Technologies and methods for interactive exhibit design: from wireless object and body tracking to wearable computers

Flavia Sparacino (*), Kent Larson (**), Ron MacNeil (**), Glorianna Davenport (*), Alex Pentland (*)
(*) MIT Media Lab
(**) MIT Architecture, CAVS
{flavia,gid,sandy}@media.mit.edu, {kll,ronmac}@mit.edu

Abstract
We present three interactive exhibit projects which add technology to the museum space or to the museum visitor. We propose a technological intervention which helps curators and designers to achieve a balance between leisure and learning and help them be more effective in conveying story and meaning. This is made possible by tracking people and objects with wireless and unencumbering real time computer-vision techniques, and with wearable computers which can respond to the context and presence of the visitor along the path of the exhibit. By using these techniques museums can present a larger variety and more connected material in an engaging manner within the limited physical space available. They can also enrich and personalize the visit with a wearable computer-driven visual and auditory storyteller which can adapt and guide the public through the path of the exhibit. All these systems enhance the memory of the visit and help build a constructivist-style learning experience for the public.

Introduction

Applications of technology to museums have so far mainly focused in making extensive and attractive web sites with catalogues of exhibits. Occasionally these web sites also present introductory or complementary information with respect to what is shown inside the physical space of the museum. However, unless the public is interested in retrieving a specific information about an artist or artwork, they will end up spending time scrolling across photographs and text in static pages, and likely are not involved in an engaging or entertaining experience. Presenting large bodies of information in the form of an electronic catalogue, usually does not stimulate learning or curiosity.

Museums have recently developed a strong interest in technology, as they are more than ever before in the orbit of leisure industries. They are faced with the challenge of designing appealing exhibitions, handling large volumes of visitors, and conserving precious artwork. They look at technology as a possible partner which can help archive a balance between leisure and learning as well as help them be more effective in conveying story and meaning.

One of the main challenges that exhibit designers are faced with is that to give life to the objects on display by telling their story within the context determined by the other objects in the exhibit. Traditional storytelling aids for museums have been panels and labels with text placed along the visitors' path. Yet the majority of visitors express uneasiness with written information. Usually time spent reading labels interrupts the pace of the experience and requires a shift of attention from observing and contemplating to reading and understanding [Klein, 1986]. Successful labeling is sometimes hierarchical and allows the viewer to select the desired degree of information.

Another challenge for museums is that of selecting the right subset of representative objects among the many belonging to the collections available. Usually, a large portion of interesting and relevant material never sees the light because of the physical limitations of the available display surfaces.

Some science museums have been successfully entertaining their public, mainly facilitated by the nature of the objects they show. They engage the visitor by transforming him/her from a passive viewer into a participant by use of interactive devices. They achieve their intent, amongst other things, by installing button-activated demonstrations and touch-sensitive display panels which provide supplementary information when requested. They make use of proximity sensors to increase light levels on an object when a visitor is close-by and/or to activate a process.

Other museums -- especially those which have large collections of artwork, like paintings, sculptures, and manifactured objects -- use audiovisual material to give viewers some background and a coherent narrative of the works they are about to see or that they have just seen. In some cases, they provide audio-tours with headphones along the exhibit. In others, they dedicate sections of the exhibit to the projection of short audiovisual documentaries about the displayed material. Often, these movies that show artwork together with a description of their creation and other historical material about the author and his times, are even more compelling than the exhibition itself. The reason is that the documentary has a narration, the visuals are well orchestrated and come with music and dialogues. The viewer is then offered a more unified and coherent narration than in the fragmented experience of the visit. A visit to a museum demands, as a matter of fact, a certain amount of effort, knowledge, concentration, and guidance, for the public to leave with a consistent and connected view of the material presented.

Motivation

Based on the above observations, we have identified two areas of technological intervention which contribute to engage the public and enrich its experience during the museum visit. They are: Information Overlay in Smart Rooms (adding technology to the museum space) and Spatialized Interactive Narrative with Smart Clothes (adding technology to the visitor).

