish this there are several steps: predict the appearance of the user in the new image using the current state of our model; for each image pixel and for each blob model, calculate the likelihood that the pixel is a member of the blob; resolve these pixel-by-pixel likelihoods into a support map; and update the statistical models all blob models.

Individual pixels are assigned to particular classes: either to the scene texture class or a foreground blob. A classification decision is made for each pixel by comparing the computed class membership likelihoods and choosing the

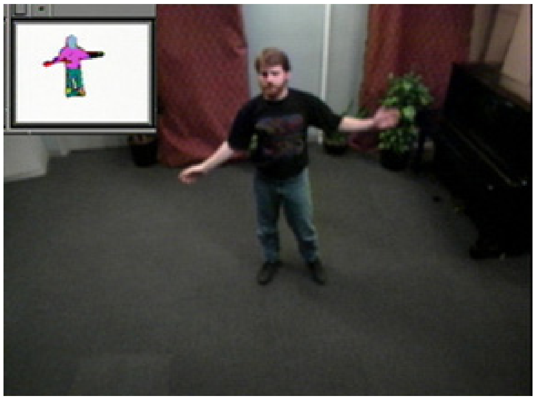

Figure 1: Pfinder tracking a person in the IVE studio

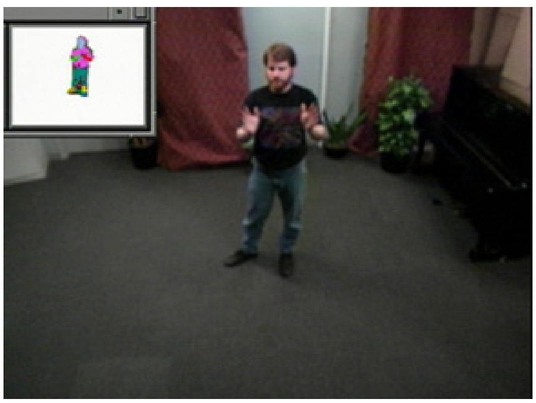

Figure 2: Notice correct hand tracking even when the hands are in front of the body

best one (in the MAP sense).

Connectivity constraints are enforced by iterative morphological "growing" from a single central point, to produce a single region that is guaranteed to be connected. The first step is to morphologically grow out a "foreground" region using a mixture density comprised of all of the blob classes. This defines a single connected region corresponding to all the parts of the user. Each of the individual blobs are then morphologically grown, with the constraint that they remain confined to the foreground region.

This results in a set of connected blobs that fill out the foreground region. By extracting the contour of of this foreground region we obtain the silhouette of the body in real time.

Pfinder is implemented on the SGI architecture using the VL (Video Library) interface. It works in real time on an SGI O2. For input we use a JVC-1280C, single CCD, color camera which provides an S-video signal to the SGI digitizers.

3 The Home Studio: the IVE Space

Our "home studio"—called IVE (Interactive Virtual Environment)—is a room sized area (15'x17') whose only requirements are good, constant lighting and an unmoving background. IVE is an interactive space developed at the MIT Media Lab, in which the public interacts with visual material presented on a large projection screen which occupies one side of the room (see Figure 3).

A downward pointing wide-angle video camera mounted on top of the screen allows the IVE system to track a player. By use of real-time computer vision techniques [3, 1, 4] we are able to interpret the participant's posture, gestures, identity, and movement. A phased array microphone is mounted above the display screen for audio pickup and speech processing. A narrow-angle camera housed on a pan-tilt head is also available for fine visual sensing. The only constraints are a constant lighting and an unmoving background. IVE was built to enable people to participate in immersive interactive experiences without wearing suits, head-mounted displays, gloves, or other gear. Remote sensing via cameras and microphones allows people to interact naturally and spontaneously with the material shown on the large projection screen. IVE currently supports one active person in the space and many observers on the side. We are in the process of extending the tracking technology to support many people at once. The

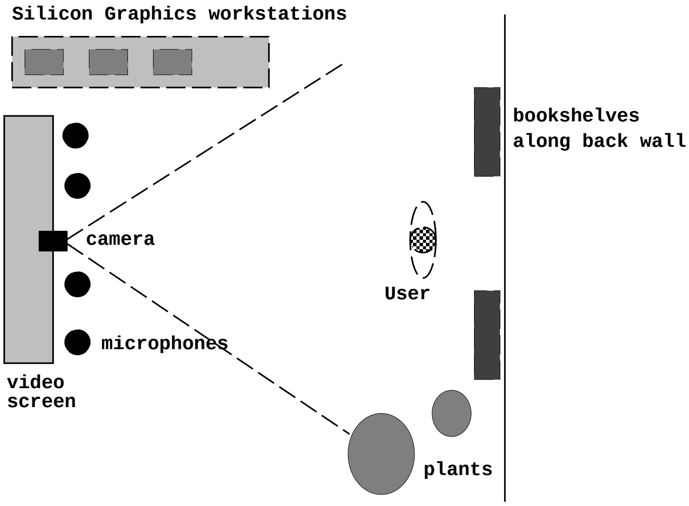

Figure 3: The IVE Space

IVE environment was originally developed for the ALIVE project [5] and has since become our main development platform for interactive experiences [6].

