The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2005)
M. Mudge, N. Ryan, R. Scopigno (Editors)

Exploring Digitized Artworks by Pointing Posture Recognition

C. Malerczyk, P. Dähne and M. Schnaider

ZGDV e.V. Computer Graphics Center, Darmstadt, Germany

Abstract

This paper describes a Mixed Reality-supported interaction system to explore digitized artworks like two-dimensional paintings and three-dimensional sculptures. Using an easy and intuitive pointing gesture recognition system, the user is able to interact directly with the artworks, which leads to a deeper involvement with and understanding of the art pieces. The usage of a video-based gesture tracking system ensures seamless integration of Mixed Reality technologies into a traditional museum's environment. Furthermore, it addresses even technically unversed users since no additional physical devices are needed and even no training phase is necessary for the interaction.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Interaction techniques H.5.2 [Information Interfaces and Presentation]: Interaction styles

1. Introduction

Cultural organizations such as museums and galleries face the situation of competition for the favor of visitors or the public in general with numerous suppliers of different educational or entertaining offers. As a consequence, they are looking for novel ways of attractively present their cultural assets, helping them to increase demand for cultural related edutainment. Interactivity here is one of the key factors to attract visitors. Hence information technology in general and in recent years Mixed Reality technology in particular received attention as these promise a wide range of interaction potentials. But anyhow, just applying technology is not enough to reach full success. Bob Raiselis, an exhibit developer, summarized in his article 'What makes a good Interactive Exhibit' [Rai05] some key characteristics that should be considered when trying to build interactive exhibits, like

- Be inviting, i.e. the exhibit raises the interest of visitors and invite them to spend some time with it;
- Be understandable, i.e. the navigation scheme of the exhibit should be instantly understandable and the visitor should be able to get the exhibit to 'work';
- Be explorative, i.e. invite visitors to explore the exhibit and let them discover things by themselves;
- Be accessible to people of varying ages and development.

In this paper we propose a number of interactive scenarios matching the accustomed association of artworks in a museum such as the presentation of paintings on a canvas or the presentation of three-dimensional sculptures with innovative human-computer-interaction methods. The combination of intuitive interaction techniques and the presentation of multimedia content on a large screen resolution system is used to generate a novel experience during an exhibition visit. Interacting with virtual exhibits should instantly increase the level of interest of the user and thus the impact of quality of education through hands-on experiences [MS03].
We describe the idea of a Mixed Reality-supported interaction system for museums and large galleries and its prototypical installation. The visitor is able to create his/her own exhibit and can choose between different digitized paintings and three-dimensional sculptures for exploration. Once selected, he/she can interact with the selected artwork in an easy and intuitive way just by pointing at the interaction canvas. The display of digitized paintings and sculptures on an interactive screen is usable for museums that are too limited in space to present all their art pieces in a traditional way. Furthermore, the direct interaction with art pieces typically leads to a deeper involvement with and understanding of the art pieces, whereas the manipulation of original paintings and sculptures is obviously prohibited. Exploration of

the paintings is e.g. achieved by giving the user the possibility of looking at details of the paintings, which he/she normally only can see using tools like a magnifying glass or by helping the to obtain additional information about the painting, which normally have to be read in the museums guide.
Due to the fact that all visitors of the museum should be able to use the system, the input device has to be as easy and intuitive as possible. We address this issue using an innovative video-based hand pointing recognition system as the input device. The usage of a video-based tracking system addresses even technically unversed users since no cumbersome physical devices need to be used and no training phase is necessary to perform interaction.
Rendering software and image processing algorithms are used for data management, which enables displaying and manipulating high-resolution 2D images of digitized paintings and 3D scanned sculptures in real time.

2. Technichal Setup

The equipment for the interactive museum exhibit consists of one single standard PC, which is used for both rendering of the scenario application and pointing gesture recognition and tracking. A standard video beamer or a large plasma display is connected to the PC, displaying high-resolution images, virtual additions to the image or control objects like sliding menu frames on the canvas in front of the exhibit visitor. For tracking purposes two Firewire (IEEE1394) cameras are connected to the computer feeding the system with greyscaled images of the interaction volume in real time. It is possible to use special lenses with infrared diodes to flood the scene with additional infrared light to ensure safer light conditions and to enhance the robustness of the video-based tracking system.

