View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by DSpace at University of West Bohemia

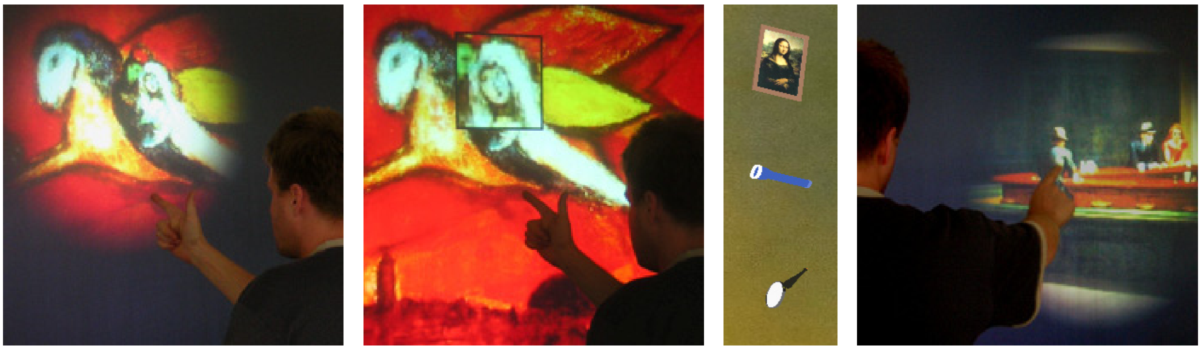
Interactive Museum Exhibit Using Pointing Gesture Recognition

Cornelius Malerczyk
ZGDV Computer Graphics Center
Fraunhoferstr. 5
64283 Darmstadt, Germany
cmalerc@zgdv.de

ABSTRACT

This paper describes a Mixed Reality-supported interactive museum exhibit. Using an easy and intuitive pointing gesture recognition system, the museum visitor is able to create his/her own exhibit choosing between different painters, artistic topics or just between different images. The usage of a video-based gesture tracking system ensures a seamless integration of Mixed Reality technologies into the environment of a traditional museum. Furthermore, it addresses even technically unversed users due to the fact that no physical devices need to be used and even no training phase is necessary for the interaction. The display of digitised paintings on an interactive screen is usable e.g. in museums that do not have the space to present all of their paintings in the traditional ways. Furthermore, direct interaction with art pieces leads to a deeper involvement with and understanding of the art pieces, whereas the manipulation of original paintings is obviously prohibited. Exploration of the paintings is achieved by giving the user the possibility of looking at details of paintings, which he/she normally can only see using tools like a magnifying glass.

Keywords
Gesture Recognition, Human-Computer-Interaction, Image Processing



1. INTRODUCTION

Museums have come to feel a necessity to look at new technologies to enhance the attractiveness of their exhibits. Due to the fact that Mixed Reality (MR) is commonly understood as a seamless integration of the real world and the virtual world, it provides the technical possibilities to create a deviceless interaction with digitised art pieces.

In this paper we propose an interactive exhibit scenario matching the accustomed association of artworks in a museum such as the presentation of paintings on a canvas with innovative human-computer-interaction methods. The combination of intuitive interaction techniques and the presentation of multimedia content like digitised paintings on a projection screen is used to generate a novel

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

WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.
Copyright UNION Agency – Science Press

experience during an exhibition visit. Interacting with virtual exhibits directly increases the level of interest of the user and thus the impact of quality of education through hands-on experiences.

This paper describes the idea of a Mixed Reality-supported interactive museum exhibit. The visitor is able to create his/her own exhibit and can choose between different paintings for exploration. Once selected, he/she can interact with the selected painting in an easy and intuitive way by just pointing at the interaction canvas. The display of digitised paintings on an interactive screen is usable for museums that are too limited in space to present all their paintings in a traditional way. Furthermore, the direct interaction with art pieces leads to a deeper involvement with and understanding of the art pieces, whereas the manipulation of original paintings is obviously prohibited. Exploration of the paintings is 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.

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. From the technological point of view, this leads to 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 due to the fact that no physical devices need to be used and even no training phase is necessary for the interaction.