Information Overlay in Smart Rooms

By using a variety of wireless sensors (primarily cameras) and by placing audiovisual devices (projectors/speakers) in the museum area, we can use technology to virtually enlarge and augment the exhibit surface. It is then possible to select and show more objects from the ones available in the collection -- in the form of images which can be virtually layered on each other, and selcted by meaningful interaction with the public.

A typical example is that of the curator who wants to show photographs of the facade of a monument at different historical times. If for example 20 large prints were made, framed, and hung on walls they would cover a large surface and take room away from other relevant material of the exhibit. The 20 images can instead be scanned at high resolution and be projected interactively on one dedicated surface. An interaction device, ranging from a simple slider, to a computer camera tracking the visitor from above or the front, offers the public a selection device which can control which image is projected at one time. This method has also other advantages over the traditional one. The first is that the information overlay highlights changes and evolution of the facade through time whenver a transition from one image to the next is triggered by the visitor. The second advantage derives from the constructivist nature of the interactive experience. Empowering the public with the ability to cause the transition which shows the evolution of the facade through time enriches the learning with an actual experience and stimulates memory and curiosity. The learner in the museum is then engaged in an "active process in which s/he uses sensory input and constructs meaning out of it" [Hein, 1991].

Spatialized Interactive Narrative with Smart Clothes

Technology can help construct a coherent narrative of the exhibit for the visitor by creating experiences in which the objects on display narrate their own story in context. Wearable computers have recently raised to the attention of technological and scientifical investigation [Starner, 1997] and offer an oportunity to "augment" the visitor and his perception/memory/experience of the exhibit in a personalized way. Wearable computers can be used to simulate a museum curator, or to dynamically edit a documentary about the shown artwork which is interactively edited according to the path of the visitor inside the physical space of the museum. As opposed to creating interactive experiences which are driven by one or more visitors, and can be seen by all the surrounding visitors, such as the ones described above, wearable computers target the individual visitor with special learning needs or curiosity. A wearable computer usually comes in the form of a jacket or vest with a small embedded computer, an on-board sensing system, and a lightweight headmounted display, or glasses with a miniature computer monitor in them. Sensors placed in key locations of the exhibit site, signal to the wearable computer its vicinity to a selected location, and trigger an appropriate response on the personalized display. Usually the display is placed only in front of one eye, and therefore the viewer sees the external world as it is, as well as the superimposed graphics, text, or images, interactively played by the wearable.

Projects

Information Overlay in Smart Rooms: Responsive Portraits

We presented Responsive Portraits, our first interactive museum project at ISEA 97, in Chicago. We described a technique to display photographs of people in a more engaging manner, such that the people portrayed would react to the presence, actions and expressions of the viewer [Sparacino, 1997a]. Responsive portraits challenge the notion of static photographic portraiture as the unique, ideal visual representation of its subject. When compared to film, photography seems to carry an intrinsic narrative poverty, because of its static nature. In the case of portraiture, portraits are usually read not as stories but as symbols, short visual poems that describe a unique and immediate perception of reality. Moreover editing single photographs for magazines or exhibits can be a frustrating experience for the artist as it requires discarting a number of photographs which all contribute to define the story or personality of the portrayed subject. Responsive portraits seem to be filling some of these gaps by incorporating a story behind the photographic portraits, by letting the photographs tell the viewer their own story through the interaction. Here the meaning of a photograph is enriched by its relationship to the other photographs in the set, and to the story line attached to them.

A responsive portrait consists of a multiplicity of photographs virtually layered on a high-resolution digital display. The image which is shown at a given time depends on how the viewer approaches and reacts to the portrayed subject. A small active computer-controlled camera is placed right above the display. By using real-time computer vision techniques we are able to determine how close/far the viewer is to the portrayed character, her viewing angle, and we can also interpret some of her facial expressions, such as smile, laughter, suprise or disappointment.