Figure 1: *Camera and (optional) infrared light beamer.*

Equipped with 4mm lenses the cameras are mounted at the ceiling at approximately 2.5 meters from ground with a distance of 2-3 meters from left to right.
Since the human pointing posture is naturally not as precise as a technical device like e.g. a laser pointer, it is important to permanently provide the user of the system with a visual feedback for comprehensible perception of his/her interaction instead of calculating the pointing direction as precise as possible.
The position of the user is somewhat pre-defined with respect to the camera set-up. Inducing the user to take the correct position can easily be achieved by adding markers like footsteps or other indicators on the floor in front of the interaction canvas. The actual position of the cameras with respect to the user and the displaying canvas however depends on various parameters like

- Focal length of the camera lenses;
- Dimension of the rendering canvas; and
- Designated speed and accuracy of the tracking system.

The above-mentioned distances of the cameras are proposed values for a canvas of approximately 2.0m width and 1.5m height and a user position located at about three meters in front of the canvas.
The project hardware used for the demonstrator consists of off-the-shelf components. The system is running on a standard PC. The current configuration features a Pentiumő 4 processor with 3.2 GHz, 1GB of memory and a NVIDIA GeForce 6800LE graphics board. The software runs under MS Windows XP.
As output device any standard video beamer with a resolution of 1024 * 768 pixel or higher and a projection canvas of approximately 2.0m width and 1.5m height can be used. The current set-up of the implemented prototype uses a back projecting system by OTLO VR Systeme GmbH, Rostock, Germany (http://www.otlo.de/) with a Liesegang ddv 1800 beamer. In addition, the system is tested with a large scale monitor (NEC 61" Plasma Display Monitor PX-61M1A, http://www.nec.com/) with a resolution of 1360 * 768 pixels. As no hardware interaction devices need to be connected to a computer, the complete technical set-up can be concealed from the user. The display canvas, the user is interacting with, is the only piece of technical equipment visible to the user.

3. Tracking Software

The purpose of the tracking module is to recognize and to track a static pointing gesture of the user to enable intuitive interaction with the scenario application. The definition of an abstract gesture description allows the tracking system to recognize individual pointing gestures without any learning procedure. The video-based module uses two cameras, which observe the user in front of the display canvas identifies if the user is pointing at the canvas and extracts the pointing direction. Due to a background communication with the rendering module the user gets a direct visual feedback on the canvas in real time.

The tracking module is separated into two different operation modes:

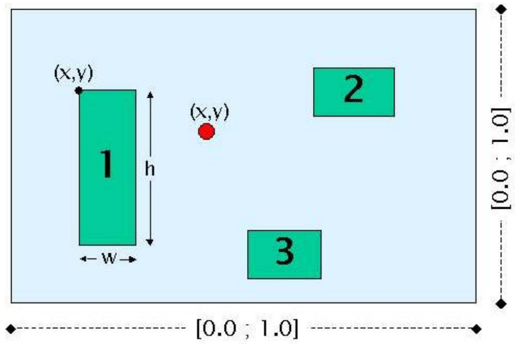

Figure 2: *Sketch of the virtual interaction area: pointing position (dot) and different button regions (boxes) defined.*

- Tracking of the pointing direction and its target point at the canvas and
- Observation of predefined regions that can be used like virtual buttons (pointing for a period of $\sim \frac{1}{2}$ second at a button object leads to a 'button selected'-event handled by the scenario application)

Both modes (pointing direction and region selection) are usable in parallel.

Hand gesture recognition in computer vision is an extensive area of research that encompasses anything from static pose estimation of the human hand to dynamic movements such as the recognition of sign languages [Koh05]. The demands on the tracking software used for this application arise from the scenario itself. In a public place such as a museum, a wide range of different visitors are expected to use the system. Therefore, it is necessary to have a tracking system at hand that is able to handle the interaction of different users, no matter if they are left- or right-handed, if they use just the index finger for pointing or even the opened hand. Furthermore, it is obvious that the tracking system needs to be usable without a visitor specific training phase. A museum visitor should instantly be able to interact with the exhibit without reading operating instructions first.

