Art – a Perfect Testbed for Computer Vision Related Research

Peter Peer and Borut Batagelj

Abstract Contemporary art nowadays tries to exploit modern technology to better address and enlighten specific problems and ideas of our time. Our interest in wider impact of modern technology on society and the interest in contemporary art, brought our attention also to the applicative field of use of computer vision methods in art. This chapter walks us through a few projects, proving that art is definitely a perfect testbed for our research: 15 Seconds of Fame, Dynamic Anamorphosis, Virtual Skiing, Smart Wall, Virtual Dance and Virtual Painter, from face detection, motion following, depth recovery, touchless human-computer interaction to pop-art, constant eye gaze of a person on the portrait, regardless of where the spectator stands, immersion into different virtual worlds without the need for any special equipment.

1 Introduction

Human ability to function in multiple disciplines, communities is again becoming very important. Latest observations in computer vision community are also showing the need for collaboration between our community, computer graphics community and contemporary art community. We are talking about somekind of convergence, eventhough the idea has been around for some time now.

A wider discussion is always appreciated as similar problems are enlightened from different perspectives. In such a way a specific professional terminology becomes community-independent (or at least less dependent), similar solutions are

Peter Peer
Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia, e-mail: peter.peer@fri.uni-lj.si

Borut Batagelj
Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia, e-mail: borut.batagelj@fri.uni-lj.si

refactored, joint, improved, the ideas become clearer, new applicative areas are found.

By following this philosophy, we initiated collaboration between our faculty and the Academy of Fine Art and Design at the University of Ljubljana. To be more precise, the collaboration between the students of both faculties within the undergraduate subject Communication Methods at our faculty was initiated by prof. Franc Solina (Faculty of Computer and Information Science) and prof. Srečo Dragan (Academy of Fine Art and Design). The successfull collaboration resulted also in the establishment of the formal association called ArtNetLab [4], which core purpose is to support fusion of science and art [21, 22].

This is way the next section talks about the way the students collaborate and points to some latest projects. Because of our interest in contemporary art and also because of seeing art as the perfect testbed for our research, a new and interesting applicative area for us, we also developed some installations out of the scope of the mentioned undergraduate subject. Thus, in section 3 we present a project 15 Seconds of Fame, an interactive art installation which elevates the face of a randomly selected gallery visitor for 15 seconds into a work of art. In section 4 we describe another project called Dynamic Anamorphosis, an installation that enables constant eye gaze of a person on the portrait, regardless of where the spectator stands. The Virtual Skiing idea is discussed in section 5. The Smart Wall project, which provides a platform for a rapid prototyping of computer supported interactive presentations that sense human motion is presented in section 6. Section 7 reveals the Virtual Dance idea, which allows to define a set of visual markers and to associate them with visual icons in the interactive video. The conclusions are given in section 8, touching the discussed projects, future projects, ideas and applicability of mentioned solutions in other areas than art.

2 How communities collaborate?

At the Academy of Fine Art and Design they have a Department for Video and New Media. Postgraduate students studying at this department develop an idea that incorporates new technologies into it. Since the art students are less experienced with technical development, have less engineering skills, we form a project group around them. The project team normally consists of one art student and two to four undergraduate students from the Faculty of Computer and Information Science in their final year of study. The advisor is also assigned to them to help them successfully carry out the project in the given time-frame, to monitor the progress, to suggest, point out the right approach.

Each year we form around 10 teams and they have a bit less than a semester to finish the project, present it and write a final report. The projects are then also presented, exhibited at different festivals.

The Smart Wall and Virtual Dance ideas (sections 6 and 7) are actually two projects done in such teams last year. Other projects include titles like: Sinking;

Morphing; Demon; An Invitation for a 20th Century Dinner; Interactive Meeting – just to name a few dealing with computer vision methods in art. Some interesting projects are wider in the sense that they are integrating new technologies and not specifically computer vision, for instance: Touch Animation; Protocols, Communications and Expanded Media (A Place of Home); DataDune.

Everybody involved in the process gain a lot: an artist gets the possibility to put his/hers idea into life, experiment freely with new technology and try to invent better and new ways of interfacing with virtual worlds [12, 17], science/engineering students get hands on practice, go through the project phases, learn to listen to and understand the customer, they all learn a lot, and at the end they are all enthusiastic because of the working installation, product.

All descriptions of the mentioned projects are available on the Internet [4].

