Narrative Spaces:
bridging architecture and entertainment via interactive technology

Flavia Sparacino
MIT Media Lab
flavia@media.mit.edu

Abstract

Our society's modalities of communication are rapidly changing. Large panel displays and screens are being installed in many public spaces, ranging from open plazas, to shopping malls, to private houses, to theater stages, classrooms, and museums. In parallel, wearable computers are transforming our technological landscape by reshaping the heavy, bulky desktop computer into a lightweight, portable device that is accessible to people at any time. Computation and sensing are moving from computers and devices into the environment itself. The space around us is instrumented with sensors and displays, and it tends to reflect a diffused need to combine together the information space with our physical space. This combination of large public and miniature personal digital displays together with distributed computing and sensing intelligence offers unprecedented opportunities to merge the virtual and the real, the information landscape of the Internet with the urban landscape of the city, to transform digital animated media in storytellers, in public installations and through personal wearable technology. This paper describes technological platforms built at the MIT Media Lab, through 1994-2002, that contribute to defining new trends in architecture that merge virtual and real spaces, and are reshaping the way we live and experience the museum, the house, the theater, and the modern city.

1. A new architecture for the information society

Architecture is no longer simply the play of masses in light. It now embraces the play of digital information in space.
William J. Mitchell, Dean of MIT's School of Architecture and Planning, in: e-topia, pg. 41

Our daily lives are characterized by our constant access-to and processing-of a large quantity and variety of information. In the last decade, the rapid diffusion of the information superhighways, the amazing progress in performance and processing power of today's computers, paralleled by a drop of their cost, has determined a profound transformation of western world societies. In addition to being transmitted by the traditional media, such as television, radio, the newspaper, the book, the telephone, the mail, information is conveyed to us in electronic form by the home or office computer, public billboards, private hand-held PDAs, cellular phones, and soon even by our wrist-worn watch and clothes. The potential offered by the rapid and efficient exchange of data, globally, between individuals and organizations, delineates new social, economic, and cultural models based on the exchange of knowledge. The information society is defined by the primary role of information, such that power and growth are associated to our ability to receive, store, process, and transmit information instantaneously.

Screens are everywhere, from the large billboards commonly embedded in the contemporary urban city-scape, to the video walls which welcome us in the entry-hall of corporate headquarter buildings, the desktop computer monitor in our home, the PDA in our pocket, or the tiny private-eye screens of wearable computers [1] [figures 1 and 2].

Figure 1. Aerial view of the Shibuya district in Tokyo showing of a large variety of LCD displays and animated screens distributed in the city-scape

Figure 2. Fashionable wearable display projecting an image on the viewer's eye through the private-eye

We split our daily activities between the real and the digital realm. More and more frequently we go shopping virtually on the internet, go to the library on the internet, meet and chat with people over the internet, manage our finances, play, plan our entertainment, and even date through the information superhighways. These profound transformations of our life-style demand a new architecture that supports these new modalities of communication and living. Space needs to be redesigned to favor information exchange through video walls, across portable devices, and private-eye displays. Computation and sensing will move from computers and devices into the environment itself. The space around will increasingly be instrumented with sensors and displays, to reflect the diffused need to combine together the information space with our physical space. "Augmented reality" and "mixed reality" are the terms most often used to refer to this type of media-enhanced interactive spaces.

Several scenarios arise from the encounter and blending of media design and architectural disciplines [2]. Yet architecture's agenda needs to encompass not just the design of new media- and information-enhanced spaces, but should also extend itself to investigate natural modalities of human-computer interaction which facilitate communication through cyberspace [3][4]. Indeed still today, the interfaces available for people to communicate electronically are limited to the primitive and low-bandwidth keyboard and mouse attached to desktop or mobile computers.

This paper describes technological platforms built at the MIT Media Lab, through 1994-2002, that contribute to defining new trends in architecture that merge virtual and real spaces, and are reshaping the way we live and experience specifically the museum, the theater, the house, and the modern city. These platforms are grounded in research in real-time computer vision based human-computer interaction, as well as sensor fusion, and mathematical modeling of perceptual intelligence. Our focus is in narrative spaces, that is sensor-enabled, people-driven, media-augmented interactive indoors or outdoors spaces that convey information as (audio-)visual micro-stories or more simply as three dimensional information landscape visualizations.

This paper also wishes to highlight that, in the author's view, it is not a coincidence that the contribution to the new architecture for the information society illustrated in this paper comes from within a research group with a strong technical and scientific background, combined with creativity, a sense for architecture and space design, knowledge of filmmaking, and attention towards social change and needs. Specifically, the author's knowledge of statistical modeling, image processing, Bayesian networks, has enabled her and her collaborators to build a whole series of platforms and experiment with various scenarios of interactive spaces, to iterate the design process several times as needed, to progressively adapt the science and technology of interaction to the emerging design issues, so as to bring both the architectural and the technical aspects of the presented projects to the desired level of quality and performance.