The viewer's proximity to the image, head movements, and facial expressions elicit dynamic responses from the portrait [Figure 1], driven by the portrait's own set of autonomous behaviors, which are modeled by a computer program. This type of interaction reproduces an encounter between two people: the viewer and the character portrayed. The experience of an individual viewer with the portrait is unique, because it is based on the dynamics of the encounter rather than the existence of a unique, ideal portrait of the subject. As the participant observes the portrait s/he is also being observed by it: the notion of "who is watching who" is reversed: the object on display becomes the subject, and the viewer is observed. The uniqueness of the portrait is transferred to the uniqueness of the encounter between the viewer and the portrayed character. In this sense, the viewer and the artist cooperate in creating an artistic experience. Another shift then happens which leads from the reproducibility of a work of art to the uniqueness of the experience happening in the exhibition gallery or art museum. Following on the Bahaus concept of a "Modern Exhibition", exhibited art should not retain its distance from the spectator. It should be brought close to him, penetrate and leave an impression on him. It should explain, demonstrate, and even persuade and lead him to a planned reaction. In this sense exhibit design can borrow from the psychology of advertising.

Responsive portraits are created in two steps. First the photographer goes on assignement and shoots an extended set of portraits of her subject in a variety of poses, expressions, gestures, significant moments. We feel that is important that at this stage the artist concentrates on connecting with its subject and postpones editing choices to the next step. Later, editing happens. In the case of responsive portraits the photographer can choose at this stage not only what the public will experience but also how it will be experienced. The artist can edit a set of pictures which map her experience approaching the subject. Otherwise she can choose another set which represents a landscape of portraits of a person which changes according to the point of view of the observer. It is important to notice that at this stage the artist does not do a final edit of what the viewer is going to see. The artist only sets up the terms of the encounter between the public and the portrayed character by choosing a basic content set and a mapping. Mapping is done by autonomous agent based modeling of content. In this work we used our previous research and implementation of Media Actors [Sparacino, 1999a]. Media Actors are autonomous agents with goals, behaviors, and sensors. A Media Actor knows whether its content is text, image, movie clip, sound, or graphics and acts accordingly. It also has a notion of its role and "personality attributes."

Figure 1. Reactions of a responsive portrait to the viewer's proximity and behavior.

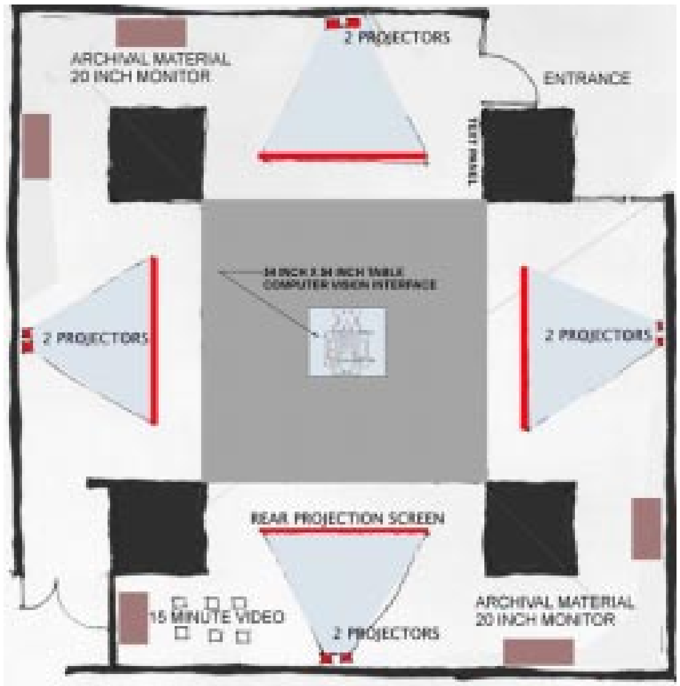
According to the type of chosen mapping we are gathering content and implementing three types of Responsive Portraits: 1. The Extended Portrait; 2. The Responsive Hologram; 3. The Photographic Essay. The Extended Portrait maps single aspects of the personality of the portrayed subject to the "personality" of a Media Actor. Extended Portraits include: "The Chronological Portrait", which layers photographs of a person across time; "The Expressive Portrait" which sets up a communicative facial expressions game between the portrayed character and the viewer; and the "Gestural Portrait" which uses a wider framing of the subject, including hands, to engage the public in the interactive experience.