3 15 Seconds of Fame

15 Seconds of Fame is an interactive installation which every 15 seconds generates a new pop–art portrait of a randomly selected person from the audience [23, 25]. The installation was inspired by Andy Warhol's ironical statement that "In the future everybody will be famous for 15 minutes". The installation detects human faces in digital images of people who are standing in front of the installation. Pop-art portraits are then generated from randomly chosen faces in the audience by applying randomly selected filters. These portraits are shown in 15 second intervals on the flat-panel computer monitor, which is framed as a painting. Electronic copies of each displayed portrait can be ordered by e–mail.

3.1 Motivation

Warhol took faces from mass media, banal in their newspaper everydayness, and transformed them into portraits. Warhol portrayed in this fashion celebrities from arts and politics. The installation tries to make instant celebrities by reversing Warhol's process – making Warhol-like celebrity portraits of common people and putting them on the gallery walls to make the portraitees implicitly famous. Since 15 minutes would hardly make the installation interactive the fame interval was shortened to 15 seconds. The faces for the portraits made by the installation are selected by chance out of all people in front of the installation to allude that fame tends to be not only short-lived but also random. In his film and video projects Andy Warhol was in fact fascinated with celebrification of "nobodies" which marks the beginning of an era in which media attention became the new mirror of the individual's self-perception.

3.2 How the installation works?

The visible part of the installation consists of a computer monitor framed like a painting. A digital camera is hidden behind the frame so that only a round opening for the lens is visible. Pictures of gallery visitors which are standing in front of the installation are taken every 15 seconds by the digital camera using a wideangle lens setting (Fig. 1). The camera is connected to a computer, which detects all faces in each picture, randomly selects a single face, makes a pop-art portrait out of it and displays it for 15 seconds on the monitor.

Fig. 1 People in front of the 15 Seconds of Fame installation

The color-based nature of our face detection makes it sensitive to illumination. Since it is not always possible to exhibit the installation under daylight or white-balanced studio illumination, we improved our face detection results by applying color compensation methods to make the whole system more flexible [25].

To make his celebrity portraits Warhol segmented the face from the background, delineated the contours, highlighted some facial features, started the process with the negative, overlaid the image with geometric color screens etc. We tried to achieve similar effects with a set of filters that achieve effects similar to segmentation. The filters drastically reduce the number of different colors by joining similar looking pixels into uniform regions. They combine three well known filters: posterize, color balance and hue-saturation with an additional process of random coloring. In this way, we achieve millions of different effects.

Our primary goal was not to mimic Andy Warhol's pop-art portraits per se, but to play upon the celebrification process and the discourse taking place in front of the installation. In comparison to other video camera based art installations, ours does not require exact positioning of observers due to automatic face detection with the additional benefit that a group of people can interact with the installation simulta-

neously. The interaction is technically very simple – no visible interface is actually involved – but unpredictable and socially revealing.

4 Dynamic Anamorphosis

Fig. 2 On the bottom of the painting appears a diagonal blur, which appears as a human skull when viewed from the upper right (The Ambassadors by Hans Holbein)

In [24] we define the concept of dynamic anamorphosis. A classical or static anamorphic image requires a specific, usually a highly oblique view direction, from which the observer can see the anamorphosis in its correct form (Fig. 2). Dynamic anamorphosis adapts itself to the changing position of the observer so that wherever the observer moves, she/he sees the same undeformed image.

4.1 Motivation

The dynamic changing of the anamorphic deformation, in concert with the movement of the observer, requires from the system to track the 3D position of the observer's head and the recomputation of the anamorphic deformation in real time. This is achieved using computer vision methods which consist of face detec-

tion/tracking of the selected observer and stereo reconstruction of its 3D position while the anamorphic deformation is modeled as a planar homography. Dynamic anamorphosis can be used in the context of art installation, in video conferencing to fix the problem of the missing eye contact and can enable an undistorted view in restricted situations. Anamorphosis serves as a model for the concept of the gaze, which suggests that visual appreciation rather than passive "looking" requires active "observing".

4.2 How the installation works?

We use a face detection method to determine the position of the user's face in the pictorial plane. Face detection is now a mature technology and can run in real-time. For head tracking we must detect the faces in every frame. To improve the tracking we use addition clues such as motion, skin color or near-infrared image. By using two or even more cameras and the principle of stereo reconstruction of distances we can further determine the position of the user's head in 3D space. The most difficult problem in stereo reconstruction is the correspondence problem – to find for a given point in the left image the corresponding point in the right image. Since the number of possible matches goes into thousands of points this is a computationally intensive task. The correspondence problem in this particular case is solved by finding faces in both images first. Next, only correspondences between faces need to be established.