A combination of different basic computer-vision and image processing algorithms ensures a fast and robust identification of an eventually existing pointing gesture [SHB98]. The approach is based on the recognition of the human fingertip within a calibrated stereo system. Therefore, position and orientation of the cameras are determined with respect to the world coordinate system by swaying a small torch light for a few seconds in the designated interaction volume [SM02]. This calibration procedure has to be performed only once after setting up the cameras.
During runtime of the system, difference images are used to detect moving objects, which then are analyzed and the probability of a pointing posture and its direction in 3D space is calculated (see Figure 4). Intersecting the pointing ray with a virtual and normalized representation of the display canvas triggers the respective visual feedback or selection events. Smoothing of the tracking results using smoothing splines to reduce jittering effects [SE00] leads to an immersing experience during the interaction without the need of any technical device. The tracking system is nearly self-calibrating. Only a few parameters like the dimensions of the regions of interest in the images and a segmentation threshold have to be set or adapted during and after the installation of the system. Furthermore a simple graphical interface ensures the easiest handling of the tracking application.
Due to the separation of the tracking module and the scenario application, it is easy for the support staff of the museum to change or replace the content on the scenario side of the virtual exhibit without any need of changing parameters in the tracking software.

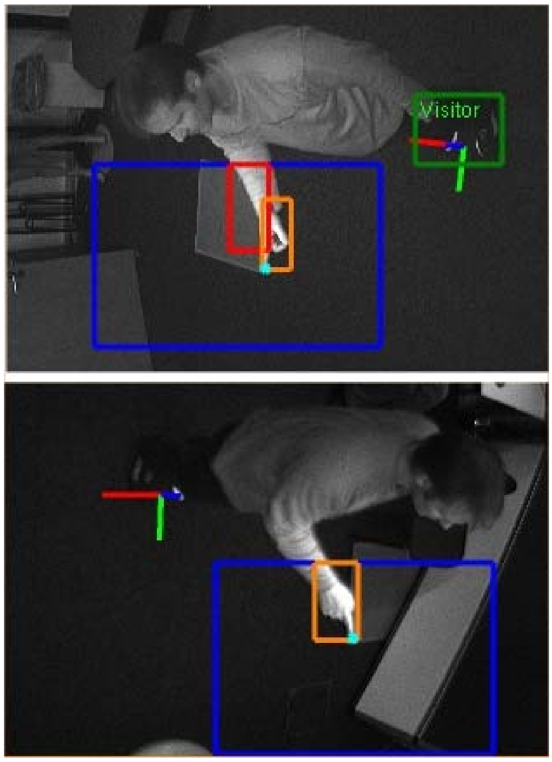

Figure 3: *Screenshot of pointing tracking software showing two (superimposed) camera images.*

4. Scenario Applications

The goal of the scenario application is to create an intuitively usable experience for any museum visitor, who is curious enough to explore digitized paintings on a technical exhibit canvas with a new interaction paradigm like the pointing recognition system.

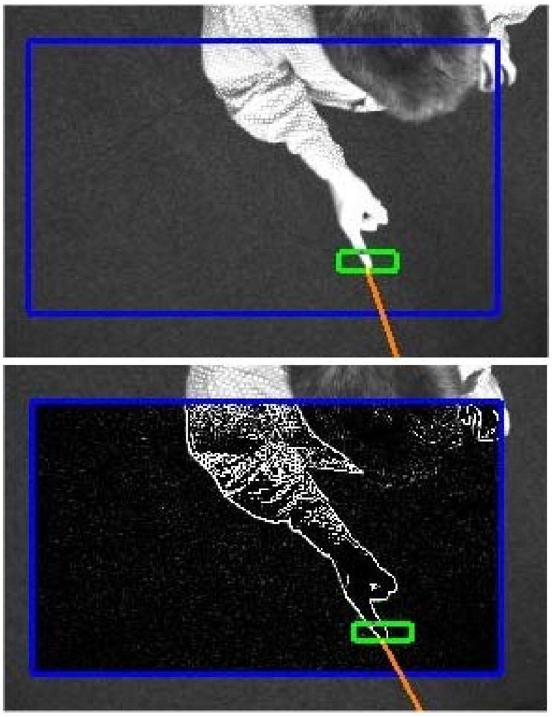

Figure 4: *Original camera image (top) and edge-of-difference image (bottom), both superimposed.*