2. Enabling Technologies

The problem, in my opinion, is that our current computers are both deaf and blind: they experience the world only by way of a keyboard and a mouse. Even multimedia machines, those that handle audiovisual signals, as well as text, simply transport strings of data. They do not understand the meaning behind the characters, sounds, and pictures they convey. I believe computers must be able to see and hear what we do before they can prove truly helpful. … To that end, my group at the Media Laboratory at the Massachusetts Institute of Technology has recently developed a family of computer systems for recognizing faces, expressions, and gestures. The technology has enabled us to build smart rooms … furnished with cameras and microphones that relay their recordings to a nearby network of computers. The computers assess what people in the smart rooms are saying and doing. Thanks to this connection, visitors can use their actions, voices and expressions – instead of keyboards, sensors or goggles – to control computer programs, browse multimedia information or venture into realms of virtual reality.
Alex Pentland, head of the Perceptual Computing Group at the MIT Media Lab, in: Scientific American, April 1996, pg 68 and 71.

The architect who wishes to reshape our surrounding space and body, and transform them into technology-augmented devices for information exchange and communication needs: sensors that are reliable and robust, and (mathematical) modeling tools which allow the system to understand the public's intentions and coordinate a narration. Information authoring tools need to be able to take input from people, and deliver a (personalized) story articulated not only over time but also over space. Just as humans use vision as their main sensing modality to perceive and understand their surroundings, the narrative spaces here presented use predominantly real time computer vision to locate people in space and understand what they are doing. This section offers a brief overview of the technologies and requirements the author believes are important to enable people-driven narrative space design. These are the technologies utilized to build the narrative spaces presented in the following section.

2.1. Sensing

Applications such as unencumbered virtual reality interfaces, performance spaces, and information browsers, all have in common the need to track and interpret human action. The first step in this process is identifying and tracking key features in a robust, real time, and non intrusive way. Computer vision is a tool capable of solving this problem across many situations and application domains. By use of real-time computer vision techniques [5][6][7] we are able to interpret the people's *posture*, *movement*, *gestures* , and *identity*. Wren and others [5] have shown that a system which has an image based two dimensional description of the human body as a set of adjacent color regions (head/torso/hands/feet), a MAP estimator for color pixel classification, and a Kalman filter applied to the features to track, is a very powerful tool to mathematically and computationally describe the human body in motion in real time. Similar maximum likelihood statistical approaches are also effective in stereo vision tracking to locate body features more accurately in 3-D space, when pointing direction, and accurate depth information are needed Hidden Markov Models and more recently Bayesian networks, have been successfully used to classify human movements and gestures [8][9].

2.2. Robustness of multimodal perception

Robust sensing is the premise for the correct interpretation of the user's intention. Mono-sensor applications which rely on one unique sensor modality to acquire information about people are brittle and prone to error. For how well that one sensor works individually, whether that be a camera, or a radar, or an electric field sensor, it only provides the system with a single view of what is going on. In order for a body driven interactive application to offer a reliable and robust response to a large number of people on a daily basis, as needed in a museum, or to meet the challenges of the variable and unpredictable factors of a real life situation, we need to rely on a variety of sensors which cooperate to gather correct and reliable measurements on and about the user. Cooperation of sensor modalities which have various degrees of redundancy and complementarity can guarantee robust, accurate perception [10]. We can use the redundancy of the sensors to register the data they provide with one another. We then use the complementarity of the sensors to resolve ambiguity or reduce errors when an environmental perturbation affects the system.

2.3. Context-based data interpretation

To make good use of reliable measurements about the user, we need to be able to interpret our measurements in the context of what the user is trying to do with the digital media, or what we actually want people to do to get the most out of the experiences we wish to offer. The same or similar gesture of the public can have different meanings according to the context and history of interaction. For example the same pointing gesture of the hand can be interpreted either as pushing a virtual character, or more simply, as a selection gesture. In a similar way, the system needs to develop expectations on the likelihood of the user's responses based on the specific content shown. These expectations influence in turn the interpretation of sensory data. Following on the previous example, rather than teaching both the user and the system to perform or recognize two slightly different gestures, one for pushing and one for selecting, we can simply teach the system how to correctly interpret slightly similar gestures, based on the context and history of interaction, by developing expectations on the probability of the follow-on gesture. In summary, our systems need to have a user model which characterizes the behavior and the likelihood of responses of the public. This model also need to be flexible and should be adaptively revised by learning the user's interaction profile [11][18].

2.4. Narrative engines

It is difficult to produce compelling applications simply by direct mapping of sensor measurement inputs with digital media output. While this strategy may work for very simple interactive environments, it is not effective to orchestrate digital information effectively.