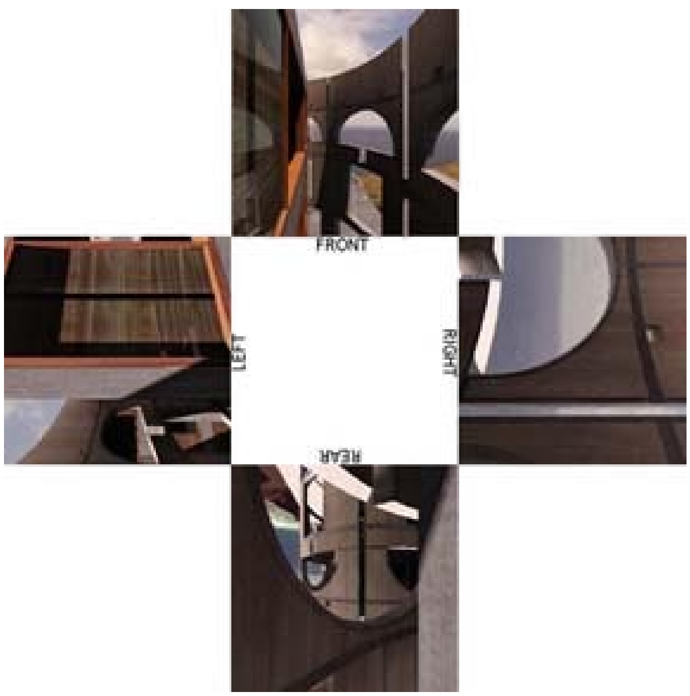
Responsive Holograms are portraits which react as a function of the viewing angle of the observer as certain well known holograms. These holograms show a sequence of an action as the viewer moves her head horizontally across the display. In our system the portrayed subject changes her pose/expression according to the observation point of the viewer. The metaphor here is that we tend to see people according to our own emotional and experiential perspective coordinates: as these coordinates change we acquire new knowledge and understanding of the people surrounding us. Lastly, the "Photographic Essay" addresses the challenge of letting the public edit a photographic narrative piece through the interactive feedback of distance from the subject, point of view, and facial expression.

The interactive interface is a real-time computer vision system named LAFTER [Oliver, 1997]. LAFTER is an active-camera, real-time system for tracking, shape description, and classification of the human face and mouth. By using only an SGI INDY computer it is able to provide a wide range of information about the person appearing in the frame, such as: the center of the bounding box of the head and mouth, the rotation angle of the face and mouth about the axis given by the standing body, size of face and mouth, distance of the viewer from the camera, head motion, facial expression recognition -- the person is: surprised, smiling, laughing, sad, neutral. The system runs at a speed which varies from 14 to 25 Hz on a 200MHz R4400 indy, according to whether or not parameter extraction and mouth detection are activated in addition to tracking.

Information Overlay in Smart Rooms: Unbuilt Ruins

Following on the experienced earned with the previous project, in January 1999, we created another museum exprerience which satisfies both the needs of the curator -- who needs to be able to feature a great quantity of material in a limited physical space -- and those of the viewer -- who benefits from a coherent narration of the spatially fragmented objects on display. Prompted by architect Kent Larson and designer Ron MacNeil, we designed an exhibit space which shows a variety of architectural designs by the influential XXth century american architect Louis Kahn. The exhibition interactively features radiosity-based, hyper-realistic computer graphics renderings of 8 unbuilt masterworks by Luois Kahn. It is housed in a large room, approximately 50'x50', and contains in the center a square table above which are mounted a camera and a projector pointing downwards. The table is surrounded by 4 large back projection screens, parallel to the 4 sides, and situated a few feet away to provide an ideal viewing surface. The table contains the projection of the image of one of the eight projects' schematic floor plan and 3D printed physical models of each of the projects, on bases 3"x3", located two to a side, at the perimeter of the table [Figure 2]. Each model contains a digital tag which identifies it.