The ability to enter the home studio just by stepping into the sensing area is very important. The participants do not have to spend time "suiting up," cleaning the apparatus, or entangling wires. IVE was built for the more general purpose of enabling people to participate in immersive interactive experiences and performances without wearing suits, head-mounted displays, gloves, or other gear. Remote sensing via cameras and microphones allows people to interact naturally and spontaneously with the material shown on the large projection screen. The advantages of such a system are: ease of entry/exit, social context, maintenance, hygiene, greater safety, expandability.

Furthermore, social context is often important when using a virtual environment, whether it be for game playing or designing aircraft. In a head mounted display and glove environment, it is very difficult for a bystander to participate in the environment or offer advice on how to use the environment. With IVE, not only can the participant see and hear a bystander, the bystander can easily take the user's place for a few seconds to illustrate functionality or refine the work that the original user was creating.

Another advantage of IVE is that it does not involve special equipment. An entry level SGI Indy can be used to run vision, graphics, and sound. While it may be a few years before such a system can be economically produced for a "set-top box," the availability of the equipment from a major vendor reduces the cost of acquiring, maintaining, and upgrading the apparatus in the future.

4 Virtual Studios meet Entertainment

We use Pfinder's ability to do real-time compositing of the participant's video image in a 3D world, without the need for a blue screen, to build a networked application that engages participants from remote locations. We envision home based entertainment systems which allow for multiple people to be virtually present in the same 3D set. A presenter or coordinator would be guiding the show from any desired physical location. Participants can be given different interactive powers within the 3D set according to their role in the show or game. The TV coordinator can activate the main software server which does the rendering of the graphical virtual set and decides when, or which ones among the participants, can connect (as software clients) and be seen in the virtual set. Pfinder also provides the Z channel mixing information for correct compositing with respect of depth and occlusion inside the set.

While many interactive experiences and video games tend to focus on exploring new worlds and killing monsters, we have developed a virtual studio application based on the theme of transformations. Participants are engaged in a game of transformations which involves appearance/disappearance of objects, scaling the image of other participants very large or small, triggering events such as firing a cannon woman from a virtual cannon, or interacting with butterfly creatures which inhabit the set. All of these actions are interpreted and guided by the real-time gesture recognition and compositing features of our computer vision system as well as the responsive and behavioral authoring of the virtual set.

We believe it is important to develop, together with the new technology, new genres of collaborative experiences which offer room for exchange, communication, and maybe transformation. Digital Circus moves some steps towards a more educational and artistic genre of networked game playing.

4.1 Motivation

We see the circus as a communication process with its own language. In the circus information is transmitted mainly in a non-verbal, iconic and symbolic way. A circus performer can show his/her number in any country, without the need to know much about the language or the culture of that particular country. In this sense s/he is a transcultural performer. Moreover, although the circus is usually connected to childhood, the circus performer usually addresses (and enchants) everyone, without any distinction of age.

The language of the circus is essentially transformational. Objects have a meaning according to their role in the circus number. When a circus object appears, it brings a whole set of possible meanings. This set depends on the "encyclopedia" of the audience, while its actual meaning is chosen depending on its role in the circus number. An example of this "functional sliding" of the meaning of circus objects is given by the chair. It is a defense tool for a lion tamer; it can be a "character" or a musical instrument for a clown; it is a co-performer for a tight-rope walker that places it on a rope and then sits on it. When a chair is presented just as a chair, it is usually for horses and elephants to sit on it and functions as a humanizing factor for the animals.

The tradition of the circus is not an invariable repetition of the same tricks, but a set of cultural transformations rules. These rules are represented in a ritualized manner to disguise the "transgression" implicit in the message. The circus represents typical daily life situations. Nevertheless these situations—objects, values, beliefs, paradigms of the mainstream culture—are manipulated and modified by the circus rules. What is usually related is presented as unrelated. The cause precedes the effect. An elephant might use a phone, play a musical instrument, or eat a meal at the table, like a human. A clown produces sequences of inconsistent behavior. Similarly the basic rules of balance are challenged or denied.

