

Motion Swarms: Video Interaction for Art in Complex Environments

Quoc Nguyen, Scott Novakowski, Jeffrey E. Boyd, Christian Jacob^{*}, Gerald Hushlak[†]
Departments of Computer Science and Art
University of Calgary
Calgary, Alberta, Canada T2N 1N4
{nguyenq,scottn,boyd,jacob}@cpsc.ucalgary.ca, hushlak@ucalgary.ca

ABSTRACT

We create interactive art that can be enjoyed by groups such as audiences at public events with the intent to encourage communication with those around us as we play with the art. Video systems are an attractive mechanism to provide interaction with artwork. However, public spaces are complex environments for video analysis systems. Interaction becomes even more difficult when the art is viewed by large groups of people. We describe a video system for interaction with art in public spaces and with large audiences using a *model-free, appearance-based* approach. Our system extracts parameters that describe the field of motion seen by a camera, and then imposes structure on the scene by introducing a swarm of particles that moves in reaction to the motion field. Constraints placed on the particle movement impose further structure on the motion field. The artistic display reacts to the particles in a manner that is interesting and predictable for participants. We demonstrate our video interaction system with a series of interactive art installations tested with the assistance of a volunteer audience.

Categories and Subject Descriptors

J.5 [Computer Applications: Arts and Humanities]: Fine Arts; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—*Animations, video*; H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Input devices and strategies*; I.4.8 [Image Processing and Computer Vision]: Scene Analysis—*Motion, tracking*

General Terms

Algorithms, Design, Experimentation, Human Factors

^{*}Nguyen, Novakowski, Boyd and Jacob are in Department of Computer Science

[†]Department of Art

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'06, October 23–27, 2006, Santa Barbara, California, USA.
Copyright 2006 ACM 1-59593-447-2/06/0010 ...\$5.00.

Keywords

Audience interaction, swarm art, motion analysis, motion history image, interaction through video, art installation

1. INTRODUCTION

As “The new electronic interdependence recreates the world in the image of a global village” [20] people can find themselves isolated by new technologies. It seems that people prefer to talk to those that are elsewhere than to talk to the person standing to them. Vonnegut [25] suggests that “Electronic communities build nothing. . . . We are here on earth to fart around.” In response, we *fart around* with technology to create interactive art that can be enjoyed by groups such as audiences at public events. This group interaction has the pleasant side effect of encouraging communication with those around us as we play with the art.

Video systems are an attractive mechanism to enable interaction with art. They are inexpensive, easy to install, and impose few physical constraints. However, public spaces are complex environments for video analysis systems. The number of people seen by a camera can vary from none to many. There may be motion in the background. Light and weather conditions may vary. Clothing varies. For effective interaction, the video system must accommodate all of these factors.

Interaction becomes even more difficult when the art is viewed by groups of people. Imagine spectators at a sports event watching an artistic display while interacting with and manipulating the display through their motion. In this scenario, video interaction faces all the confounding factors in any public space in addition to the scene complexity arising from the number of people interacting (and some people choosing not to interact). Figure 1 illustrates the complexity with an audience at an international speed skating event. The photographs show people clapping and banging objects to make noise, waving national flags, moving up and down stairs, and traversing the walkway behind the bleachers.

In this paper we describe a video system for interaction with art in public spaces and with large audiences. Rather than attempting to extract detailed information about people, objects, and trajectories from the video input, we use a *model-free, appearance-based* approach. In doing so, our system extracts parameters that describe the field of motion presented to the system independent of any object-specific knowledge related to the scene. This yields a robust system



Figure 1: Examples of audience activity at an international speed skating event illustrating scene and motion complexity.

that is immune to many of the factors that can confound video interaction. Although our *model-free* approach obviates the need for object-specific information, it must still impose some structure upon the motion field to avoid an apparently random reaction to complex motion. We get this structure by introducing a swarm of *particles* that moves in reaction to the underlying motion field. Constraints placed on the particle movement impose further structure on the motion field. The artistic display reacts to the agents in a manner that is interesting and predictable for participants.

We demonstrate our video interaction system with artistic displays generated from swarms of autonomous agents inspired by Boyd et al. [7]. The combination of esthetic displays and our interaction system yields engaging interactive installations.

Procedures for human-subject experiments and observations described in the paper were reviewed and approved by the ethics review board at the University of Calgary.