For the creation of new content for the exhibition it is important to have standardized and easy to use authoring tools and rendering components at hand. We use Avalon [Ava05] [BF98] [BDR04] for the rendering part of the interactive museum exhibit. Avalon is an open environment for VR applications developed at the Computer Graphics Center (ZGDV) in Darmstadt, Germany. Avalon uses VRML/X3D as the programming language for the virtual worlds the user interacts with. Like most traditional toolkits, Avalon uses a scene-graph to organize the data, as well as spatial and logical relations. In addition to the scene description, VR applications like the interactive museum exhibit need to deal with dynamic behavior of objects, and the user interaction via non-standard input devices. The use of VRML/X3D as an application programming language leads to a number of advantages over a proprietary language [BDR04]:

- It is integral to an efficient development cycle to have access to simple yet powerful scene development tools. With VRML97/X3D, the application developer can use a wide range of systems for modeling, optimizing and converting.
- The interface is well defined by a company-independent ISO standard.
- Due to platform independence, development and testing can even be done on regular standard desktop computers.
- VRML and JavaScript are much easier to learn than the low-level interfaces often provided by traditional VR toolkits.
- There are a great number of books and tutorials available.

5. Paintings Exploration Scenarios

One of the most obvious applications for the presentation of and interaction with artworks in a museum is the exploration of two-dimensional digitized masterpieces. In museums it is often strictly forbidden to approach the original canvases. Therefore, a lot of visitors are in the dilemma that nonetheless they would like to look at details of the paintings. This problem could be solved by using the interaction system, where the visitor is able to explore a digitized copy of the masterpiece having e.g. a virtual magnifying glass or a virtual pocket lamp at his/her fingertip.

Furthermore, additional information about the original art pieces often has to be retrieved from books like printed museum guides or by using audio guides. We propose to use the interaction system to provide background information about the painting itself, the artist or even other paintings of that day directly by hands-on experience.

There is a large number of possible and useful applications how to interact with digitized paintings. The following sections describe three different applications we implemeted for demonstration and testing purposes.

5.1. Rousseau Scenario

As a first proof-of-concept of the interactive museum exhibit, we developed a simple application for the exploration of three different paintings of the French post-impressionist painter Henri Rousseau (1844-1910). The visitor of the museum exhibit is invited to take position in front of the canvas indicated by footstep markers on the floor indicating the designated interaction position. The application directly starts

Figure 5: *Image selection menu with permanent visual feedback.*

with the full screen exploration of one of the three images.

At the right hand side border of the image a virtual button is displayed indicating an image selection menu. Pointing for at least $\frac{1}{2}$ second at this button will activate the menu, which smoothly slides into the canvas and displays the selectable images (see Figure 5). During the interaction with an activated menu a small red point (understandable as a laser pointer metaphor) provides direct visual feedback during the selection phase. After the selection of a new image, the menu smoothly disappears and allows the exploration of the selected painting. At the left hand side border of the image another virtual button provides the exploration tools menu. Here, the visitor of the exhibition is currently able to switch between two different interaction modes, using the pointing posture as a virtual pocket lamp or as a virtual magnifying glass.

The spot light-based exploration, activated by selecting the according button of the tool selection menu, helps the visitor to focus on interesting parts and to blind out currently uninteresting parts of the image. If pointing does not take place the canvas is left black. Only the virtual button objects at the left and right hand edges of the canvas are visible. Pointing at the canvas leads to the effect of having a virtual pocket lamp at the fingertips of the user.

The magnifying glass-based exploration, activated by selecting the according button of the tool selection menu, allows the user to focus on interesting parts of the currently displayed image and to zoom in on details at which the user points. With the virtual magnifying glass at hand the user is able to let the lens slide over the image.