We approach the stereo matching problem as a matching between homologous faces, instead of point matching. The main idea is to determinate a unique disparity value for the whole face region and no longer for individual pixels. After we detect the position of faces in both stereo images we construct a graph for each image where face blobs are represented as nodes in the graph. To find homologous faces in both stereo images we perform graph matching. The computational process is simple and fast since we consider only complete face regions.

At the end we deform the projected image of the face in such a way that it looks undeformed from the viewpoint of the observer (Fig. 3).

Dynamic anamorphosis disassociates the geometric space in which the user moves from the visual cues she/he sees, since wherever the observer moves, she/he sees the same image. The installation promotes a human face (Fig. 3) with the eye gaze directed straight ahead to meet the eyes of the installation user. It requires a dark room with the video projection over an entire wall so that the only visible cues seen by the user are given by the projection. The light reflected back into the room from the projected image must sufficiently illuminate the scene that face detection can be performed. Since the installation can truly be experienced only by a single user, the entrance to the room with the installation should be controlled.

Fig. 3 Transformed frame of a video clip, when the user views it under 30° angle from the right (Big Brother from the film after George Orwell's novel 1984)

5 Virtual Skiing

An interactive installation Virtual Skiing [26] enables a visual immersion into the feelings of gliding on snow through a winter landscape. The computer rendered winter landscape is displayed over the entire wall in front of the skier (Fig. 4). As on real skis you can regulate the speed of descent by changing the posture of your body so that the air resistance is decreased or increased. By shifting the weight of your body to the right or left ski you can make turns down the slope between the snow capped trees. The interface to the virtual world is implemented by computer vision techniques, which capture the posture of the skier's body using a video camera placed in front of him and processed on a PC in real time to drive the projected animation of the virtual slope.

5.1 Motivation

Real-time interaction of people with virtual environments is a well established concept but finding the right interface to do it is still a challenging task. Wearing different kinds of sensors attached to the body of the participants is often cumbersome. Computer vision offers the exiting possibility to get rid of such sensors and to record the body movements of participants using a camera [20]. People, their appearance
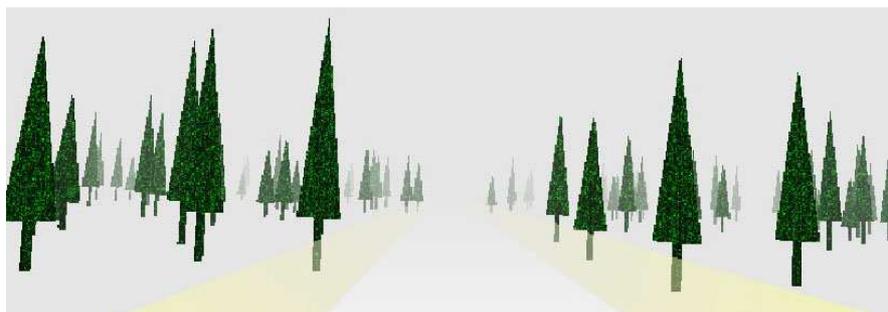
Fig. 4 The virtual slope is covered with sparsely populated trees among which the skier must find his way

(i.e. face), their emotions and the movements of their bodies are becoming on the other hand an important object of study in computer vision research [11].

The number of application areas for virtual environments is growing since the cost of technology for making virtual environments is in general going down. Sporting games in general are an attractive area for using virtual technology. Many training machines for cycling, running, rowing are enhanced with a virtual world to make the training more interesting. Instead of a static scene in a fitness room one can get the feeling of moving along a real scene or even to race against other real or virtual competitors. Virtual exercisers are sophisticated simulations that deliver the demands, stresses, and sensations of a sport or exercise with unprecedented verisimilitude and precision.

Skiing is a popular sport, which is unfortunately restricted to appropriate climactic and terrain conditions. Therefore various attempts have been made to bring this sport closer to anyone using modern technology. A very costly and direct method is to build an artificial slope with artificial snow or some other surface that enables sliding. Much more cost effective is now virtual technology although the whole ensemble of sensations experienced in the virtual world is not as realistic.

There have been quite a number of skiing games and skiing simulators played on a regular computer interface or on dedicated platforms such as the video game "Alpine Racer" by Namco. In the mid 1990's a special robotic platform was built by Jim Rodnunsky for the Vail ski center. The Ski and Snowboard Simulator is a surging, rolling, swaying, pitching, heaving, and yawing hydraulic recreation which took $4 million to develop. More recently, complete 3D models of existing ski centers have been build for promotional goals. They enable the user to freely move around and attain additional information or to follow a predefined path for the user to experience a virtual downhill race [1, 2].