2. BACKGROUND

2.1 Large-Scale Audience Interaction

Event organizers often try to incorporate the audience into their events. By engaging the audience, they give people a sense of participation, reinforcing the notion that the audience is important to the organizers. In many cases, the mood of the audience can determine the success of an event. Therefore, many event organizers devise methods to engage an audience to keep them happy and entertained.

Many sporting events try to engage their audiences. For example, mascots interact with the audiences and influence them to cheer for a team. Video screens cue audiences to clap and make noise. To re-inforce the mood of the event, video footage of the excited audience is displayed on these screens. Television shows engage live audiences, such as the show “Who Wants to be a Millionaire”, where the audience can help a contestant through a voting system. Various universities entice students to participate during lectures by using the Classroom Performance System(CPS) [12].

Audience motion and behavior is complex, making unconstrained audience interaction practically impossible. To make interaction succeed, it is necessary to constrain the interface in some way.

One option is to constrain interaction to a physical input device, e.g., a device with which a person can enter responses to stimuli. eInstruction [12] has developed such input devices for classroom environments to give a lecturer immediate feedback from the students. Scheible et al. [23] demonstrated audience interaction using a cell phone as a voting device to select music videos. Input devices are an effective way to receive audience input but there are certain constraints associated with these devices including distribution and price of the devices. The devices can be cumbersome or difficult to use, having a negative impact on the audience.

Maynes-Aminzade et al. [19] devised a series of games to engage a large audience. Their primary method of interaction was through basic computer vision techniques with the aid of some simple props. In one example, the audience leans left or right to control a paddle in the video game *Pong*. In this case, constraining the allowable audience motion makes interaction possible. In another example, the audience throws a ball around. The ball casts a shadow onto a screen which was then tracked using computer vision techniques. The audience used the ball to control input to the *Missile Command* video game. The ball acts as a mechanism to reduce complexity and facilitate interaction.

Breazeal et al. [8] devised a robot that responds to a small audience using computer vision techniques. Audiences viewing the robot compete for the attention of the robot using hand and face movements. The robot detects the movements of faces and hands using features extracted from images of the audience. In this case, the interaction is constrained by having the robot follow simple motions and limiting the quality of interaction that is possible.

In this paper, we advance the use of *model-free* computer vision techniques by imposing structure on the motion to reduce complexity and facilitate meaningful interaction.

2.2 Model-Free Vision

Historically, computer vision research in human motion analysis can be divided into two broad categories: model-based and model-free. The model is kinematic, describing the interconnected parts that make up the human body and their kinetic relationship to each other. Model-free methods are sometimes referred to as *appearance-based*. Much of the vision research on human motion is focused on either recognizing some activity or action, or recognizing an individual as in biometric gait recognition.

In the model-free category Little and Boyd [17] describe a *shape-of-motion* system that analyzes periodicities in the shapes of people as they walk. Baumberg and Hogg [2, 3] describe the shape of the silhouette of walkers. Bobick and Davis [4, 5] accumulate motion information into a motion energy image and a motion history image to recognize human activities. Boyd [6], Liu and Picard [18], and Polana and Nelson [21] look at oscillations in image intensities due to gait in order to analyze gaits. Cutler and Davis [10] describe a system that identifies periodicities in the motion of objects in order to identify human and animal gaits. Hari-toglu et al. [14] describe a system that classifies a silhouette of a human into one of four main postures, and then further classify it into one of three view-based appearances.

In the model-based category Hunter et al. [15], Rowley and Rehg [22], and Bregler and Malik [9] use models that represent joints and limbs in a human. They estimate the

kinematic pose of a person by extracting joint angle information from a sequence of images. Wachter and Nagel [26] describe a system that derives pose by tracking individual body parts using constraints imposed by the known kinematic structure. While these approaches use an explicit kinematic model, the model need not be so detailed. Wren et al. [27] track a set of *blobs* that are parts of a person. Their system identifies some body parts based on their expected locations in typical views. They do not strictly enforce a kinematic model. Fujiyoshi and Lipton [13] use a simplified model of a person that maps a *star* skeleton onto the silhouette of a moving person.

Also within the model-based category are motion capture systems. These systems, pioneered by Johansson [16], typically use markers placed on the limbs of the subject that are then tracked. Johansson originally used marker motion as a stimulus in psychological experiments, but currently, the most common use is to recover the kinematic pose of the subject for applications in computer games, graphics and research in kinesiology.