Rendering software and image processing algorithms are used for data management, which enables the display and manipulation of high-resolution 2D images of digitised paintings in real time.

2. TECHNICAL SETUP

The equipment for the interactive museum exhibit consists of two standard PCs with approximately 2GHz processors and 1GB RAM each: one is used for the rendering of the scenario application and the other for pointing gesture recognition and tracking. A standard video beamer is connected to the rendering PC, displaying high-resolution images and virtual 3D objects on a canvas in front of the exhibit visitor. Furthermore, the PC used for tracking purposes is equipped with a frame grabber card. Two progressive scan cameras are connected to the frame grabber card feeding the system with grey-scaled images of the interaction volume in real time. It is possible to attach infrared filters to the cameras and to flood the scene with additional infrared light to ensure constant light conditions and to enhance the robustness of the video-based tracking system (see Figure 1). Assistance from infrared light beamers is only necessary, if the setup is running under extreme light conditions (e.g. if at least one of the cameras is looking directly into a window or if the environment is very dark to maximise the contrast of the display screen).

Figure 1: Camera and (optional) infrared light beamer

There are some constraints with the camera set-up: Both cameras have to be placed nearly orthogonal - one above the user at a distance of approximately three metres from the ground, the other at the side of the user at a distance of about 2m from user and 1.5m above the ground. The usage of this constrained camera set-up allows skipping a time consuming and complex camera calibration procedure to determine 3D camera positions and orientations in a defined world coordinate system. Due to the fact that the human pointing posture is naturally not as precise as a technical device like e.g. a laser pointer, it is more important to give the user of the system permanently a visual feedback and a smooth impression of his/her interaction instead of calculating the pointing direction as precise as possible.

The position of the user is pre-defined with respect to the camera set-up. Getting the user take the correct position can easily be achieved by adding markers like footsteps on the floor in front of the interaction canvas. The position of the cameras with respect to the user and the displaying canvas 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 3m width and 2m height and a user position located about three meters in front of the canvas. For the rendering device a standard video beamer with a resolution of up to 1024 * 768 pixels is used. Due to the fact that no technical interaction devices have to be connected to a computer, the complete technical set-up can be hidden from the user. The display canvas, the user is interacting with, is the only piece of technical equipment visible to the user.

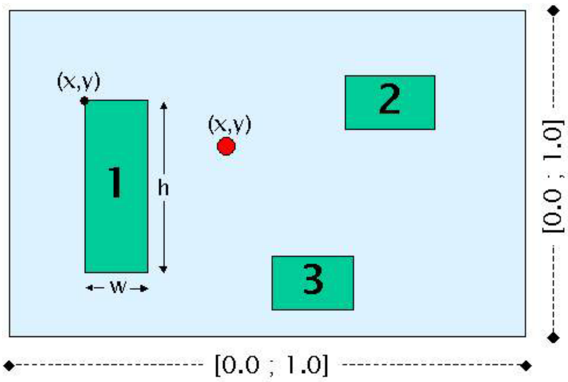


Figure 2: Sketch of the virtual interaction area (pointing position (dot) and three different button regions (boxes) defined)

3. TRACKING SOFTWARE

The purpose of the tracking module is to recognise and to track a static pointing gesture of the user to enable the most intuitive interaction with the scenario application. The definition of an abstract gesture description allows the tracking system to adapt to different and even technically unversed users without a 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 extract 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 (see Figure 2):

- 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 ~1 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 at the same time.

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 [Koh00]. 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 will 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 has to be usable without a visitor specific training phase. A museum visitor should directly be able to interact with the exhibit without reading operating instructions first.

