HUM, an interactive and collaborative art installation

ABSTRACT

This paper describes *HUM*, an interactive art installation which interprets the behavior of the visitors on different time scales to render visual and sonic artwork in real-time. *HUM* was presented at BRASS cultural center (Brussels, Belgium) in May 2009.

Keywords

Interactive installations, real-time sound and visual rendering, motion tracking analysis

1. INTRODUCTION

1.1 Related works: interactive installations

Interactive installations describe a wide class of artworks designed to transform the perception of a space by involving the audience acting on it. During the last century, following to the progress in technologies, the domain of architecture has been showcasing many inventive constructions, as highly documented and illustrated in Kronenburg's books [3]. Lots of immersive multimedia installations have been created by artists lately, as documented in [4, 1]. Most of these interactive installations relying on straightforward and commonly used computer vision techniques to analyze visitors' behavior, we believe there is still a high research potential for developing new ways of interaction between the artwork and its audience in such context.

1.2 Artistic intention

The following citation of the French painter Leopold Surnage (1879-1968) perfectly summarizes the approach we adopted for the creation of *HUM*: *'An abstract and static shape does not say a lot. Round or pointed, long or squared, simple or complex, it only produces an extremely confused sensation, as a simple graphical notation. This is only when it starts moving, transforming and meeting other shapes that it becomes able to evoke an emotion'.*

HUM is a generative sonic and visual artwork where interactivity is used to enhance the dialog between the artwork and its audience. Through their motion, the visitors handle visual and sonic shapes in a virtual scene and, in the same time, feed and educate *HUM*. As soon as a visitor stops moving within the installation, *HUM* gives the visitor a response through the same media by mixing the long-term trends learned since the beginning of its life, highlighting the potential richness of the installation. The aim, and unique criteria of quality, is to increase the energy freed by the visitor by encouraging him to go out of standard behavioral schemes. *HUM* is also a form of interactive and collaborative live painting: indeed, after its stay in the installation, each visitor leaves a visual track in the background of the scene, as a stamp of its life in the structure. The visitor is thus both creator and spectator, guiding and guided by *HUM*.

1.3 Technical overview

This section presents a brief technical overview of the installation[1], which will be then detailed in sections 2 and 3.*HUM* is a visual and sonic interactive installation relying on a motion tracking analysis; it interprets the behavior of the visitors on various time scales within the installation to control both visual and sonic synthesis in real-time. The navigation scheme in the installation is as follow: when a visitor enters in the space, he/she is attributed a visual shape which is added to the virtual scene rendered on the screen. This shape, whose color depends on the dominant color of clothes worn by the visitor, follows his/her displacements in the installation. According to his/her motion, the basis shape will also evolve in the scene and further graphical elements, such as points, lines, petals or replications of the original shape will be added on the screen according to the overall quantity of energy get out by the visitor.

The setup of *HUM* is described in Fig.1 ; the installation was presented in an 30 m^2 cubic structure, closed up to avoid light coming from outside, where visitors could move around and interact with the artwork. The setup consisted of a large 15 m^2 projection screen placed vertically against a wall of the structure, a retroprojector, 4 loudspeakers located at the edges of the structure, a video camera up to the screen and 3 computers running the video analysis and the visual and sound rendering.

[1] a video presenting the installation including authors comments is available online: http://www.tele.ucl.ac.be/~jjfil/ACMMM10.html