While marker-based motion capture is the most precise method, it is the least practical for interaction with art due to the need for markers on the participants. Other model-based methods eliminate the need for markers, but still require strict control over imaging conditions that eliminate the possibility for interaction in a complex environment. Furthermore, recovering kinematic pose for many people in an audience simultaneously is not practical.

Model-free methods provide the most likely method to allow an audience to interact with art. However, existing methods focus on the actions of a single person or a small group. Motion patterns become more complex as the number of people increases. To use model-free methods for interaction requires further refinement.

3. MOTION SWARMS

Motion history images [4, 5] (MHI) provide a representation of a field of motion that is independent of the number of people or scene complexity. However, for useful interaction, it is necessary to find some order in the motion field. Bobick and Davis [4, 5] do this by characterizing the shape generated by a single person in the field with a set of moments. This approach is prohibitive for a large group of people since no small set of moments can capture the richness and complexity of group activities.

We introduce the concept of a *motion swarm*, a swarm of particles that moves in response to the field representing a motion history image. Structure imposed on the behavior of the swarm forces a response to the motion history images. To create interactive art, the art responds to the motion swarm and not the motion directly. The appearance of the swarm gives structure to the interaction so the art responds to the audience in a meaningful way and does not appear to be random.

3.1 Motion Intensity Field

Pixels in an MHI indicate how recently a pixel has changed in intensity. Brighter pixels have changed more recently while darker pixels have not changed. Changes in intensity are presumed to be due to motion.

Let $T_k(\mathbf{x})$ be a binary image that indicates whether or not the pixel at $\mathbf{x} = [x, y]^T$ changed significantly in video frame k . We can compute T_k using adaptive background

subtraction as follows:

$$T_k(\mathbf{x}) = \begin{cases} 1 & |I_k(\mathbf{x}) - \tilde{I}_k(\mathbf{x})| \geq \tau \\ 0 & \text{otherwise} \end{cases},$$

where I_k is the image at time k , \tilde{I}_k is I smoothed in time, and τ is a threshold that determines what intensity change is significant. The temporal smoothing of I over a wide time window allows the *background* to adapt to slow changes in scene illumination. A recursive, infinite impulse response (IIR) filter allows for computationally efficient smoothing over broad temporal windows.

The MHI at time k is given by

$$M_k(\mathbf{x}) = \max(cT_k(\mathbf{x}), M_{k-1}(\mathbf{x}) - 1).$$

Note that $cT_k(\mathbf{x}) \in \{0, c\}$. Thus, when a pixel changes, the corresponding pixel in the MHI is set to c , otherwise, the value is decremented, never going below zero. The constant, c sets the *persistence* of the motion history.

Since we desire to have swarm particles respond in a *natural* and predictable manner, we smooth $M_k(\mathbf{x})$ in space to get

$$\tilde{M}_k(\mathbf{x}) = M_k(\mathbf{x}) \otimes G(\mathbf{x}; \sigma),$$

where \otimes indicates convolution and G is a Gaussian kernel. This also broadens the basin of attraction for particles described in Section 3.2. In our application we select σ large, and a recursive filter [11] provides a computationally efficient implementation for arbitrary values of σ .

3.2 Particles

In public places and large audiences, the motion field is complex. To extract some structure from the field, we introduce swarms of *particles* that move in response to the MHI by treating the MHI as a field. The gradient of the MHI yields *forces* that act on the particles. We then introduce other forces such as friction and neighborhood interactions. Simulation of the forces acting on the particle leads to the particle motion that forms the basis for interaction.

Let $\mathbf{x}_k = [x, y]^T$ be the position of a particle, $\mathbf{v}_k = [v_x, v_y]^T$ be its velocity, and $\mathbf{p}_k = [p_x, p_y]^T$ be its momentum at time interval k . Δt is the time sample interval. The following equations (used in physics simulations and video games) simulate particle movement in response to a force F acting on it:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{v}_{k-1}\Delta t, \quad (1)$$

$$\mathbf{p}_k = \mathbf{p}_{k-1} + F_k(\mathbf{x})\Delta t, \quad \text{and} \quad (2)$$

$$\mathbf{v}_k = \frac{\mathbf{p}_k}{m}, \quad (3)$$

where m is the particle mass. In the context of our system where particles are imaginary points with no mass, m is a tunable constant.