The circus is a transgressive territory: transgression of laws of ordinary physics by the acrobats, (or better transgression of the laws of physics for ordinary people), transgression of politeness by the clowns, transgression of perception by the magicians, transgression of roles by the animals. Normally, lions are not mild, donkeys do not do arithmetic, ordinary people do not walk on tight-ropes, elephants do not eat at the table, objects/people do not disappear/appear or transform in an instant, people do not throw cakes at each other's face if they disagree. In this respect, cartoon is a genre close to the circus. There are a specific cartoon physics, cartoon sense of humor, cartoon description of a character. Cartoon characters are close to humans but not quite humans. They have magic powers, they follow law of physics that violate regular physics and usually interact in a way that would be judged violent or impolite in our regular way of interacting with other people.

This process of transformation, inversion, parody, and caricature of the mainstream culture and its objectives/values produces an ambivalent attitude towards the circus. A strong interest is often blended with a scornful smile. For this reason the circus is usually mostly appreciated by those individuals that haven't been completely integrated within a culture, but instead are marginal, like children, artists or poets. These are among the reasons that make of the circus an interesting playground for our research: its structural transformational nature and its elected audience. In this respect we believe it can appeal to an audience at least as large as that of the current consumers of videogames, if not larger. Its appeal would also be based on the fact that, along with Buissac [7], we see the circus not as a collection of numbers but as a set of transformation rules that can generate a potentially infinite series of numbers.

Once we have established that the circus communicates through its own gestural language, we still need to ask what is the narrative structure of a circus number and what is its relationship with the audience in the process of co-construction of meaning. Most circus numbers evolve through successive steps that bear close resemblance to the development of the folktale [7]. They are, with some variance: identification of the hero (performer); qualification trial; principal trial; glorification trial (usually preceded by drum playing); and final recognition from the public. We have so far established a grammar of interaction among participants and among participants and the virtual set. Our current work is aimed at endowing the digital circus with a more robust story architecture.

4.2 The Digital Circus

We created a digital circus composed by a basic architectural setting: a big top, a circus cannon, a chair, a gramophone, together with a series of magic objects that appear or disappear according to the needs of the participant. This digital environment is shown on a large screen that comprises one side of our home virtual studio. The audience experiences the circus by seeing their video image projected onto the virtual environment. All the participants' images are shown in the virtual environment.

We start our game with our basic transformational circus grammar by placing a chair in the real space. This chair projects symmetrically onto a multicolored virtual chair in the digital circus. When a person from the audience sits on the real chair she will see her video avatar sitting on the virtual chair. The system then recognizes the sitting gesture and it triggers "magic events." A gramophone suddenly appears and starts to play music. If the person stands up again the gramophone disappears and the music stops.

This intrinsic transformational nature of the circus makes the circus story similar to an "open-eyes dream." In particular the inversion of cause-effect relations, the "functional sliding" of objects, as well as the iconic-symbolic representation of most elements recalls the "logic of dreams." To convey this point, we have programmed a transformation rule into the system that pops up an umbrella when a member of the audience raises up her hand, as though she was actually holding an umbrella. Our system also recognizes a flying gesture that virtually lifts the audience up in the clouds, driving a digital plane. This operates literally as a 3D scene cut, that radically transforms the environment and drags the audience into another dream.

Two people in the circus can "play the clowns" with each other (Figure 4). The first who touches a "magic box" and then lifts or lowers his arms causes the other participant to grow very tall or very short. When short, the participant needs to hide so as not to be kicked by

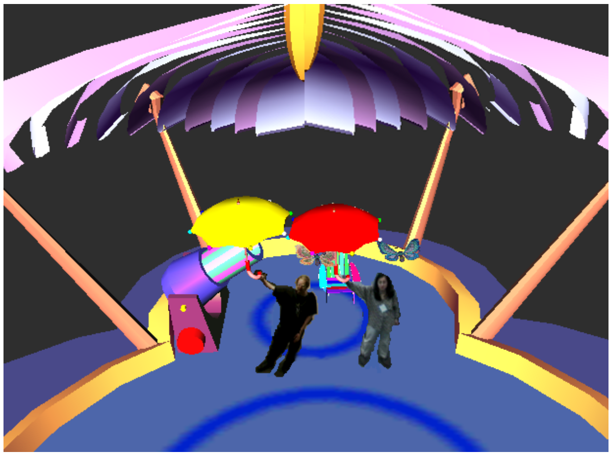

Figure 4: Two remotely communicating participants in the circus with their respective butterfly