Many current interactive systems are defined by a series of couplings between user input and system responses. The problem with these systems is that they are often repetitive: the same action of the participant always produces the same response by the system. Alternatively, most existing CDROM titles are scripted: they sequence micro-stories in multi-path narrative threads. While the content presentation in these applications tends to be more engaging, they often impose a rigid interaction modality and become boring after a while. The participant's role is confined to clicking and choosing the sequencing of the narrative thread without real engagement or participation in the narrative. In order to create compelling narrative spaces we need to be able to simulate encounters between the public and the digital media acting as a character. To accomplish this goal we need to be able to model the story we wish to narrate in such a way that it takes into account and encompasses the user's intentions and the context of interaction [11]. Consequently the story should develop on the basis of the system's constant evaluation of how the user's actions match the system's expectations about those actions, and the system's goals.

These guidelines have driven the research described below on smart rooms [3][figure 3], smart desks [4][figure 4], and wearable computing, with special emphasis on interactive information presentation, digital storytelling, and cultural communication.

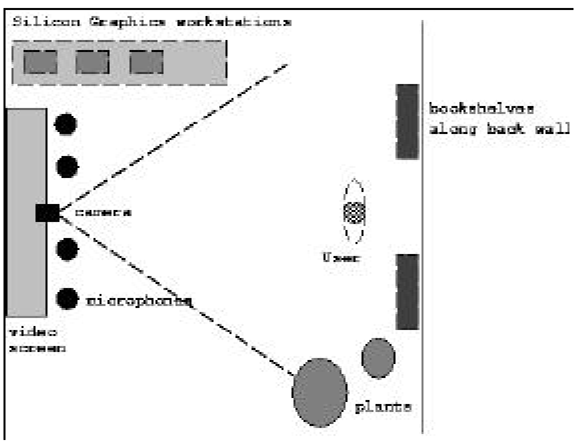

Figure 3. Smart r oom setup including a single camera for body tracking, a microphone array for speaker identification, and a large projection screen for content viewing

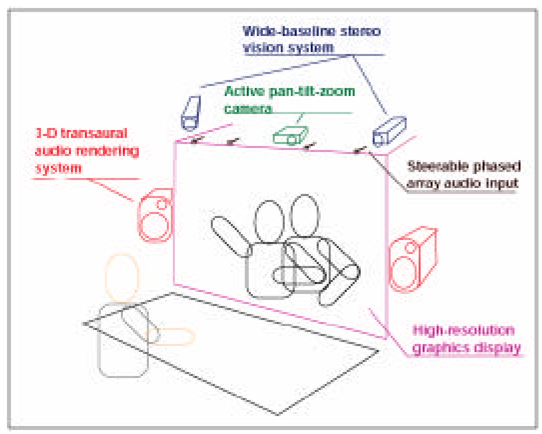

Figure 4. Smart desk setup with stereo vision, active camera for face expression recognition, microphone for speech recognition, and a large screen

3. From theme parks to culture parks: a family of narrative spaces

Now by embedding intelligence and interconnectivity in material products and creating systems of tags and sensors… we can construct spatially extended smart spaces from collections of interacting smart objects. Real desktops, rooms, and other settings – rather than their electronically constructed surrogates – can begin to function as computer interfaces … As a result, our actions in physical space are closely and unobtrusively coupled with our actions in cyberspace. We beco me true inhabitants of electronically mediated environments rather than mere users of computational devices.

William J. Mitchell, Dean of MIT's School of Architecture and Planning, in: e-topia, pg. 43

This section briefly describes a full generation of interactive spaces and platforms of interaction the author built at the MIT Media Lab through 1994-2002. While the reader will find more detail on each of these in the bibliography, the purpose of the following presentation is to illustrate the family of body driven interfaces for smart spaces in its entirety. This will serve as a basis to for the next section in which the author comments upon how the collection of such interaction platforms constitute a set of paradigmatic setups leading towards a renewed mixed-reality architecture for the information society.

3.1. Video Wall and interactive carpet: MetaSpace

MetaSpace [12] [figures 5 and 6] is large scale installation which uses two projection surfaces: one vertical and one horizontal. The horizontal surface is a large map projected on the floor. People physically walk onto different locations of the floor map and trigger consequently the front surface projection to show the correspondent visual and auditory information. One can see the floor map projection like a mouse pad, the person walking onto the map like the mouse driving the system, and the vertical large projection screen as the computer display. A computer-controlled infrared camera detects people's presence and location on the floor map, as well as their pointing gestures (selection gestures) towards the objects displayed on the vertical screen. MetaSpace installations have been shown at Ars Electronica 98, SIGGRAPH 99, and are currently being installed in a museum setting. It is the extension of the author's previous work on Hyperplex [13].

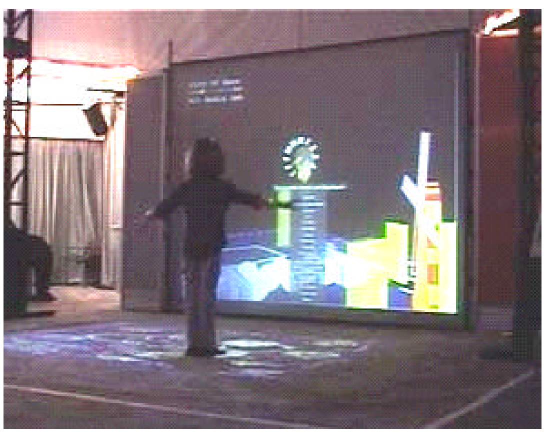

Figure 5. MetaSpace at the SIGGRAPH 99 Millennium Motel show.