For a particle at position \mathbf{x} , the force due to the motion field is

$$F_{M_k} = \nabla \tilde{M}_k(\mathbf{x}) \approx \frac{1}{2} \begin{bmatrix} \tilde{M}_k([x+1, y]^T) - \tilde{M}_k([x-1, y]^T) \\ \tilde{M}_k([x, y+1]^T) - \tilde{M}_k([x, y-1]^T) \end{bmatrix} \quad (4)$$

If we let $F = F_{M_k}$, then the particles will tend to move up the gradient of the MHI. Since the brightest pixels in the MHI are where the motion has most recently occurred, the particles tend to follow the motion. Alternatively, the motion can repel the particles by either setting $m < 0$, or letting

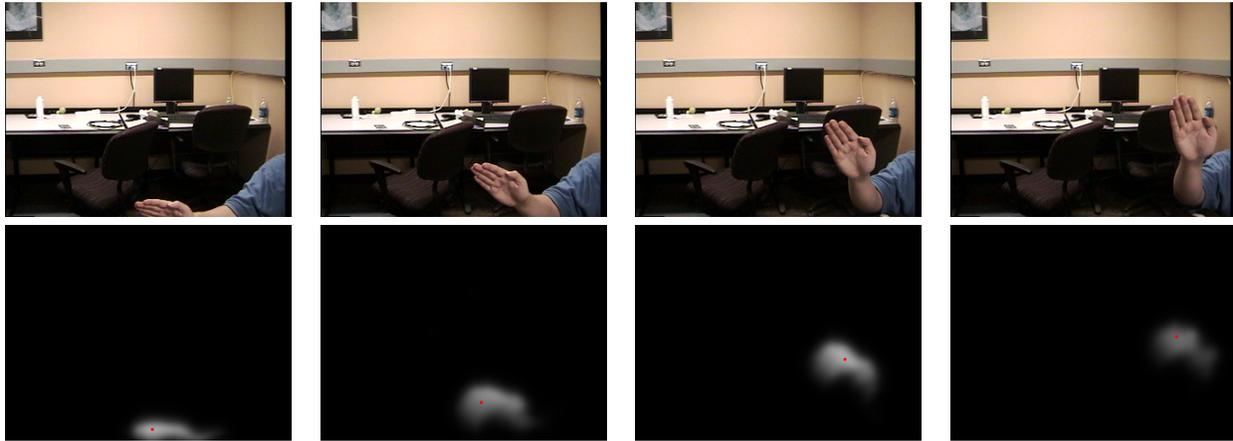


Figure 2: Example showing motion particle generation. The top row shows a sequence of images with a hand moving upward. The bottom row shows the corresponding motion fields (smoothed MHI) with a single particle following the motion superimposed.

$F = -F_{M_k}$. Figure 2 illustrates the MHI and motion particles with an example.

3.3 Additional Structure

The introduction of moving particles to the motion field adds structure to the representation of the motion by sampling at a fixed number of points (the particle positions) and by forcing the particles to react to the motion field. Further constraints to the particle motion add more structure to the particle representation.

Additional constraints that we use in our work include:

- **Friction:** Frictional forces act in opposition to particle velocity. Without friction, the particles continue to accelerate, getting faster every time they encounter motion. Friction allows the particles to slow down in the absence of motion, and can prevent particles from *shooting past* regions of motion because their velocity is too high.
- **Momentum Limits:** Particles that move fast are not conducive to interaction. People cannot keep up with them. Therefore, we place an upper and lower bound on the magnitude of the particle momentum. The upper bound prevents the particle from moving too fast, and the lower bound prevents the particle from coming to a complete stop.
- **Bounding Box:** It is useful to constrain the particles to move only with a defined bounding box. At the very least, a bounding box the size of the images is necessary to keep the particles from going beyond the image boundaries. However, bounding boxes can also be useful to define sub-images of interest. For example, in an interactive game between two groups of people, we can define a bounding box for the image region corresponding to each group.
- **Anchors:** We can *anchor* a particle by defining a central position and adding a force that propels the particle toward that position. This is useful when we want to maintain a distribution of points throughout the image space. Without the anchors, the particles can

simultaneously follow motion to one part of the image and leave large portions of the image unsampled, and therefore unavailable for interaction. We model the anchoring force as a spring between the center position and the particle.