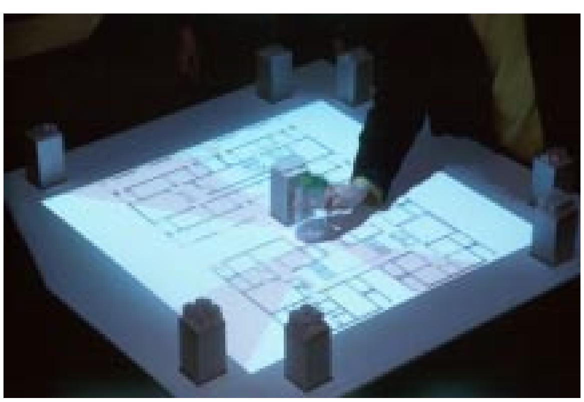
Figures 2 and 3. Schematic diagram of the Unbuilt Ruins exhibit and projected views from one hot-spot from the architectural plan of the Meeting House of the Salk Institute, La Jolla, California, 1959-65.

A combined use of real time computer vision techniques, grown from a long tradition of work in the field at MIT [Wren, 1997], and tag sensing, allows the visitor to explore the eight projects inside the museum space. The visitor selects the project to view by placing the corresponding 3D model in the 3"x3" depression in the center. A commercially available tag reader placed under the table, below the depression, identifies the 3D model with the attached tag. When the reader finds that a physical model has been placed in the central receptor, a square image of its schematic floor plan is projected onto the table from the ceiling projector. At the same time, conceptual images (slides) of the selected buildings are projected on the 4 rear projection screens.

The schematic plan projected on the table contains graphics symbols indicating camera positions for the selected projects (hot-spots). A standard color camera aimed at the a table, linked to a computer running a real time computer vision system, tracks a color-tagged cylindrical object (active cursor). A view is selected when the viewer places the active cursor over a camera position symbol. When a view is selected the side screens shows a rendering of what a visitor would see if they were standing inside the physical construction in the location determined by the position of the symbol on the schematic plan and looking towards the direction indicated in the position symbols [Figures 3,4,5,6,7]. If no view is selected within one minute, the system automatically scrolls though all views of the selected project until a new view or architectural plan is chosen.

With the eight unbuilt projects presented in the exhibition, Kahn developed and tested ideas quite different from the modern architecture of his time: a configuration of space as discrete volumes, complex ambient light and deep shadow, a celebration of mass, the use of materials with both modernist and archaic qualities, monumental openings uncompromised by frames, and a concept of "ruins wrapped around buildings." The exhibition attempted to shed light on how Kahn came to make the architecture he did, by looking in depth at projects left unbuilt.





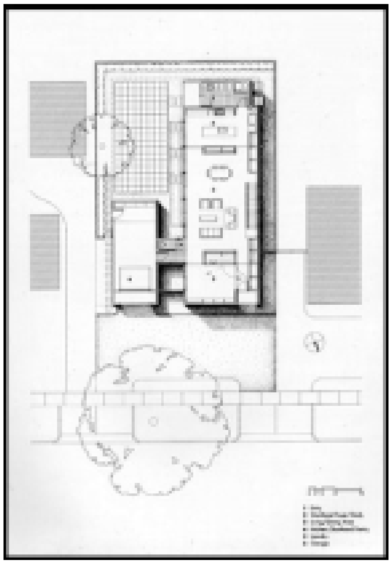

Figures 4 and 5. Images of the Unbuilt ruin Exhibit taken at the Compton Gallery at MIT.
Figures 6 and 7. Visitors placing the active cursor on a hot spot on the map and discussing the displayed views.

Since these eight projects were incomplete and schematic, the design of each had to be resolved by extrapolating from archival material at the Kahn collection of the Architectural Archives at the University of Pennsylvania. A series of three dimensional digital stage sets were created, and photographs from Kahn's Salk Institute, Exeter Library, and the Yale Center for British Art were used to develop high-resolution texture map montages for each unique surface plane. Essential to the accurate interpretation of these projects was the physically accurate treatment of light. Images were rendered with Lightscape, a radiosity-based renderer, to capture the complex inter-reflections of ambient light in space.