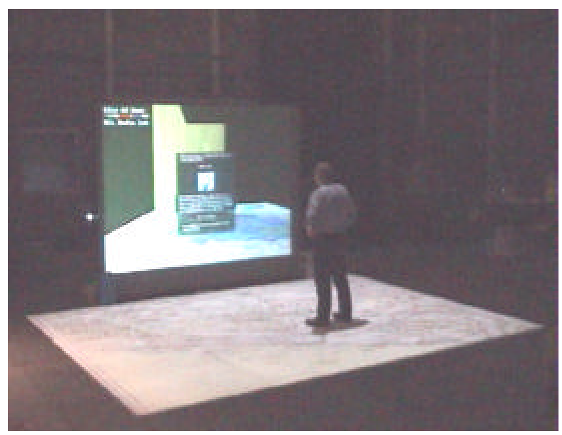

Figure 6. MetaSpace in the MetaCity Sarajevo setup, built for Ars Electronica 98

3.2. Presentation Table

The presentation table is an interactive horizontal display which narrates a story guided by the position of the objects on it. It is a playful interface which allows the public to explore the non linear contents of a book, CDROM, or audiovisual story [figure 7]. The table uses tag sensing and computer vision to locate and track physical content-selector objects on the table. Typically two types of objects are needed: a chapter-selector and an item-selector object. Visual narrative maps are associated to each chapter [figure 8]. When people place a chapter-selector object in the marked space in the center table, the visual table of contents for that chapter is projected on the table. When people select to learn more about one of the items on the visual content map, the corresponding text, image, audio, or video are displayed. Examples of the presentation table are: Unbuilt Ruins, an exhibit space, orchestrated by the table, which shows a variety of architectural designs by the influential XXth century

American architect Louis Kahn [14] and MOMA's Unprivate house exhibit's interactive table, modeled on the previous one.

Figure 7. The Unbuilt Ruins presentation table featuring projects by architect Louis Kahn

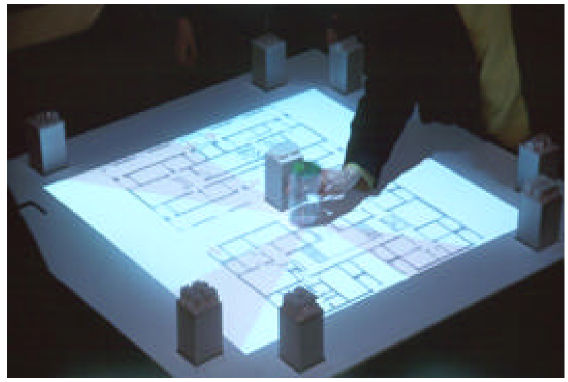

Figure 8. Chapter-selector objects around the table and item-selector object (green)

3.3. Gesture browsing through the information city: City of News

City of News is an immersive, interactive web browser that reads web pages from the internet and embeds them inside a three dimensional city-scape. It fetches and displays URLs so as to form skyscrapers and alleys of text and images through which the user can "fly" using body movements. Known cities' layout, architecture, and landmarks are given as input to the program and are used as organizing geometry and orientation cues [figure 9]. This virtual internet city grows dynamically as new information is loaded: following a link causes a new building to be raised in the district to which it belongs, conceptually, by the content it carries, and content to be attached onto its "façade". By mapping information to familiar places, which are virtually recreated, City of News stimulates in its users association of content to geography. The spatial, urban-like, distribution of information facilitates navigation of large information databases, like the Internet, by providing the user with a cognitive spatial map of data distribution [15]. To navigate this 3-D environment, users sit in front of a large screen and use hand gestures to explore or load new data. Pointing to a link will load the new URL building. The user can scroll up and down a building by pointing up and down with either arm. Side-pointing gestures allow users to navigate along an information path back and forth. Raising both arms drives the virtual camera above the internet city and gives an overall color-coded view of the urban-like information distribution [figure 10] [16].

Figure 9. Aerial view of City of News. Example based on the map of the city of Stuttgart

Figure 10. Gesture-driven navigation through the City of News in the smart desk

3.4. Telepresence: Virtual Studio

Virtual Studio is a 3-D set inside which people can meet and interact among themselves or with other 3D characters. Thanks to advanced computer vision techniques, as in a magic mirror, participants see their full body video image composited in 3-D space – without the need for a blue screen background [17] [12] [figure 11]. Such a setup, requiring only a camera, a computer with a graphic card, and a fast network connection per participant, allows people to have access to real time compositing technology today available only in high end costly production studios. In addition to compositing the partcipants' video image inside a 3-D set the system uses gesture recognition techniques to give networked participants the added feature to modify, activate, and manipulate the elements of the 3D set in which they are immersed. The participant's image is subjected to all graphics transformations that can apply to graphical objects, including scaling. According to the participant's position in the space, his/her image occludes or is occluded by virtual objects in respect to the 3-D perspective of the virtual scene. Multiple people can connect from remote locations, therefore such setup can be used for collaborative storytelling, visual communication from remote locations, or game playing.