5.2 How the installation works?

The virtual skiing installation consists of a video camera in front of the skier which records the skier's movements, a computer with a 3D graphics card, a sound system to convey the basic sound events (i.e. colliding with a tree) and a video projector to project the virtual slope.

From the grabbed video frame we first subtract the background. The background image is taken at the beginning and periodicity when there is no moving in front of the camera. In such a way we also get rid of the problems with changing illumination during the exhibition. Additionally we binarize the images to final separate the figure of the skier from the background (Fig. 5). The threshold for binarization is determined dynamically using a histogram method each time a new participant steps into the scene. The image subarea where binarization takes place can be adjusted interactively to suit different setups of the installation.

Fig. 5 The silhouette of the skier as result of the background subtraction and binarization of the input video image

Each time a user enters the scene as seen from the camera, the height of his silhouette is recorded and the center of his silhouette is determined by finding a pixel balance between his left and right side. As the skier shifts his body to the left or right side this initiates on modern carving skis a turn in the same direction. Turning can be therefore controlled just by shifting the center of the body/silhouette. When the user flexes his knees and lowers his body to achieve a lower air resistance the height of his silhouette is decreased and this is interpreted as an increase of the skier's speed (Fig. 6). The information of the body's position relative to the upright position is given to the skiing polygon rendering engine, which displays the virtual

slope. In reality, the biomechanics of skiing is much more complicated but for the interaction with a virtual environment such actions are quite realistic.

When the user comes in the virtual world too close to a tree, a collision is triggered. The collision algorithm is quite simple and written to suit the dedicated terrain rendering engine. It first checks if any of the trees are close along the direction of the movement of the skier (z axis). Then it checks if any of these trees are also in the x axis range (x axis of the projected image). When the skier collides, three parameters are taken into consideration. The skiers speed, tree height and the side from which the skier hit the tree (left, right, or direct middle). The life points are decreased according to the skier's speed and tree height. The side of the tree is used to "bounce" the skier off the tree to the side that the skier came from. Hitting the tree on its left side, would bounce him back to the left.

The software is written in C++ and uses the DirectShow interface to capture video data. For displaying the terrain the cross-platform SDL library and a custom rendering engine based on OpenGL is used. The rendering of the virtual terrain is done in OpenGL using the cross-platform SDL library. The trees have been modeled in 3D Studio and imported into the rendering engine. At startup this engine computes random positions and heights for 400 trees, which make up all of the trees seen. Additionally, 7% of the trees are flagged as "hearts". This information is used when the user is skiing in the survival mode of the game play to draw hearts. Another mode is plain skiing where the user skis the virtual slopes with no obvious objective other than to enjoy the experience.

6 Smart Wall

The main goal of the Smart Wall project is to provide a platform for a rapid prototyping of computer supported interactive presentations that sense human motion. The system is composed by a front end application, where the developer defines a
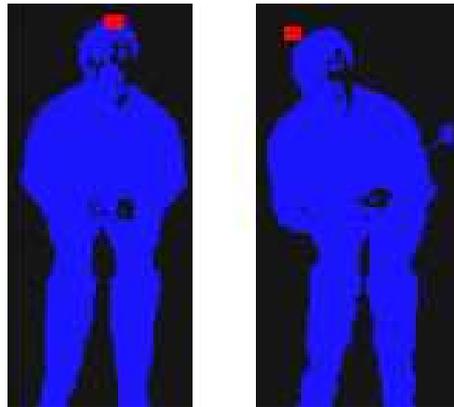

Fig. 6 The skier performs the same movements as on real skis to turn on the virtual ski slope. The position of the video camera can be seen at the bottom of the projection

number of hot spots in a camera view, a Hotspot Processor, which senses the activity in each of the hot spots, and a Player, which displays interactive content triggered by the activity in hot spots. By associating actions or sequences of actions in the environment to actions in the interactive presentation, a variety of complex interactive scenarios can be developed and programmed with ease. Due to the modular architecture, the platform supports distributed interaction, connecting physical activity and content display at remote locations.

6.1 Motivation

Computer vision is widely used to sense the presence and actions of humans in the environment. Novel surveillance systems can reliably track and classify human activity, detect unusual events and learn and retrieve a number of biometric features. Due to the low cost and the ubiquity of personal video technology, the research has recently shifted towards developing novel user interfaces that use vision as the primary input. In the area of personal computing, the most prominent areas of research are desktop interfaces that track gestures [18], and novel multi touch interfaces which use camera based tracking of fingers in contact with a translucent display [13]. On a wider scale, human motion can be used to interact with smart environments [20], or to trigger smart public displays [3].