The exhibit was open for one month at the Compton Gallery at MIT in February 1999, and later in the year for another month at the University of Pensilvania's gallery exhibit space. It raised enthusiastic feedback from visiting people, architects, curators, and classrooms.

Unbuilt Ruins was also used as a model for a MOMA exhibit called "The Un-private House" curated by Terence Riley and shown at MOMA in the summer of 1999. We built an initial prototype in which we adapted the original setup we had at the MIT Compton Gallery to display architectural plans of houses projected on a table. A custom tag system, developed at MIT, was used to select which architectural plan would be explored by the visitor at one time. The tags were embedded inside postcards of the shown houses, and when placed on the table, they would be detected, and the corresponding floor plan would be projected. The computer vision system tracked three objects, in the form of small dolls. When the visitors placed an object over a hot-spot on the plan, they triggered the system to display photographs showing what would be seen from that point in the house, as well as text which would function as a guide of the exhibit. The three dolls represented the point of view of the curator, the architect of the displayed house, and the owner. Each of these dolls placed on a hot-spot told the story of the exhibit from a slightly different viewpoint. The communication between the tag reader, the computer vision software, and the graphical display software was networked such that more communicating tables with displays would be able to exchange information, or be aware of the behavior/exploration of the other visitors of this exhibit [Figures 8,9,10].





Figures 8,9,10. Floor plan and views of one of the chosen houses for the prototype of "The Un-private house" interactive exhibit at MOMA.

Spatialized Interactive Narrative: Wearable City

In the context of the SIGGRAPH99 Emerging Technology exhibit [Sparacino, 1999b] we presented a wearable setup that can easily be customized for museum use. Wearable City is the mobile version of a 3D WWW browser we created, called "City of News." City of News [Sparacino, 1997b] fetches and displays URLs so as to form skyscrapers and alleys of text and images which participants can visit as if they were exploring an urban landscape of information [Figure 11]. By associating information with geography people can better remember what path has led them to a particular web page, and recall its content, or find it again on the web, based on its spatial location. In addition, the three dimentional architectural landscape offers a context to associate related information by its geographical position. "Wearable City" is the mobile version of the above. It shows a 3D WWW browser with a SIGGRAPH map superimposed to the real world on the participant's head mounted display. The sensing system is made by infrared transmitters/receivers distributed around the physical location of the exhibitor. When the wearer is near an infrared tagged location Wearable City shows a web page of that contributor from the right location in the map. The receiver is a small infrared diode situated appropriatedly on the headmounted display and connected to a small transducer which communicates with the wearable computer's serial port. The computer is made by a thin laptop's motherboard placed in the back of the jacket. We used conductive thread to sew a small keypad on the jacket's sleeve so that people could request more information or browse by pressing one of the VCR-like keys [Figure 12].

We also described possible applications of wearable computing to performance in [Sparacino, 1997c]. In this case a similar setup is used as an augmented storytelling device which allows the wearer to see the actual perfomance with the added context/explanation/virtual characters superimposed and coordinated with the performer's movement though the viewer's headmounted display. We are currently implementing a real-time computer vision system which can run on a separate processor inside the wearable computer so as to adapt the same device for museum use.

Figure 11. Screen shot of City of News. Figure 12. MIT Media Lab's departmental head, Professor Alex Pentland, modeling a wearable computer with a display embedded in the glasses.