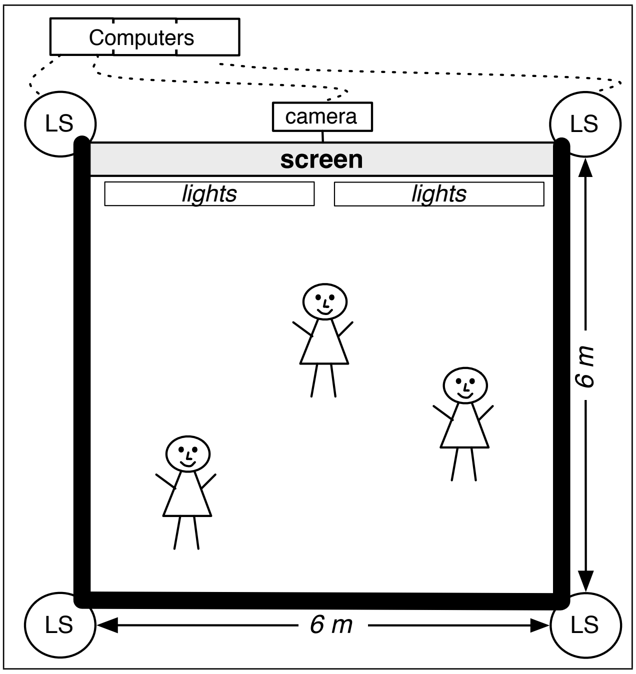

Figure 1: Setup of the installation.

2. ANALYSIS OF THE SCENE

In *HUM*, the interaction between the system and the audience is based on a two-levels analysis of the behavior of the visitors in the installation (Fig.2):

- a short-term analysis, mainly consisting of a motion tracking analysis, provides instantaneous features characterizing the position, the size and the quantity of energy get out by the visitor.
- a long-term analysis allows to consider the temporal evolution of the instantaneous features on various time spans, in order to understand how the visitor evolves during its stay in the installation.

The short-term analysis of the scene was implemented in Java whereas both the long-term analysis and the sound rendering used the real-time multimedia programming environment Max-MSP. The graphical synthesis relied on Processing, an open source programming language widely used by the interactive arts community built on the graphical capabilities of Java. Each of these software components communicated in real-time through the OpenSoundControl protocol. The following sections detail the two layers of the scene analysis implemented for *HUM*.

2.1 Short-term analysis

This first level of analysis provides a short-term description of the scene: it takes as input the video stream grabbed by the camera placed on top of the scene and provides a number of instantaneous features characterizing the current behavior of the visitors in the installation. The video stream is thus used to feed a motion tracking algorithm, which analyses the video frames by applying a background subtraction and outputs the location of 'moving objects', also named

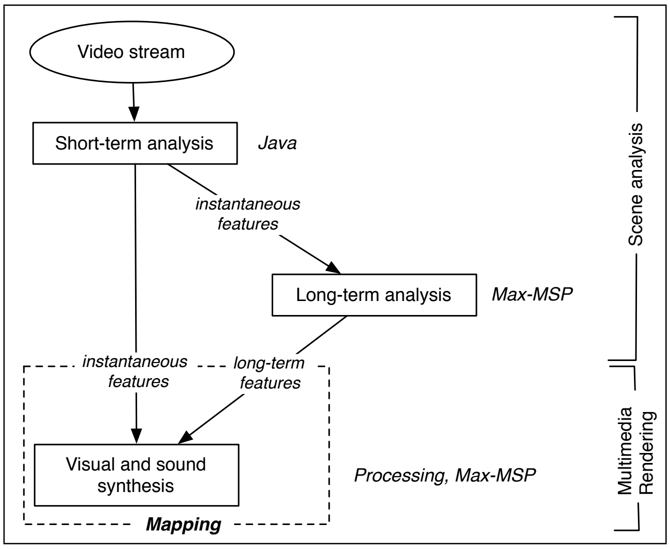

Figure 2: A software architecture based on a two-levels scene analysis

'*blobs*', within the video frame. These blobs, corresponding to the visitors in the installation are characterized by their instantaneous features, i.e. their size and position in the installation. A third kind of instantaneous feature is also provided by the motion analysis algorithm: it is described as 'quantity of energy' and refers to the overall amount of detected movement, including both displacement (from a point A to a point B) and movements of limited parts of the body, such as arms or legs, which do not necessarily imply a displacement of the visitor.