Each of these parameters introduces tunable parameters to the system. Tuning the parameters alters the behavior of the particles significantly and alters the nature of the interaction. We demonstrate a number of different behaviors achieved by tuning parameters in the exhibits described in Section 5.

4. BUILDING ART INSTALLATIONS

Our interactive art installations consist of two basic components:

1. **Video Processing:** Video processing acquires the images, computes the MHI, smooths the MHI to get a motion field, and does particle simulations.
2. **Artistic Display:** The artistic display produces the images that viewers of the installation actually see.

This section describes our implementation of these components.

4.1 Video Processing

We do all video processing on a dedicated system acting as a server. We code in Python, relying on C- and assembly-based libraries to do the fast inner-loop computations required for the video. The server communicates via network, exchanging XML documents and image data via HTTP protocol. We configure the server and tune particle parameters by sending XML documents to the server. The server broadcasts source video, MHI, and XML documents containing the particle positions.

4.2 Artistic Display

The artistic display communicates with the video server to acquire images and particle positions. It then does simulation and rendering to produce an esthetically pleasing



Figure 3: Photograph of the trial in progress. The audience views the projected display from seats in a large lecture theater. A camera beside the projection screen views the audience providing input to the artistic display.

display. We use either *Breve* [24] or *Quartz Composer* [1] for simulation and rendering.

Breve is a swarm simulation system with excellent capabilities for rendering swarms for visualization. Our use of motion swarms for interaction is a natural match for *Breve*. We added modules to *Breve* to allow it to interact with the video processing system and read XML documents. In addition to its rendering capabilities, *Breve* also facilitates production of sound, which we exploit in one of the installations described in Section 5.

Quartz Composer is a high-level development system used for generating three-dimensional images. The graphical language used to program in *Quartz Composer* allows us to quickly generate impressive three-dimensional rendering. We use this extensively in the installations described in Section 5.

5. AUDIENCE TRIALS

Live audiences suitable for testing interactive art are not easy to come by. Organizers responsible for events that draw a large audience are reluctant to allow an untried installation to become part of their event. Conversely, without an audience, we cannot test the installation to acquire the evidence necessary. Thus, to demonstrate our installations, we are restricted to working with an artificially assembled audience. In our case, we assembled a group of approximately 25 volunteers to participate as a makeshift audience in a university lecture theater. We placed a video camera at the front of the room and pointed it towards the audience, and then used the available projection equipment to display the renderings for all to see (see Figure 3). Participants were encouraged to move as they desired, but no instruction on how to interact were given. We provided cardboard squares the participants could wave to produce motion for the cameras. This was particularly useful for those at the back whose images would otherwise have been quite small in the camera due to foreshortening.

We recorded images of the rendered display and of the audience directly. Lighting was challenging. We wanted dark-



Figure 4: Interactive music with motion swarm interaction. Participant movement propels the balls. The blue band at the top of the image emulates a keyboard that plays music when struck by the balls.

ness to enhance the projected display, but also wanted light so that the camera could see the motion. The compromise we found was to use the *Night Shot* feature on the camera which removes the infrared filter in the camera and provides some infrared illumination to compensate for dimmed lights. While this was effective, it does result in washed-out images devoid of color, as can be seen in the examples that follow. As liquid crystal and plasma displays grow in size (and our budget too) we can move to alternate display technologies that are better for simultaneous video interaction and display.

The remainder of this section describes four installations we tested with our makeshift audience and our observations.

5.1 Music

The first trial installation gets the mundane yet descriptive label, *Music* since it allows audience to produce music. Simulation, rendering, and sound generation for *Music* was done using *Breve*. Figure 4 shows a photograph of the rendered display.

The audience sees an image of themselves, reversed left-to-right to give the effect of a mirror. Superimposed upon the image is a blue band at the top of the image and a set of green balls. The ball positions correspond to the positions of particles simulated by the video system. Parameters are tuned so that motion *repels* the particles/balls. As the audience waves hands and swats at the balls, they propel them around the display. The blue band at the top of the image behaves as a virtual keyboard. When a ball hits the blue band, the keyboard plays music and the audience is rewarded with the sounds they generate.