Figure 6: *Exploring Hieronymus Bosch's "The Haywain" triptych with a virtual magnifying glass.*

5.2. Hieronymus Bosch Scenario

The second scenario application exclusively addresses the exploration of a digitized painting using a virtual magnifying glass. For this scenario we have chosen the well known triptych "The Haywain" (see Figure 6) by the Netherlandish painter Hieronymus Bosch (c. 1450-1516), which is originally located at the Prado Museum in Madrid, Spain. Paintings of Hieronymus Bosch perfectly fit for an exploration using a virtual magnifying glass since Bosch is well known for his complex painted panels featuring fantastic and very detailed portrayals of demons, fools and other creatures from Eden to hell. The application directly starts with the full screen exploration of the painting. While no menu bars or other objects disturb the visual impression of the digitized painting, the visitor is able to focus solely on the painting and its details. Due to the panoramic aspect ratio of the original triptych we use a large 16:9 plasma display monitor with a diagonal screen size of 61" for the presentation. An additional post processing step in the pointing gesture tracking module allows an extremely stable position of the magnifying glass, if the user is bringing an interesting detail into focus.

Figure 7: *Guardi (1712-1793), The Marcus place with the clock tower, visible image (top) and x-ray image (bottom), Courtesy of the Picture Gallery of the Academy of Fine Arts Vienna.*

5.3. Guardi Scenario

Another scenario application allows the exploration of a single painting of Francesco Guardi with a virtual x-ray beam. The painting of the Italian Rococo Era painter shows the Marcus place in Venice, Italy (see Figure 7, top). The

original masterpiece is located at the picture gallery of the Academy of Fine Arts Vienna in Austria. X-ray photographs of the painting brought to light that Guardi used an already painted canvas [Sch03]. The x-ray photograph shows a painting of a manger scene of an unknown artist (see Figure 7, bottom), where several details of the original painting like the face of the Madonna, Jesus hold by his mother, two of the three Magi or even the head of the donkey are clearly visible.
Instead of showing the x-ray image on a separated canvas beside the original painting and using an information board or the museum guidebook for explanations, we use the pointing gesture tracking for an interactive exploration of both layers of the canvas. Without interaction, only the original painting of the Marcus place is visible. Pointing at the canvas enables a virtual x-ray view on the manger scene behind (see Figure 8). In addition to the visual feedback during the exploration, several regions of interest are predefined. If the visitor finds an interesting detail in the x-ray layer (indicated by pointing for at least $\frac{1}{2}$ second at the virtual region), a voice gives further information on this detail. The voice can either be a pre-recorded voice of a speaker, but it can also be synthesized during runtime of the system using a speech synthesis module generating speech from a given information text. The later allows easy creation or variation of the speech annotations incorporated in the scenario.

Figure 8: *Image exploration with virtual x-ray functionality.*

6. 3D Object Exploration Scenario

As a consequence of using an open X3D/VRML environment for immersive applications for the rendering purposes the next logical step is to build applications beyond the limitations of interacting with two-dimensional paintings and to enhance the system by the possibilities of exploring three-dimensional objects like statues, sculptures or other artefacts. The argumentation why to use digitized copies of the real artwork for exploration is the same as mentioned for two-dimensional paintings: The original 3D masterpieces are not allowed to be approached or even to be touched by the visitors.
Unfortunately, the presentation of three-dimensional content is not as easy as the presentation of digitzed paintings. That applies to the creation of suitable content as well as to the interaction with the virtual world itself. The major problem of the creation of 3D content is the fact that the original masterpiece has to be scanned to achieve a high resolution model of the object, which needs a much higher technical effort to spend than to take just one digital picture of a painting.
Furthermore, the interaction with a 3D virtual world is not as intuitive as just pointing at a specific position on the screen. While the museum visitor is able to look upon a real 3d sculpture from all sides, he/she has to rotate the virtual copy of the object in 3D space. There are two major possibilties how to interpret the pointing posture as an input device for the rotation of a three-dimensional object:

- The position on the displaying canvas the user is pointing at indicates directly the direction of the rotation of the object. In practice, it is neccessary to define an empty area in the center of the screen where no interaction takes place to avoid permanent rotation, whenever the user is pointing at the canvas.
- The virtual world is enhanced by additional virtual buttons for the rotation of the 3D-object. Whenever the user is pointing at one of these button, the object rotates left, right, up or down with a predefined rotation speed.