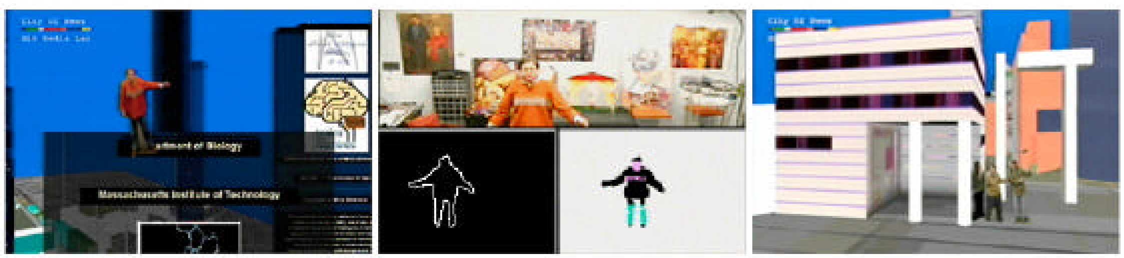

Figure 11. Virtual Studio: the user is composited in real time inside a 3-D set. In the last panel right: multiple people from remote locations are composited altogether in front of a virtual model of the MIT Media Lab.

3.5. Responsive Portraits

Responsive Portraits [18] [figure 12] challenge the notion of static photographic portraiture as the unique, ideal visual representation of its subject. A responsive portrait consists of a multiplicity of views – digital photographs and holographic 3D images, accompanied by sounds and recorded voices – whose dynamic presentation results from the interaction between the viewer and the image. The viewer's proximity to the image, head and upper body movements elicit dynamic responses from the portrait, driven by the portrait's own set of autonomous behaviors. This type of interaction reproduces an encounter between two people: the viewer and the character portrayed. In this installation the whole notion of "who is watching who" is reversed: the object becomes the subject, the subject is observed.

Figure 12. Responsive Portraits installation

3.6. Museum Wearable and Wearable Cinema/Wearable City

The museum wearable is a wearable computer which orchestrates an audiovisual narration as a function of the visitor's interests gathered from his/her physical path in the museum and length of stops [19] [figures 13, 14]. The wearable is made by a lightweight and small computer that people carry inside a shoulder pack. It offers an audiovisual augmentation of the surrounding environment using a small, lightweight eye-piece display (often called private-eye) attached to conventional headphones. Using custom built infrared location sensors distributed in the museum space, and statistical mathematical modeling, the museum wearable builds a progressively refined user model and uses it to deliver a personalized audiovisual narration to the visitor. This device will enrich and personalize the museum visit as a visual and auditory storyteller that is able to adapt its story to the audience's interests and guide the public through the path of the exhibit. Wearable City extends the previous application outdoors, to the city, using GPS location sensing. Wearable cinema explores the idea of a movie distributed in space and time and whose segments are triggered by the viewer's path in a theme park or city quarter [20].

Figure 13. The Museum Wearable worn by a visitor at the MIT Museum

Figure 14. The Museum Wearable: an explanatory movie clip plays as the visitor approaches a printed photograph of an artwork on display

3.7. DanceSpace and Improvisational Theater Space

When applied to the stage, our work augments the expressive range of possibilities for performers and stretches the grammar of the traditional arts rather than suggesting ways and contexts to replace the embodied performer with a virtual one [21]. In dance, the author has conducted research towards musical and graphical augmentation of human movement. This has led to DanceSpace: a stage in which music and graphics are generated on the fly by the dancer's movements. A small set of musical instruments is virtually attached to the dancer's body and generates a melodic soundtrack in tonal accordance with a soft background musical piece. Meanwhile, the performer projects graphics onto a large backscreen using the body as a paint brush [figure 15]. In theater, the focus done work in gesture, posture, and speech augmentation. In Improvisational TheaterSpace a human actor can be seen interacting with his own thoughts in the form of animated expressive text projected on stage. The text is just like another actor able to understand and synchronize its performance to its human partner's gestures, postures, tone of voice, and words [figure 16].

Figure 15. DanceSpace

Figure 16. Improvisational Theater Space

The next section attempts to draw scenarios on how the interaction platforms and instrumented spaces described above will likely influence and hopefully enrich the way we learn, explore, entertain, live, and specifically on how the traditional spaces that house such experiences are being reshaped to accommodate for the social changes in act.