Observations: Although the audience received no instructions, they were quick to grasp the mode of interaction (Figure 5(a)). In a few seconds, the audience was swatting at balls and producing music. The music they produce, although not strictly random and constrained to play on a natural scale, was not sophisticated enough to sustain interest after a three or four minutes of play. The number of balls in the display was not conducive to team play. People merely swatted at whatever ball was closest. There was no



Figure 7: Photograph of hockey installation where team logos move with the motion swarm particles. Logos of *opposing* teams appear on opposite sides of the screen while participants compete to move the logos of their team. The logo particles are constrained to return to a *home position*.

need to cooperate since no matter what one did, another ball was sure to come by in a moment.

5.2 Volleyball

We labeled our second trial installation *Volleyball* because the interaction allowed the audience to move a ball back and forth as in a volleyball game. Simulation and rendering were done using Quartz Composer. A series of images showing this installation in action is in Figure 6.

Again, the audience sees a mirror image of themselves on the display. Superimposed on the display is a single ball. In this case, the ball is *attracted* to motion. As the ball moves, the audience sees that their image is rendered on a surface that moves with the ball. In fact, the displayed position of the ball never changes. Only the image of the audience moves. A tail flowing behind the ball emphasizes the ball's motion.

Observations: Initially the audience has trouble interacting (Figure 5(b)). This may be because of the preceding trial where their motion repelled balls propelling them around the display. Now when a person moves near the ball, the ball moves to them and is *stuck* there until they stop moving. This difficulty does not last long and the audience eventually starts to pass the ball around the room. It is difficult because the participants must coordinate their motion, first moving to attract the ball, then stopping to allow the ball to move on, all synchronized among the group. This was our most successful trial with respect to team work. A single ball and the need to coordinate group motion encouraged cooperation and interaction among audience participants.

5.3 Hockey

Our third trial we labeled *Hockey* due to the use of professional hockey team logos in the display. Simulation and rendering were done by Quartz Composer. Figure 7 shows a photograph of this installation in action.

The audience sees a display with the names of rival hockey teams, one above each side of the display. Six copies of each team logo appear below their team name. As the audience

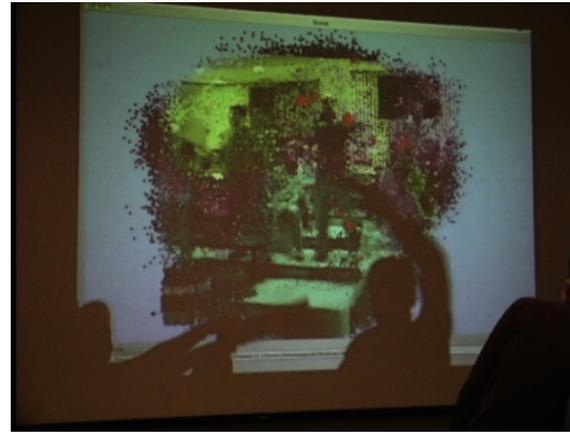


Figure 8: Photographs of the *Spray Paint* installation. Images of the audience are digitally *sprayed* onto the display at the locations of motion swarm particles.

moves, the logos are *attracted* by the motion causing the logos to move. As the logos move, they create a brilliantly colored tail and the team name begins to pulse. The audience is able to engage in a sort of competition to get their team name and logo to be the most animated. Each logo corresponds to a motion particle. The particle motion is constrained to remain in its side of the image by a bounding box. Also, the logos are anchored to maintain coverage of the audience.

Observations: The goal is simple and the audience grasps it quickly, working to create motion and keep the logos moving (Figure 5(c)). No coordinated behavior is required, but a rivalry between the two sides quickly emerges. The interaction is fun, but the high energy output required to maintain the motion exhausts enthusiasm after one to two minutes.

5.4 Spray Paint

Our fourth, and final installation we call *Spray Paint*. Simulation and rendering are done in Breve. Figure 8 shows photographs of the installation.

The audience sees an image with a set of three balls superimposed. The balls follow the positions of motion particles, acting as virtual spray cans, spray painting the image of the audience onto the display. Where the balls are present, we



(a)



(b)



(c)



(d)

Figure 5: Photographs of audience interaction with (a) *Music*, (b) *Volleyball*, (c) *Hockey*, and (d) *Spray Paint*.