The major motivation of Smart Wall framework is to provide an open environment to develop and experiment with interactive presentations in public space. The interactive area can be any public area that is observed by a camera connected to a computer. The field of view of the camera can then be divided in a grid of cells which define the smallest areas where activity will be detected. The resolution and, consequently, the size and the number of the cells, is defined by the developer. Multiple neighboring cells on the grid can then be joined in a single "hot spot", representing thus an interactive entity that promotes its state to other modules.

Each hot spot can define an action that occurs when the activity in the hot spot area exceeds a certain threshold (i.e. a person enters the area). The presentation on the screen is controlled by such actions or by a sequence of actions.

6.2 How the installation works?

Smart Wall system consist of three parts: HotSpot Processor, HotSpot Definition Interface and Player. All three parts communicate through XML files and can be installed on different systems. Furthermore, a Player with any other content can be implemented and by reading generated XML a variety of other actions could be achieved.

HotSpot Definition Interface (HDI) is a Flash application for defining hotspots (Fig. 7). HDI imports the picture of the floor and overlays the floor with grid of

selected resolution. In first step user can choose density of the grid and in second step user can define hotspots (a unique name can be applied for easier XML reading by users). In the last step HDI exports the XML via web service, so it can be saved on any computer.

HotSpot Processor is a C# application which uses AForge [15] library for computer vision. HotSpot processor input is a XML file from HDI in which hotspots are defined with the name, threshold and squares that are attached to it. HotSpot processor compares current image from the camera with the snapshot of the grid at the beginning or any other time if the user defined so. The output of the HotSpot Processor is a XML file in which every hotspot is presented with 0/1 value. If a change is detected on a hotspot then the hotspot value is 1 otherwise the value is 0.

The third part of the system is the Player. We used Adobe Flash in our project, as it can produce eye-catching, dynamic content and support many multimedia formats, however it can be implemented in any other technology. The player runs the presentation in loop and reads the HotSpot Processor output XML. When it detects a change on a hotspot it runs the action defined on the hotspot. HotSpot actions must be defined in the player.

Each part of the system can be used separately and for many applications (Fig. 8).

7 Virtual Dance

The Virtual Dance project [10] presents a system that is able to track dance movements and integrates the motion in an interactive video. The idea is to provide a flexible framework which allows the artist to set up an interactive virtual dance performance by defining markers, videos and interactive visual icons associated with markers. The system is then able to interact between the real movement and the vir-
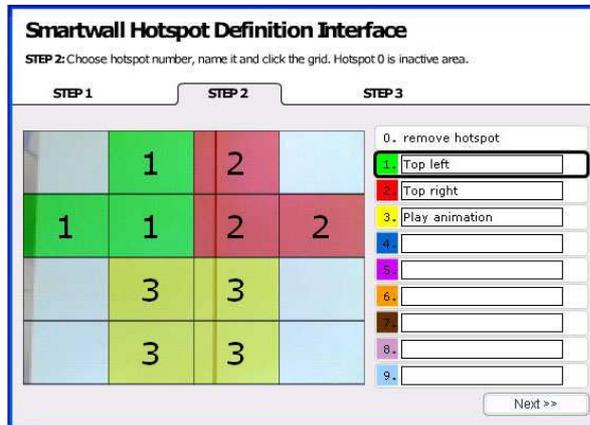

Fig. 7 HotSpot Definition Interface (HDI)

Fig. 8 Smart Wall in action

tual movement. We used standard tracking methods and modified them to support fast moving markers, small markers and discontinuous tracking of markers.

7.1 Motivation

Consequently, by dancing, the dancer creates two patterns of movement in two different spaces [19]. The first pattern is produced in the three dimensional real world. The second pattern is reflected in the two dimensional virtual space. These two representations are inseparable. The virtual dance is shown on the screen and in this way the real world dance extends to a new complementary dimension of expression [27].

7.2 The real and the virtual dance

The real dance is recorded by the camera in real time (Fig. 9). While the dance is limited by space, no spatial constraints should be imposed by the camera and other technology that processes the video. Therefore, the system must handle tracking of markers that leave the camera field of view and reappear, or markers that are temporarily occluded by the dancer, by objects in the scene, or by other dancers. In that way, the dancer can also leave the scene and reappear multiple times during the performance, and the system must be able to robustly track the markers thorough the performance.

The virtual dance is produced as a consequence of real dance. The user defines the virtual space while the dancers dance, so she manipulates the virtual dance and

Fig. 9 The dancer is marked with four hypothetical markers. The positions of the markers were chosen by the dancer

its presentation [7]. The application produces a movement of the virtual representations that is in correlation with markers on the dancer's body. The user interactively chooses the representation of a marker in the virtual video dance. The representation can be a picture or an animated two dimensional sequence, e.g. a movie clip of the dancer representing a pre-recorded movement. The user can also choose the background of the virtual dance and can change the size and the colour of virtual markers (Fig. 10).