A combination of basic different computer-vision and image processing algorithms is used to ensure a fast and robust identification of the eventually existing pointing gesture [Son98]. The approach is

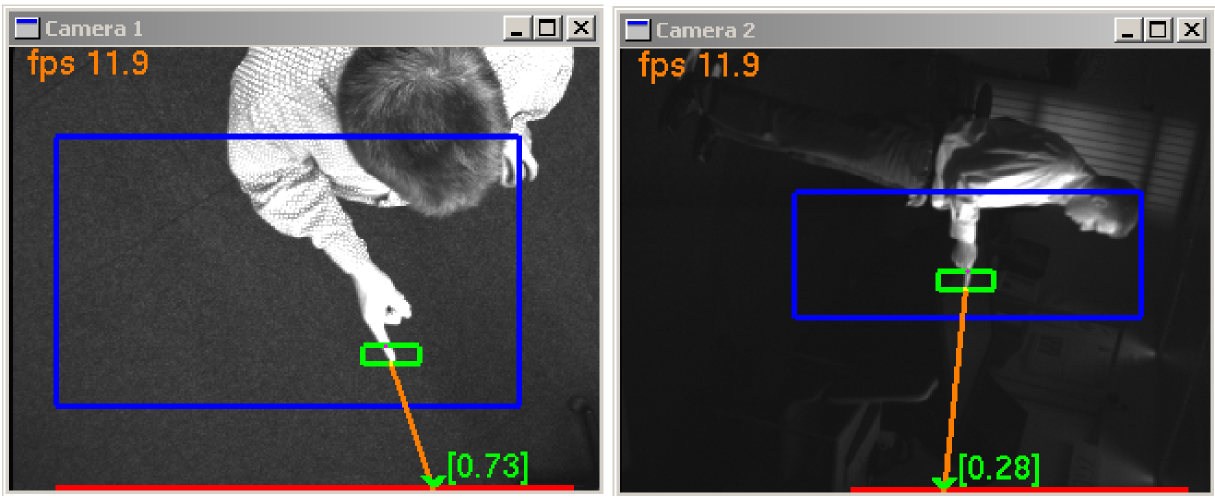


Figure 3: Screenshot of pointing tracking software showing two (superimposed) camera images (top view, side view)

based on the recognition of the human fingertip. Difference images are used to detect moving objects, which are then analysed and the probability of a pointing posture and its direction is calculated (see Figures 3 and 4). Intersecting the pointing ray with a

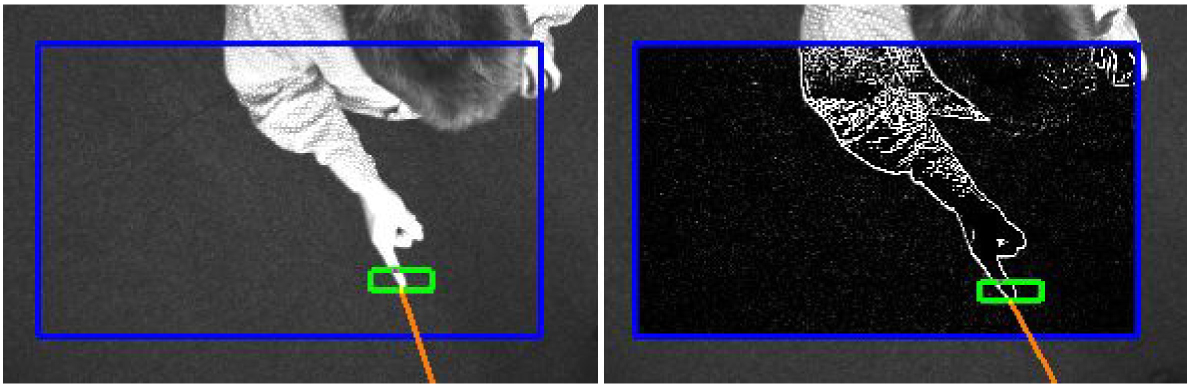


Figure 4: Original camera image (left) and edge-of-difference image (right), both superimposed

virtual and normalised representation of the display canvas (see Figure 2) triggers the respective visual feedback or selection events. Smoothing of the tracking results using *smoothing splines* to reduce jittering effects [Sun00] leads to an immersive 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.