As a first application dealing with the presentation of three-dimensional objects we have chosen the exploration of a bust of the Greek mythological creature Medusa. The user is able to look upon the bust from every angle by rotating it using four virtual buttons arranged at the right hand side and the bottom of the screen (see Figure 9). A small red cursor ensures permanent visual feedback during the interaction.

Figure 9: *Interactive rotation of the Medusa bust, Courtesy of the Picture Gallery of the Academy of Fine Arts Vienna.*

7. Conclusion

The paper describes an approach for the creation of interactive exhibits in the context of cultural heritage. It addresses the specific needs in the context of interactive cultural heritage applications and offers extended interaction and utilization of digitized art works. The interactive museum exhibit has been tested and evaluated at different public places e.g. showing the Guardi scenario described above at the picture gallery of the Academy of Fine Arts in Vienna, Austria, which allows drawing some first conclusions with respect to acceptance and handiness of Mixed-Reality technology in museums and other cultural heritage sites.
The approach offers a sufficient set of interactivity at a very generic level and hence meets the requirement to be as intuitive and understandable as possible. It enables exploration of various forms of digitized art works, such as resulting from the digital preservation actions currently fostered by the European Commission in the field of cultural heritage. However, it is yet to be discovered whether or not interacting with virtual exhibits will directly lead to an increase in the level of interest of the user and thus the impact on the quality of education through hands-on experiences. This is considered to be a next step, which will include the incorporation of the approaches into real exhibitions at cultural institutions.

8. Acknowledgements

Parts of the work presented here were accomplished with support of the European Commission through the art-e-fact project; contract number IST-2001-37924 [Art05], and the SIMILAR Network of Excellence; contract number IST-2002-507609 [Sim05].

References

[Art05] ARTEFACT: Project homepage of the art-e-fact project. Retrieved July 2005 from http://www.art-e-fact.org. 7

[Ava05] AVALON: An open x3d/vrml-environment for virtual and augmented reality applications. Project homepage, Retrieved July 2005 from http://www.zgdv.de/avalon. 4

[BDR04] BEHR J., DÄHNE P., ROTH M.: Utilizing x3d for immersive environments. In *Web3D 2004 Proceedings* (2004). 4

[BF98] BEHR J., FROEHLICH A.: Avalon, an open vrml vr/ar system for dynamic applications. 4

[Koh05] KOHLER M.: Vision based hand gesture recognition systems. University of Dortmund, Website, Retrieved July 2005 from http://ls7-www.cs.uni-dortmund.de/research/gesture/. 3

[MS03] MALERCZYK C., SCHNAIDER M.: Video based interaction for arts and cultural heritage applications. In *1st International Workshop on Information and Communication Technologies (ICTs), Arts and Cultural Heritage* (May 2003). 1

[Rai05] RAISELIS B.: What makes a good interactive exhibition. Retrieved February 2005 from http://www.montshire.net/stacks/exhibits/goodexhibits.html. 1

[Sch03] SCHREINER M. R.: X-rays in art and archaeology: History, present state and perspectives. Denver X-ray Conference, Denver, Colorado, USA. 6

[SE00] SUN S., EGERSTEDT M.: Control theoretic smoothing splines. *IEEE Transactions on automatic control 45*, 12 (2000). 3

[SHB98] SONKA M., HLAVAC V., BOYLE R.: *Image Processing, Analysis, and Machine Vision*. PWS Publishing, 1998. 3

[Sim05] SIMILAR: Project homepage of the network of excellence. Retrieved July 2005 from http://www.similar.cc. 7

[SM02] SCHWALD B., MALERCZYK C.: Controlling virtual worlds using interaction spheres. In *Proceedings of 5th Symposium on Virtual Reality (SVR) 2002* (2002), C.A. Vidal B. C. S., (Ed.), pp. 3–14. 3