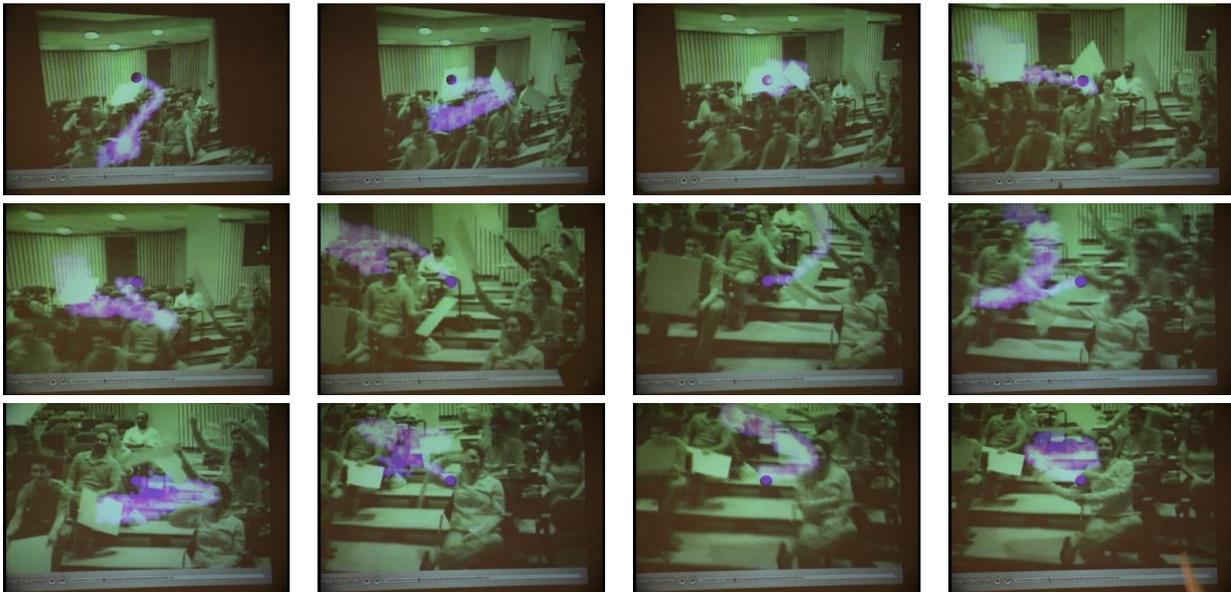


Figure 6: Sequence showing interaction with *volley ball* exhibit (order is left to right, top to bottom). Participants manipulate a single *ball* that is attracted to motion. The display warps the image of the audience so that the ball remains at the center of the display. Participants can co-operate to move the *ball* around the audience.

see moving video rendered with a spray-paint effect. Where the balls are absent, the image is frozen as it appeared when last sprayed.

Observations: When the audience was seated, this installation proved less interesting than the first three. However, when the audience begins to move around the room, the spray-painting effect comes to life (Figure 5(d)). As people walk in front of the camera, across rows in the lecture theater, and up and down the stairs, they produced intriguing images and engage in play with the installation. The temporal effects where some of the display is moving and other parts are frozen was particularly fun as people tried to control the motion to steer the balls and update the image.

6. DISCUSSION AND CONCLUSIONS

Public spaces and audiences are complex environments for video interaction. Meaningful interaction is only possible by constraining interaction to fewer, simpler entities. We demonstrated a video interaction system for art installations that resolves the complexity by first analyzing the motion with model-free vision techniques, then modeling a swarm of particles to impose some structure on the motion. The system is versatile in that it is tolerant of a variety of imaging conditions and will produce meaningful interaction in arbitrarily complex situations.

Our model-free appearance-based approach faces a dilemma in interactive art. In our society, we have become accustomed to computer interaction with the precision afforded by a mouse. In order to accommodate complex scenes, we abandon this precision in favor of a method that is robust and versatile. Although our volunteer audience had fun with our installations, we sense that they craved to point and click. We are inclined to resist this craving so that our installations can take people away from their routine interactions with technology into another realm that is perhaps more fun.

We did not ignore esthetics in our installations, but we certainly have not explored the esthetic potential of our video interaction/display medium. Our translation from motion particles to visual entities in the display was direct and literal. There is definitely space to play in the esthetic dimension to find even more varied and interesting displays.