We believe that these personalized systems offer a great platform of experimentation for personalized museum tours in which the spatial location of the wearer and what they are looking at, as detected by the wearable computer, help construct a narrative and a computerized museum guide based on, and edited by, the actual physical path of the public along the exhibit. This wearable computer guide is only complementary to and does not replace current museum guides. Its purpose is to present images, animations, text, and audio so as to integrate the fragmemted museum experience in a coherent whole, as if the wearer was driving an interactive documentary of the exhibit by choosing his path in the museum. The system can also make suggestions about possible physical paths for the wearer to follow along the exhibit, based on context and personal profile. Therefore the interactivity can be used not only to retrieve information "on the fly" and explain the exhibit, but also to direct and guide the wearer along points of interest and coherent sequencing. The ability to connect and fetch information from the web and to present it in a spatialized way is also an important element for the public to understand and be situated in the augmented experience. It facilitates the work of the curator as no additional content needs to be prepared other than what is already on the web. The wearable computer transforms the experience of the museum visit into a tour in a multimedia-augmented memory palace [Boyer, 1996] in which the memory of the visit and the learning experience are greatly enhanced by the technological innovation.

Conclusions

We have identified two areas of technological intervention which contribute to creating an engaging experience for the public. These are : Information Overlay in Smart Rooms (adding technology to the museum space) and Spatialized Interactive Narrative with Smart Clothes (adding technology to the visitor). We described applications and tools to contruct these interactive experiences for three projects: Responsive Portraits, Unbuilt Ruins, and Wearable City. These project respond to current needs of museums to attract a larger public and to offer people a memorable and fun learning experience. Through information overlay it is possible to compress in the limited space of the museum a larger portion of material relevant to the exhibit and to the public. The interactivity of the setup, in addition to the projections, stimulates people to explore and construct meaning out of the displayed artwork. The wearable computer-driven museum guide, based on the wearable city project, allows to integrate all additional information relative to exhibit along the wearer's path in a personalized way.

Acknowledgements

Rich DeVaul and Ari Benbasat designed the infrared emitter/transmitter system used in Wearable City. Christopher Wren provided technical assistance for the same project. Rehmi Post, Maggie Orth, Chris Metcalfe, Pamela Mukerji, and Peter Russo worked on the condutive-thread-based keypad for Wearable City. Ali Azarbayejani contributed to the computer vision system used at the MIT Compton Gallery. Thomas Minka was a technical consultant for the same project. MIT Professor Neil Gerschenfeld coordinated efforts towards the making of the interactive exhibit at MOMA. The tag system used in this exhibit was developed by Richard Fletcher and Olufemi Omojola. E'beth Haley helped with laser cutting and embedding the tags in the photographic prints of the displayed houses.

References

Boyer, C. *The City of Collective Memory*. MIT Press, Cambridge, MA, 1996. pp 133.

Hein, George. *Constructivist Learning Theory*, ICOM/CECA, 1991. Apperars in: Developing Museum Exhibitions for Lifelong Learning, GEM, 1996.

Klein, Larry. *Exhibits: Planning and Design*. Madison Square Press, New York, 1986. pp70-71.

Oliver, N. and Pentland, A. *Lafter: Lips and face tracking*. In CVPR97. IEEE Computer Society, 1997.

Sparacino, F., Oliver, N., Pentland, A., Davenport G. *Responsive Portraits*, The Eighth International Symposium on Electronic Art, Chicago, IL, September 22-27, 1997a.

Sparacino, F. et al., *City of News*, Ars Electronica Festival, Linz, Austria, 8-13 September 1997b.

Sparacino F., Pentland A., Davenport G. *Wearable Performance*. International Symposium on Wearable Computers, Cambridge, Massachusetts, October 13 - 14, 1997c.

Sparacino F., Davenport G., Pentland A. *Media Actors: Characters in Search of an Author*. ICMCS, Florence, Italy, 1999a.

Sparacino F., et al. *City of News*. Appears in: Conference Abstracts and Applications. SIGGRAPH 1999b.

Starner, T. et al. *Augmented Reality Through Wearable Computing*. Appears in: Presence, Special Issue on Augmented Reality, 1997.

Wren, C. et al. *Perceptive Spaces for Performance and Entertainment: Untethered Interaction using Computer Vision and Audition*, Applied Artificial Intelligence Journal, Special Issue on Entertainment and AI/Alife, Vol. 11, Number 4, June 1997.