4. Narrative Spaces

We will characterize cities of the twenty-first century as systems of interlinked, interacting, silicon- and software-saturated smart, attentive, and responsive places. We will encounter them at the scale of clothing, rooms, buildings, campuses, and neighborhoods, metropolitan regions, and global infrastructures.
William J. Mitchell, Dean of MIT's School of Architecture and Planning, in: e-topia, pg. 68

People experience their lives as a narrative. Amongst cognitive psychologists, Jerome Bruner stressed the importance of story, as the means which structures our perception and communication [22]. He reminded us that thinking cannot be reduced to mere information processing and sorting into categories and that narrative is our main instrument of making-meaning, the embodiment of culture, communication, and education. The history of architecture offers innumerable examples of places which embed and narrate a story through their spatial layout and décor. By looking at the sequence of floor plans of historical buildings through the centuries, from the Greek temple, to the Roman church, the medieval dome through today, we understand how a rectangle, a circle, a cross, or other more complex figures, transmit a message through the centuries. This message is a story about how people through times relate to life, nature, and spirituality.

The new information society architecture can more explicitly embed stories and information in its structure, thanks to the tools offered to us by the digital revolution. Following are a few scenarios relative to changing places: the museum, the city, the house, the theater, that draw from the examples presented in the previous section.

4.1. The museum

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. Technology can help construct a coherent narrative of an exhibit for the visitor by creating experiences in which the objects on display narrate their own story in context [14]. Using interactive techniques embedded in the physical space museums can present a larger variety and more connected material in an engaging manner within the limited space available. They can also enrich and personalize the visit with wearable computers which act as a visual and auditory storyteller that guide the public through the path of the exhibit. The presentation tables will be used a playful interface for the public to access and explore the body of facts, content, and stories of the exhibit. The MetaSpace can be installed as an introductory immersive cinema space. Responsive portraits show artwork that reacts to how people approach it and is capable of explaining its origin and making. All these systems enhance the memory of the visit and help build a constructivist-style learning experience for the public.

4.2. The city

In the last decade, the architectural landscape of the contemporary city has been undergoing a profound transformation. When we drive from one city to the next, when we wander in the city center, when we drive to the airport, large billboards with printed images or large LCD screens show us publicity, entice us, talk to us, inform us about our banking, clothing, insurance, entertainment, eating options. These billboards are embedded in the city architecture, and have become integrated in the visual appearance and profile of the city-scape. Most of us are no longer surprised to see building-size images decorate facades juxtaposed between skyscrapers or even historical buildings. This phenomenon has invested modern metropolis, notably Tokyo [figure 17], Los Angeles, along "The Strip", in Sunset Boulevard, a route which joins the city with Santa Monica [figure 18], New York, particularly around the Times Square area [figure 19] Yet it does not spare more classical European cities such as Milan [figure 20], Paris, or Berlin. Especially at night, Times Square and its surroundings evoke science fiction cities a la Blade Runner: the city becomes a spectacle, a visual symphony, a multi-screen receptor and display of information of all sorts. Only recently artists [23] and architects [www.streetbeam.com] have started to conceive and prototype experiences where the information flow involves communication between city billboards, passerbiers with PDA devices, and people at home connected to the internet. MetaSpace is in the process of being scaled to an outdoors experience where people walking on the carpet of light projected on a central sidewalk orchestrate images and information presentation embedded in the surrounding buildings [24]. The wearable cinema and wearable city projects add new flavor to existing research on PDA-guided tourism and entertainment.

Figure 17. Tokyo: Ginza.

Figure 18. Los Angeles: Sunset Boulevard.

Figure 19. New York: Times Square.

Figure 20. Milan: next to the city center.

4.3. The house

Compared to the enormous progress in miniaturization, computing power, and cost reduction of electronic equipment, very little progress has affected the house in the last few decades. Yet most homes today include an area, or table, with a computer that family members use for various tasks, most of them related to information browsing over the internet, communicating with friends, planning entertainment, and shopping. Soon however the current living room scenario in the home will change. In one corner a city-of-news-like gesture-browser will allow people to navigate information on the Net, represented as an evolving 3-D information city that appears in a large LCD "window over cyberspace". In another corner people could activate the virtual studio and see their video image composited, together with the ones of their remote correspondents, inside a chosen three dimensional chat room, in a virtual café. The hyperplex interface [13], together with a voice recognition system, interprets users' commands at need. In the adjacent room the kids are jumping around and making music with DanceSpace, while in the home office the presentation table assists the user with their next project.

4.4. The theater

Dance and theater performances, operas, musicals, cinema projections, talk shows, all take place in traditional spaces made by a stage on one end and the audience on the other end of the space. The lights go off and the spectator is typically passive for the duration of the show. Ongoing work by the author leads towards three different scenarios: one which augments and extends the expressive range of traditional performance, one which gives the public contextual information and explanation about what's happening on stage, and a much different one which involves technological support to bring traditional performance from the traditional stage-audience presentation setting, to the street or other unconventional performance places that allow for a more direct contact of the performer with the audience. DanceSpace, Improvisational theater space, or similar technologies will be integrated in the stage and will help create media-actors, a collection of digital expressive text, images, and video which co-act with human performers thanks to their sensors, perceptual intelligence, and context modeling abilities, as described in section 2. The presence of a public with wearable computers will also stimulate new kind of performances in which the public plays a more active role. In theater, for example, the author is interested in exploring how point-of-view transforms our perception of reality. The wearable can be used as a "semantic lens" to offer audience members a new, transformed interpretation of the story told by the performer. This lens provides a view-dependent reality augmentation, such that different people observing the show from different areas of the theater, are led to a different interpretation of the story told by the actors. Wearable street performance [25] will also contribute to change location and modalities of the contemporary performance arts.