The real dance and its virtual presentation are inseparably connected because of real time video processing [9]. Every movement in real world immediately produces a movement in virtual world. Dancers can observe the virtual dance that is produced by their movement [6]. So the dancers can also interact with the virtual space through their dance. A dancer can observe the virtual dance and produce a brand new story and movement in the virtual world. If she chooses some particular presentation of her markers in the virtual dance and if she moves in some special way, she can produce a new story in the virtual world which is not focused on tracking the real world movement, but it becomes a virtual story that has no connection with dance but is still directed by the dancer [5].
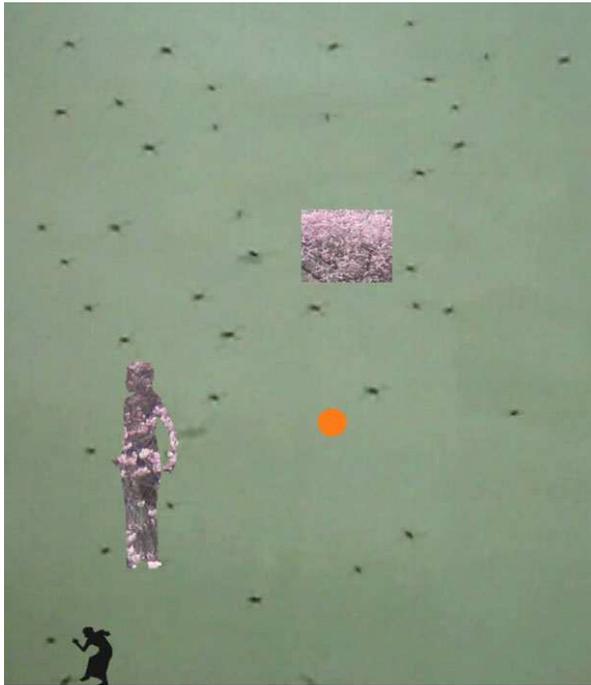
Fig. 10 This figure shows the virtual dance. Four markers were chosen in the real world, thus, the virtual world has also four markers

7.3 How the installation works?

The application can read the video either directly from the camera or from a prerecorded video file. The application then reads images from the input video, and, if markers are defined, each of the frames is processed by the marker tracker. In the pause mode, the video can be also displayed on the screen to support the definition of the marker areas (Fig. 11).

To represent a marker's appearance, we calculate the RGB histogram of chosen area on the picture where the marker is positioned. We apply a weighing schema; for example, if marker is mostly red, we give a high weight value to the red component, and a low weight value to the green and the blue component. The histogram is then a representation of the weighted distribution of red, green and blue values on the marker area (Fig. 12).

For every frame, we calculate the histogram back projection with respect to the histogram of the tracked marker. The function returns a probability of a pixel belonging to a marker area for each pixel in the image.

After we calculate the back projection, we calculate the marker position by using the Mean Shift algorithm [8]. We also consider the estimated position and size of the marker in the previous frame and modified the original Mean Shift algorithm to

Fig. 11 This figure shows an example of input video that is shown on the screen. The user has chosen that the marker will be the red textile and therefore positioned a selection rectangle on that textile
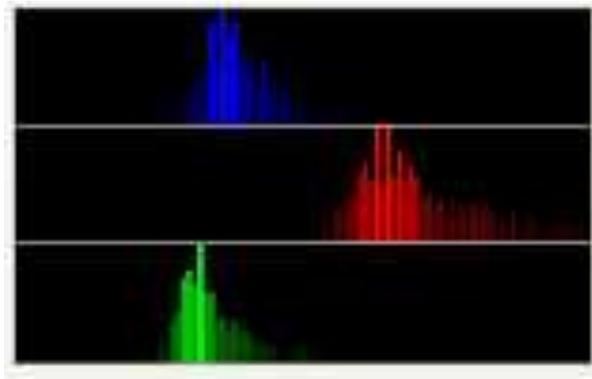

Fig. 12 Histogram of the marker area shown in 11

meet our demands. The original mean shift uses the back projection and the previous bounding box parameters of the marker. Then it finds the position inside the rectangle which has the highest probability of being part of the marker. This probability is calculated considering all probabilities of pixels that are inside the rectangle. Then, the center of the bounding box rectangle is moved over that pixel with the highest probability, and the same calculation is performed again. The calculation repeats until the rectangle is moved less than a user defined parameter at the last iteration, or for a maximum of N iterations.