7. REFERENCES

- [1] Apple Computer Inc. Working with quartz composer. Retrieved May 30, 2006, from <http://developer.apple.com/graphicsimaging/quartz/quartzcomposer.html>, 2005.
- [2] A. M. Baumberg and D. C. Hogg. Learning flexible models from image sequences. Technical Report 93.36, University of Leeds School of Computer Studies, October 1993.
- [3] A. M. Baumberg and D. C. Hogg. Learning spatiotemporal models from training examples. In *British Machine Vision Conference*, Birmingham, September 1995.
- [4] A. F. Bobick and J. W. Davis. An appearance-based representation of action. In *13th International Conference on Pattern Recognition*, Vienna, Austria, August 1996.
- [5] A. F. Bobick and J. W. Davis. Real-time recognition of activity using temporal templates. In *Third International Workshop on Applications of Computer Vision*, Sarasota, Florida, December 1996.
- [6] J. E. Boyd. Video phase-locked loops in gait recognition. In *International Conference on Computer Vision*, pages 696–703, Vancouver, BC, 2001.
- [7] J. E. Boyd, C. Jacob, and G. Hushlak. Swarmart: interactive art from swarm intelligence. In *ACM Multimedia 04*, pages 628–635, New York, NY, October 2004.
- [8] C. Breazeal, A. Brooks, J. Gray, M. Hancher, J. McBean, D. Stiehl, and J. Strickon. Interactive robot theatre. *Commun. ACM*, 46(7):76–85, 2003.
- [9] C. Bregler and J. Malik. Tracking people with twists and exponential maps. In *Computer Vision and Pattern Recognition 1998*, Santa Barbara, June 1998.
- [10] R. Cutler and L. Davis. Robust periodic motion and motion symmetry detection. In *Computer Vision and Pattern Recognition 2000*, 2000.
- [11] R. Deriche. Fast algorithms for low-level vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):78–87, 1990.
- [12] eInstruction. Classroom performance system. Retrieved May 30, 2006, from <http://www.einstruction.com/>.
- [13] H. Fujiyoshi and A. J. Lipton. Real-time human motion analysis by image skeletonization. In *DARPA Image Understanding Workshop*, Monterey, California, November 1998.
- [14] I. Haritaoglu, D. Harwood, and L. S. Davis. Ghost: a human body part labelling system using silhouettes. In *DARPA Image Understanding Workshop*, pages 229–235, Monterey, California, November 1998.
- [15] E. A. Hunter, P. H. Kelly, and R. C. Jain. Estimation of articulated motion using kinematically constrained mixture densities. In *Nonrigid and Articulated Motion Workshop*, San Juan, Puerto Rico, June 1997.
- [16] G. Johansson. Visual motion perception. *Scientific American*, pages 76–88, June 1975.
- [17] J. J. Little and J. E. Boyd. Recognizing people by their gait: the shape of motion. *Videre*, 1(2):1–32, 1998.
- [18] F. Liu and R. W. Picard. Finding periodicity in space and time. In *International Conference on Computer Vision*, 1998.
- [19] D. Maynes-Aminzade, R. Pausch, and S. Seitz. Techniques for interactive audience participation. In *ICMI '02: Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, page 15, Washington, DC, USA, 2002. IEEE Computer Society.
- [20] M. McLuhan and Q. Fiore. *The Medium is the Message*. Bantam Books/Random House, 1967.
- [21] R. Polana and R. Nelson. Detection and recognition of periodic, nonrigid motion. *International Journal of Computer Vision*, 23(3):261–282, June 1997.
- [22] H. A. Rowley and J. M. Rehg. Analyzing articulated motion using expectation-maximization. In *Computer Vision and Pattern Recognition 97*, pages 935–941, San Juan, Puerto Rico, June 1997.

- [23] J. Scheible and T. Ojala. Mobilenin - Combining a multi-track music video, personal mobile phones and a public display into multi-user interactive entertainment. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 199–208, New York, NY, USA, 2005. ACM Press.
- [24] L. Spector, J. Klein, C. Perry, and M. Feinstein. Emergence of collective behavior in evolving populations of flying agents. In E. C.-P. e. al., editor, *Genetic and Evolutionary Computation Conference (GECCO-2003)*, pages 61–73, Chicago, IL, 2003. Springer-Verlag.
- [25] K. Vonnegut. *A Man Without a Country*. Seven Stories Press, 2005.
- [26] S. Wachter and H.-H. Nagel. Tracking of persons in monocular image sequences. In *Nonrigid and Articulated Motion Workshop*, San Juan, Puerto Rico, June 1997.
- [27] C. Wren, A. Azarbayenjani, T. Darrell, and A. P. Pentland. Pfindex: real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):780–785, July 1997.