These scenarios show how, in a non distant future, the prototypes described above will likely influence the changing design of traditional living, learning, and entertainment spaces. Other possible scenarios involve the classroom, the office, the retail store, cafes, etc. In the home the reading room will be connected to an electronic book which has the touch and feel of the printed book we are familiar with. The kitchen will be made of interconnected smart appliances which assist people in cooking. The architect's agenda in the next few years will have to include, develop, and work towards turning these types of scenarios into reality.

5. Conclusions: new challenges for today's architects

Traditional urban patterns cannot coexist with cyberspace … This will redefine the intellectual and professional agenda of architects, urban designers, and others who care about the spaces and places in which we spend our daily lives … This new agenda separates itself naturally into several distinct levels … We must put in the necessary digital telecommunications infrastructure, create innovative smart places from electronic hardware as well as traditional architectural elements, and develop the software that activates those places and makes them useful… To pursue this agenda effectively, we must extend the definitions of architecture and urban design to encompass virtual places as well as physical ones, software as well as hardware, and interconnection by means of telecommunications links as well as by physical adjacencies and transportation systems.
William J. Mitchell, Dean of MIT's School of Architecture and Planning, in: e-topia, pg. 8

This paper described a series of sensor-enabled, media-augmented, people-driven narrative spaces and highlighted the role of the technologies that are key in their conception and making. Scenarios are given for how the variety of such spaces and related hardware and software platforms can be used in, and influence the way we experience more traditional "narrative spaces" such as the museum, the theater, the house, and the city.

Another contribution of the paper is also to underline that technology is not simply hardware of software that the space designer and the media artist add to their projects to make them work. It is really not sufficient to wait for technologists to develop new modalities of interaction and man-machine communication in their laboratories, to later incorporate these in space design, as software that one buys at the store. The architecture of the information society is truly driven and informed by technology, which in turn shapes the architectural thinking and project development. Unless today's architects are able to shape the tools they need to produce new space designs, then their creations and aesthetics will always be limited by their technological competence. On the other end, technologists with a robust knowledge of people tracking and statistical modeling, blended with creativity and a sense for experience design and space design, seem to be in a better position than traditional architects to contribute to new trends in architecture.

To build narrative spaces, rather than stressing the importance of collaboration among people with different backgrounds and fields of competence, the author, based on her own education and experience, wishes to show that today's architect can also be in equal measure a scientist, an engineer, and a visual artist and communicator. The role and required competencies of the contemporary architect tend to create a new professional figure characterized by a mastery of disciplines today considered belonging to separate practices and teachings. Yet this new professional figure has old and profound historical roots. The European renaissance has given birth to two typologies of intellectuals: the scientist type, incarnated by Galileo, who first established rules for scientific experimentation and scientific method and the artist-engineer, incarnated by Leonardo, involved in a creative research equally informed by art and science. In modern times Moholy-Nagy, and more closely to MIT, Gyorgy Kepes represent models of the contemporary artist-engineer. While the Galileo-scientist type has been predominant in western culture since after the renaissance, the emergence of digital media favors the re-appearance of the artist engineer, equally versatile in artistic creation and engineering abilities.

The nature of the projects here presented stresses the importance of statistical mathematical modeling techniques and corresponding technologies, such as pattern recognition and machine learning, for the field of interactive space design. The author believes that the main concepts and tools of these disciplines should become part of the current language of the today's and tomorrow's architects because they are the basis for any reliable sensor interpretation and intelligence simulation by machines, both indispensable in the new architecture of the information society.

Acknowledgements

The prototypes presented in this paper were developed under the scientific direction of Alex Pentland, head of the perceptual computing group at the MIT Media Lab. The author is also grateful to Neil Gerschenfeld, head of the MIT Media Lab Physics and Media group, who more recently provided additional and complementary scientific guidance and support. Glorianna Davenport offered advice and financial support. Dean William J. Mitchell has offered helpful suggestions and encouragement. Several MIT undergraduates contributed through the years to the research presented in this paper. Their names are: Chloe Chao, Mike Hlavac, Jeff Bender, Teresa Hernandez, Manuel Martinez, Tracie Lee, Chin Yan Wong, Eric Hilton, Sarah Mendelowitz, Anjali D'Oza. Other collaborators include Tom Minka, Christopher Wren, Ali Azarbajeani, Nuria Oliver, Yuan Qi. All featured photographs are by the author.