The original Mean Shift does not track well fast moving objects, small objects, and objects that have left video and have appeared again. These are very frequent

events in dancing, where the camera is able to record only a part of the dance floor. So we made a change to the original Mean Shift algorithm. Since we cannot rely on a search only in the immediate neighbourhood, the search has to be initiated on a number of areas covering the whole image.

We divide the picture to 23 parts vertically and 23 parts horizontally. Then, the centre of rectangle is moved in every of $23\times23=529$ parts of the picture and the probability of marker being inside the moved rectangle is calculated. That way we find the position on the picture where is the highest probability that there is the marker if we move the original rectangle there.

The new position of rectangle represents the basis to calculate the new position using the Mean shift procedure. Once the application has found the position of the marker on the image, the position is written to a XML file.

The module that displays the virtual dance reads the position of markers from XML files and updates the position of virtual markers on the screen. The artist can select whether a marker is represented by a picture or by an animated sequence. She can also choose the size and the colour of markers. The background is also selected by the user and can display a static image or a video. New pictures and videos for markers and background can be simply added by updating XML files. The resulting virtual world and the virtual dance can be then integrated into the performance shown to the public.

8 Conclusions

The main idea behind this work is to prove or at least give a very good indicator that art is a perfect testbed for computer vision based research in many cases, giving you a very multidisciplinary feeling. This indicator is provided through the summarization of the projects in the Computer Vision Laboratory dealing with the convergence of topics mostly in computer vision, graphics and contemporary art communities.

Let us look back a bit. The installation 15 Second of Fame detects human faces in digital images of people who are standing in front of the installation. Pop-art portraits are then generated from randomly chosen faces in the audience by applying randomly selected filters.

The dynamic changing of the anamorphic deformation, in concert with the movement of the observer, requires from the system to track the 3D position of the observer's head and the recomputation of the anamorphic deformation in real time.

The installation Virtual Skiing provides a user with the opportunity to interact with a virtual space. The virtual skier can navigate down the virtual slope just by changing the posture of his body. The possible actions are turn right/left and change of the speed of descent. In the process of navigating down the slope the skier must avoid the sparsely arranged trees. The interface is very intuitive since the skier just repeats the actions that he knows from real skiing and learns to control his movement in the virtual world in less than a minute. The system works in real time and is very robust to any influences from the environment, such as change of illumination. Other

games for the same virtual skiing setup could be introduced. A race course with gates could be implemented so that the skiers could compete and try to finish the course in the shortest time. A virtual model of an existing slope could be used as a promotional means for that ski resort. Virtual skiing is just one possible application of such computer vision based interface.

We introduced the Smart Wall framework for controlling different presentations only with movements in the particular hotspots. The applicability of the Smart Wall idea is very diverse, for instance, we used it also in computer games [16].

With the Virtual Dance project we presented one of the possible connections between dance, video and technology. In ideal illumination conditions, the marker does not have to be a separate add-on that is worn by a dancer, but can be defined as a part of the costume or as a part of the body that is distinguished enough to be recognized by the system. The tracked person is not necessarily aware of the tracking. Obviously, we can use the system not only for tracking in dance, but also in many other applications.

Another example of immersion into a virtual world is given by the installation Virtual Painter, which is a work in progress. As suggested by the name itself, a user of it can paint without the actual paints, while the painting forms itself on the computer screen, it can be stored, printed. As in the previous cases the camera is observing the user in front of the installation and tracks hers/his palm. The movement of the palm is then transformed into streaks in the painting. With the second hand you could simply switch between the colors on the pallet by raising the hand. An example of the interaction with the installation is given in Fig. 13.

Fig. 13 Painting in the Virtual Painter installation

We could most probably go on and on with counting up application areas of introduced approaches, but for the end we would like to mention one more, which uses unconventional human-computer interaction through camera: helping disabled persons to use computers, helping them towards better life. The author of the initial version of the Virtual Painter was inspired by it and consequently initiated AIADA (Artificial Intelligence Aided Disable Assistance) project, which enables for instance writing in MS Word, drawing in MS Paint, manipulating components

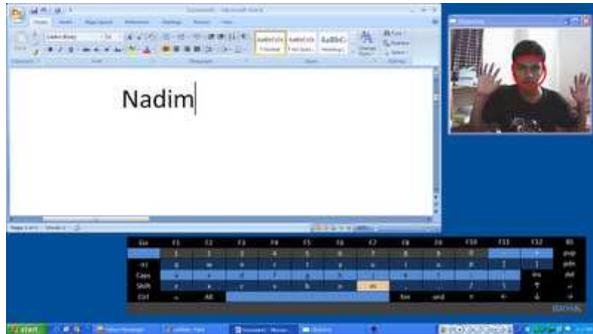of the operating system, talking, and controlling home appliances by tracking head movements (Fig. 14) [14].