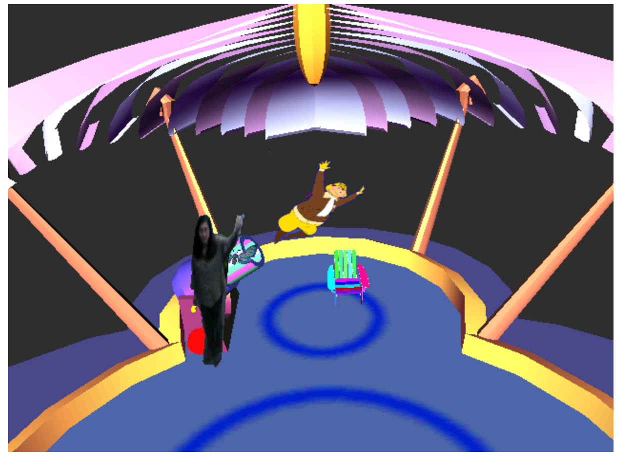

Figure 5: The circus virtual studio. Participant pointing towards the cannon woman

the other clown, otherwise s/he is transformed into a ball. The virtual butterfly can "save" the small clown and bring him/her back to the original size. If one participant is holding the umbrella, and the other touches the magic box, the umbrella lifts the person in the air. Touching the flying butterfly will bring the participant back to ground. Each participant has their own butterfly which helps them when needed. In order to gain more magic powers or to confuse the other player, a participant can create one or more clones of him/herself on the set. Clones are exact replica of the composited image of the participants, which move along with the "main" image of the person, but at a different position in the 3D set.

As of now, our circus has a main character who is a cannon woman(Figure 5). She is fired out of the circus cannon simply by pushing a button situated on the cannon. She allows us to reach a wider audience than the one physically gathered around the home virtual studio. She is fired from the cannon in the digital circus and lands on the WWW, on an applet than can be viewed by other networked participants. Through characters like the cannon woman, who are distributed across the different virtual networked home studios, we hope to gather an active audience of collaborating and engaged participants.

5 Authoring the Virtual Set

In order to create a compelling experience it is important not only to design a visually rich 3D set, but also to give participants the ability to modify and interact with the virtual objects or synthetic creatures which inhabit the set. Authoring an interactive game for multiple participants can be a very complex task, unless an appropriate methodology of modeling interaction is used. Scripted systems are those in which a central program rigidly associates people's actions or inputs with responses of the system by keeping a careful accounting of timing constraints and sequencing. Such systems are in general to avoid because authoring complexity grows very fast if a variety of interactive features are implemented. In addition they do not allow for spontaneous and unexpected actions or interactions to happen and are therefore inadequate for many real-life situations such as "live shows" or interactive games.

We have bypassed the complexity of scripted systems by building a responsive application which is defined by a series of couplings between the participant's input and the system's response. Responsive applications are easier to author because they do not have a central program which takes into account all the sequenced inputs, actions, and outputs. They define instead a geography of one-to-one mappings for which the participant's actions trigger specific system responses. Although easier to author, these systems do not take into account complex story structures and elaborate timing/sequencing information.

All the objects present in the circus are embedded with their own set of responses—the cannon, the umbrella, the chair, the gramophone, the cannon woman and even the set itself—participate autonomuosly in the game of transformations, appearances, and disappearances regulated by the participants' actions.

We have authored the virtual butterfly character instead by using a behavioral approach. Behavioral systems are those in which the response of the system is a function of the sensory input as well as its own internal state. The internal state is essentially a set of weights on the goals and motivations of the behavioral agent. The values of these weights determines the actual behavior of the agent. Behavioral systems provide a one-to-many type of mapping between the public's input and the system's response. The response to a particular sensor measurement or input is not always the same: it varies according to the context of the interaction which affects the agent's internal state. Successful behavioral systems are those which allow the public to develop an understanding of the causal relationships between their input and the agent's behavior. Ideally, the public should be able to narrate the dynamics of the encounter with a synthetic behavioral agent as they would narrate a story about a short interaction with a living entity, human or animal. This is one of the reasons why behavioral agents are often called life-like creatures. This type of authoring, inspired by Blumberg's work [5], allows the butterfly to take the right action at the right time, according to the information she gathers about the state of the participants with her virtual sensors. In particular the butterfly has senses to the size of the people's images and their location on the set.

6 Related Work

In the early days of Virtual Reality, Myron Krueger [8] built a series of interactive environments in which people would collectively participate in an artistic and entertaining experience from remote locations. Krueger's VIDEOPLACE consisted of separate rooms in which video cameras, mixers, and projectors would make it possible for people in any of the rooms to interact with the video images of others who were physically located elsewhere. Krueger had participants standing in front of backlights to help computers separate humans from the background. He then used specialized hardware to track the body movements of people. Participants would appear as "shadows" in a 2D projected environment in which they could interact with virtual objects and other shadow people.