References

[1] Starner, T., Mann, S., Rhodes, B., Levine J., Healey, J., Kirsch, D., Picard, R., and Pentland A., "Augmented Reality through Wearable Computing." *Presence*, Vol. 6, No. 4, pp. 386-398, August 1997.

[2] Rijken D., "Information in Space: Explorations in Media and Architecture". *Interactions*, May-June 1999.

[3] Wren C., Sparacino F., et al, "Perceptive Spaces for Performance and Entertainment: Untethered Interaction using Computer Vision and Audition", *Applied Artificial Intelligence* (AAI) Journal, June 1996.

[4] Wren, C., Basu, S., Sparacino, F., Pentland A. "Combining Audio and Video in Perceptive Spaces". *Proceedings of: Managing Interactions in Smart Environments* (MANSE 99), Trinity College Dublin, Ireland, December 13-14 1999.

[5] Wren C., Azarbayeani A., Darrell T., Pentland A., "Pfinder: Real-time tracking of the human body". *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):780-785, July 1997.

[6] Darrell T., Moghaddam B., Pentland A., "Active face tracking and pose estimation in an interactive room." In: *Proceedings of: CVPR 96*. IEEE Computer Society, 1996.

[7] Azarbayejani A., Pentland, A. "Real-time self-calibrating stereo person tracking using 3-D shape estimation from blob features". *In: Proceedings of 13th ICPR*, 1996.

[8] Starner, T., and Pentland, A. "Visual Recognition of American Sign Language Using Hidden Markov Models." International Workshop on Automatic Face and Gesture Recognition (IWAFGR). Zurich, Switzerland, 1995.

[9] Pavlovic, V.I. Dynamic Bayesian Networks for Information Fusion with Applications to Human-Computer Interfaces. PhD Thesis. University of Illinois at Urbana-Champain, 1999.

[10] Hall D.L. and Llinas J. "An Introduction to Multisensor Data Fusion". In: *Proceedings of the IEEE, special issue on Data Fusion*, January 1997.

[11] Sparacino F., Davenport G., Pentland A., "Media Actors: Characters in Search of an Author." *IEEE Multimedia Systems '99, International Conference on Multimedia Computing and Systems* (IEEE ICMCS'99), Centro Affari, Firenze, Italy 7-11 June 1999.

[12] Sparacino, F. "(Some) computer vision based interfaces for interactive art and entertainment installations". In: INTER_FACE Body Boundaries, (issue editor Emanuele Quinz), *Anomalie*, n.2, Paris, France, anomos, 2001.

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

[14] Sparacino F., Larson K., MacNeil R., Davenport G., Pentland A. "Technologies and methods for interactive exhibit design: from wireless object and body tracking to wearable computers." In: *Proceedings of the International Conference on Hypertext and Interactive Museums*, ICHIM 99, Washington, DC, Sept. 22-26, 1999.

[15] Sparacino F., Pentland A., Davenport G., Hlavac M., Obelnicki, M. "City of News". In: *Proceedings of the Ars Electronica Festival*, Linz, Austria, 8-13 September 1997

[16] Sparacino, F, Wren, C., Azarbayejani, A., Pentland A. "Browsing 3-D spaces with 3-D vision: body-driven navigation through the Internet city". *Proceedings of 3DPVT: 1st International Symposium on 3D Data Processing Visualization and Transmission*, Padova, Italy, June 19-21 2002.

[17] Sparacino, F., Wren, C., Davenport, G., Pentland A. "Digital Circus: a computer-vision based interactive Virtual Studio". In: *Proceedings of IMAGINA*, Monte Carlo, Monaco, 18-20th January 1999.

[18] Sparacino, F., Oliver, N., Pentland A. "Responsive Portraits". In: *Proceedings of The Eighth International Symposium on Electronic Art* (ISEA 97), Chicago, IL, USA, September 22-27, 1997.

[19] Sparacino, F. "The Museum Wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences". In: *Proceedings of Museums and the Web* (MW2002), April 17-20, Boston, 2002.

[20] Sparacino F., Davenport G., Pentland A., "Wearable Cinema/Wearable City: bridging physical and virtual spaces through wearable computing." *IMAGINA 2000*, Montecarlo, January 31st-Feb 3, 2000.

[21] Sparacino, F., Wren, C., Davenport, G., Pentland A. "Augmented Performance in Dance and Theater". In: *Proceedings of: International Dance and Technology 1999* (IDAT99), at Arizona State University, Feb. 25-28, 1999.

[22] Bruner J. *Acts of Meaning*, Cambridge, Ma.: Harvard University Press, 1990.

[23] Sparacino, F., Corsino N., Corsino N. "City-Dance: project proposal for a body driven outdoors installation in Shanghai". Unpublished.

[24] Lozano-Hemmer, R. "Relational Architecture". *Performance Research*, Routledge, London 1999.

[25] Sparacino, F., Pentland, A., Davenport, G. "Wearable Performance". In: *Proceedings of International Symposium on Wearable Computers*, Cambridge, Massachusetts, USA, October 13-14, 1997.