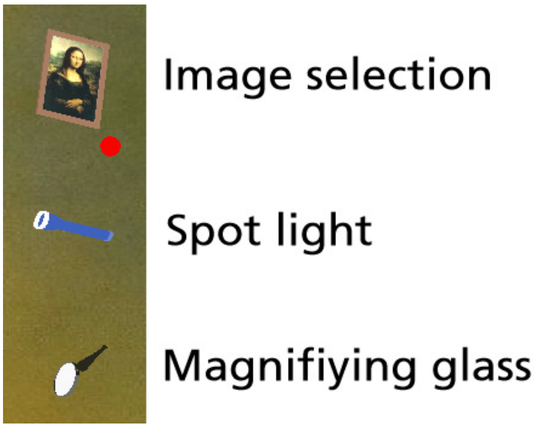


Figure 5: Button menu using rotating 3D objects (visual feedback point - the red dot - is superimposed)

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 any parameter in the tracking software.

4. SCENARIO APPLICATION

The goal of the scenario application is to create an intuitively usable experience for any museum visitor, who is curious enough to explore digitised paintings on a technical exhibit canvas with a new interaction paradigm like the pointing recognition system. The scenario application is running on a PC separated from the tracking module to ensure high performance for both applications and easy maintenance. Both computers are communicating via a network module to ensure real time interaction and visual feedback for the user.

The visitor of the virtual museum exhibition is invited to take position in front of the canvas indicated by footstep markers on the floor indicating the designated interaction position. In this scenario application, an image menu displaying nine pre-loaded images (see Figure 6) offers the selection of one of the paintings that is to be explored more detailed by the user. A small red point (understandable as a laser pointer metaphor) provides a visual feedback from the first time that the user points at the canvas (see Figure 5). Pointing at one of these thumbnail images for at least one second asks the application to replace the image menu with the selected image and to display it in full screen mode (see Figure 7). At the right hand side border of the image three rotating button objects are located, offering their different functions. These objects are standard VRML97 3D objects superimposed on the darkened border of the digitised painting. These buttons can be 'clicked' or 'pressed' by pointing at the objects in the same manner as explained above (see Figure 5):

- A rotating frame with the *Mona Lisa* of Leonardo da Vinci will bring the user back to the image selection menu;
- A rotating model of a pocket lamp indicates the spotlight exploration of the image; and

- A rotating magnifying glass enables the user to switch to an exploration, whereby details of the image are zoomed in on at the position at which the user points.

5. SPOT LIGHT SCENARIO

The spot light-based exploration, activated by selecting the according button of the tool menu, helps the visitor to focus on interesting parts and to shade currently uninteresting parts of the image. If pointing does not take place the canvas is left black. Only the rotating button objects at the right hand edge 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 (see Figure 8).

6. MAGNIFYING GLASS SCENARIO

The magnifying glass-based exploration, activated by selecting the according button of the tool 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 a virtual magnifying glass at the hand of the user, he is able to let the lens slide over the image (see Figure 9).

7. ACKNOWLEDGEMENT

The work presented here was partly funded by the European Commission within the artnouveau project; contract number IST-2001-37863 [Art03a] and the art-e-fact project; contract number IST-2001-37924 [Art03b].

8. REFERENCES

[Art03a] Project homepage of the artnouveau project, 2003, http://www.artnouveau-net.org/

[Art03b] Project homepage of the art-e-fact project, 2003, http://www.arte-e-fact.org/

[Koh00] Kohler, M., 'Vision Based Hand Gesture Recognition Systems', University of Dortmund, Website, 2000, http://ls7-www.cs.uni-dortmund.de/research/gesture/

[Mal03] Malerczyk, C., Schnaider, M., "Video Based Interaction for Arts and Cultural Heritage Applications", 1st International Workshop on Information and Communication Technologies (ICTs), Arts and Cultural Heritage, Donostia-San Sebastian, Spain, May 5th, 2003

[Son98] Sonka, M., Hlavac, V., Boyle, R., *Image Processing, Analysis, and Machine Vision*, 2nd ed., PWS Publishing, 1998

[Sun00] Sun, S., Egerstedt, M., 'Control Theoretic Smoothing Splines', IEEE Transactions on automatic control, Vol. 45, No. 12, 2000

Figure 6: Automatically generated image selection menu

Figure 7: Scenario with Marc Chagall painting and button menu on the right

Figure 8: Exploration with a virtual spot light

Figure 9: Exploration of image details with a virtual magnifying glass