Digital Circus differs from VIDEOPLACE in many ways. It has full compositing ability in real-time, i.e. it shows the human participant in the virtual environment at video resolution, and not simply as a silhouette. This enables people to see and recognize each other though the virtual world and use all the visual feedback they would normally use when interacting in bodily presence. In addition, Digital Circus is a full 3D environment and allows for more complex interaction with the virtual set, as well as the creatures and object present in it. It uses standard SGI computers and does not need any specialized hardware.

SURVIVE (Simulated Urban Recreational Violence Interactive Virtual Environment) is a networked entertainment application which shows some of the potential of body-driven home systems [9]. SURVIVE allows the participant to interact with a 3D game environment using the IVE space described in the next section. Figure 6 shows a user in SURVIVE. The gestural interpretation provided by the vision system is mapped onto the game controls for the popular Software game Doom. The player holds a large (two-handed) toy gun, and moves around the IVE room. The position of the participant in the room is fed into Doom's directional velocity controls. The hand position features, driven by the gun, which is interpreted by the vision system as a pointing arm, are used to drive Doom's rotational velocity control. SURVIVE was played by multiple players across countries in the Fall of 1995 between the MIT Media Lab in Cambridge, US, and BT Research Labs in the UK.

Digital Circus shares with SURVIVE the networked game aspect as well as the IVE setup. What is new in the circus is its ability to do compositing in real time without the use of blue screens, which opens up a wide range of possibilities for communication and the design of entertainment experiences. Our authoring methodology also allows us to construct a richer type of interactive experience.

The authors also did early work in game authoring with Hyperplex [10], an exploration game in which participants use body position and pointing gestures to select and view movie clips inhabiting a virtual palace of movies.

7 Conclusion

Research in Virtual Studios is tightly related with work in building immersive and interactive Virtual Environments in which people can interact with virtual objects, animated 3D creatures, or among themselves. We have used our experience in interactive virtual environments

Figure 6: Survive

and computer vision to build a low-cost virtual studio which does not need a blue screen to operate. Such a setup can be used at home and allows to distribute the production outside of the physical TV studio. Our studio is based on Pfinder: a real-time computer-vision body tracking and gesture recognition system. We use Pfinder to enable interaction among the composited participants, and the objects and creatures in the virtual setting. To demonstrate our system, we have built a digital circus in which multiple individual participants can connect from remote locations, see their images all composited in the virtual circus, and interact with the objects and creatures in the 3D set. We believe this work opens up new possibilities for TV production, such as distributed shows, news reports, game playing, collaborative storytelling, performance, and education.

References

[1] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 780–785, July 1997.

[2] T. Darrell, B. Blumberg, S. Daniel, B. Rhodes, P. Maes, and A. Pent-

land, "Alive: Dreams and illusions," in *ACM SIGGraph, Computer Graphics Visual Proceedings*, July 1995.

[3] T. Darrell, B. Moghaddam, and A. Pentland, "Active face tracking and pose estimation in an interactive room," in *CVPR96*, IEEE Computer Society, 1996.

[4] A. Azarbayejani and A. Pentland, "Real-time self-calibrating stereo person tracking using 3-D shape estimation from blob features," in *Proceedings of 13th ICPR*, (Vienna, Austria), IEEE Computer Society Press, August 1996.

[5] P. Maes, B. Blumberg, T. Darrell, and A. Pentland, "The alive system: Full-body interaction with animated autonomous agents," *ACM Multimedia Systems*, vol. 5, pp. 105–112, 1997.

[6] C. Wren, F. Sparacino, A. Azarbayejani, T. Darrell, T. Starner, A. Kotani, C. Chao, M. Hlavac, K. Russell, and A. Pentland, "Perceptive spaces for peformance and entertainment: Untethered interaction using computer vision and audition," *Applied Artificial Intelligence*, vol. 11, pp. 267–284, June 1997.

[7] P. Bouissac, *Circus and culture: a semiotic approach.* Bloomington, Indiana University Press, 1976.

[8] M. W. Krueger, *Artificial Reality II.* Addison Wesley, 1990.

[9] K. Russell, T. Starner, and A. Pentland, "Unencumbered virtual environments," in *IJCAI-95 Workshop on Entertainment and AI/Alife*, 1995.

[10] F. Sparacino, C. Wren, A. Pentland, and G. Davenport, "Hyperplex: A world of 3D interactive digital movies," in *IJCAI-95 Workshop on Entertainment and AI/Alife*, 1995.