Fig. 14 Writing in MS Word by tracking head movements

References

1. Almer A, Stelzl H (2002) Multimedia Visualisation of Geoinformation for Tourism Regions based on Remote Sensing Data. ISPRS – Technical Commission IV/6, ISPRS Congress Ottawa
2. Altland G (1997) Virtual Skiing Environment. http://www-personal.umich.edu/~galtland/skiVR/skiVR.html. Accessed 2 February 2009
3. Batagelj B, Ravnik R, Solina F (2008) Computer vision and digital signage. In: ICMI'08, Association for Computing Machinery, New York
4. Bovcon N, Vaupotič A (2000–2009) ArtNetLab. http://black.fri.uni-lj.si/. Accessed 2 February 2009
5. Burtnyk N, Wein M (1976) Interactive skeleton techniques for enhancing motion dynamics in key frame animation. Communications of the ACM 19:564–569
6. Calvert T (2007) Animating dance. In: Graphics Interface 2007, Canada
7. Calvert T, Wilke L, Ryman R, Fox I (2005) Applications of Computers to Dance. IEEE Computer Graphics and Applications 25:6–12
8. Comaniciu D, Meer P (2002) Mean Shift: A Robust Approach Toward Feature Space Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 24:603–619
9. Csuri C (1975) Computer animation. ACM SIGGRAPH Computer Graphics 9:92–101
10. Dovgan E, Čigon A, Šinkovec M, Klopčič U (2008) A system for interactive virtual dance performance. In: 50th International Symposium ELMAR-2008, Zadar, Croatia
11. Essa IA (1999) Computers seeing people. Artificial Intelligence Magazine 20:69–82
12. Grau O (2003) Virtual art, from illusion to immersion. MIT Press, Cambridge MA
13. Han JY (2005) Low–cost multi-touch sensing through frustrated total internal reflection. UIST 05: ACM symposium on User interface software and technology. doi:10.1145/1095034.1095054
14. Jahangir N (2008) AIADA (Artificial Intelligence Aided Disable Assistance) project. http://nadim0112358.blogspot.com/2008/07/project–aiada.html. Accessed 2 February 2009
15. Kirillov A (2009) AForge.NET framework. http://www.aforgenet.com/. Accessed 2 February 2009

16. Kreslin R, Dežman D, Emeršič Ž, Peer P (2009) Use of Unconventional User Interfaces Based on Computer Vision in Computer Games. In: potočnik B (ed) Rosus 2009, Maribor, Slovenia
17. Levin TY, Frohne U, Weibel P (2002) CTRL [SPACE], Rhetorics of Surveillance from Bentham to Big Brother. MIT Press, Karlsruhe: ZKM and Cambridge
18. Nielsen M, Moeslund TB, Storring M, Granum E (2008) Gesture Interfaces. HCI Beyond the GUI, Morgan Kaufmann, San Francisco
19. Norman IB, Stephen WS (1979) Digital Representations of Human Movement. ACM Computing Surveys (CSUR) 11:19–38
20. Pentland A (1996) Smart rooms. Scientific American 274:68–76
21. Solina F (2000) Internet based art installations. Informatica 24:459–466
22. Solina F (2004) ArtNetLab – the essential connection between art and science. In: Gržinić, M. (ed) The future of computer arts & the history of The International Festival of Computer Arts, Maribor
23. Solina F (2004) 15 seconds of fame. Leonardo (Oxf.) 37:105–110+125
24. Solina F, Batagelj B (2007) Dynamic anamorphosis. In: Enactive / 07 enaction in arts: International Conference on Enactive Interfaces 2007, Grenoble, France
25. Solina F, Peer P, Batagelj B, Juvan S, Kovač J (2003) Color-Based Face Detection in the "15 Seconds of Fame" Art Installation. In: International Conference on Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, image Analysis and Graphical special Effects MIRAGE'03, Paris, France
26. Solina F, Batagelj B, Glamočanin S (2008) Virtual skiing as an art installation. In: 50th International Symposium ELMAR-2008, Zadar, Croatia
27. Wilke L, Calvert T, Ryman R, Fox I (2005) From dance notation to human animation: The LabanDancer project: Motion Capture and Retrieval. Computer Animation and Virtual Worlds 16:201–211

Contents

Index