Motion tracking technology has interested the multimedia art community for decades and video-based motion tracking analysers have been integrated in most of the real-time multimedia programming environments such as Isadora, Eyesweb or Jitter. However, we decided to develop our own motion tracking module, inspired by the Eyesweb platform and written in the cross-platform Java language. This enabled to limit the number of software components involved in our installation and facilitate the communication between them. Fig.3 shows a snapshot of the graphical user interface of the motion tracking analyser we developed for this installation. It allows to easily set some parameters of the application, such as the sensitivity of the video segmentation, the minimum size of a blob and the maximum number of blobs detected in the scene.

2.2 Long-term analysis

A common approach adopted in interactive installations for the mapping between visitors' actions and visual and sound synthesis parameters relies on a direct one-to-one strategy; in such approach the position or displacement of a visitor can thus be used to trigger a sound event or modify a visual effect. Although this strategy provides efficient results, we believe that considering the long-term evolution of the visitors in the installation could allow to refine and enhance the interaction in such context. We also developed tools in Max-MSP enabling a long-term analysis of the behavior of the visitors in the installation. For that, our approach is

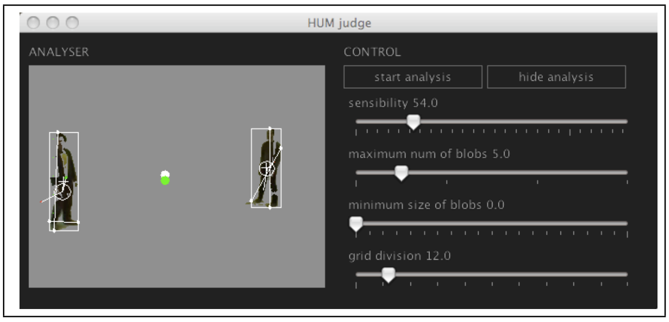

Figure 3: Snapshot of our video-based motion tracking analyser written in Java.

to characterize the long-term evolution of the instantaneous features provided by the short-term analysis, such as the position or the quantity of energy emitted by the visitor.

The first step of the long-term analysis also consists of buffering the incoming instantaneous features in a fixed-length temporal analysis window. This analysis window works as a shifting register: the new data takes the first position in the window, positions of previous data are shifted and the oldest value is dropped out of the window. The simplest way to describe the evolution of the features within the analysis window is to compute the statistical moments of the sequence, especially its mean and standard deviation. The histogram of the sequence may be also provide relevant information: it is a summary graph showing a count of the data points falling various ranges and provides a rough approximation of the frequency distribution of the data.

We also developed a Max object computing the histogram of a sequence of data (Fig.4). This object takes as input a stream of data, typically the position of a visitor in the installation, and computes its histogram. The output values of the histogram are normalized between [0-1] so that individual bins represent the fraction of the total number of positions assigned to the entire histogram. The histogram can be filtered in a way that all bins whose value is below a threshold are set to 0. Finally our object provides the n first minima and maxima of the histogram, and the bins associated to those values.

This tool has been used in *HUM* for analyzing the way a visitor occupied the space during its stay in the installation[2]. By computing the histogram of the position taken by the visitor within a temporal analysis window (typically few tens of seconds), we were able to provide a cartography of 'space occupation' corresponding to a visitor. The successive steps of this analysis are described in Fig.5: first the space, considered as a 2-D plane, is divided in a grid of $n*n$ cells, and coordinates *(x,y)* of a blob are converted in a z position in the grid defined as $z = x + n*y$ (step 1). Z positions are then stored in the analysis window (step 2), and a histogram of this sequence of positions is computed for each incoming video frame (step 3). This histogram is then used to construct a map of the areas which have been occupied by the visitor within the temporal analysis window. This map

[2]see a video demonstrating the space occupation analysis online: http://www.tele.ucl.ac.be/~jjfil/ACMMM10.html

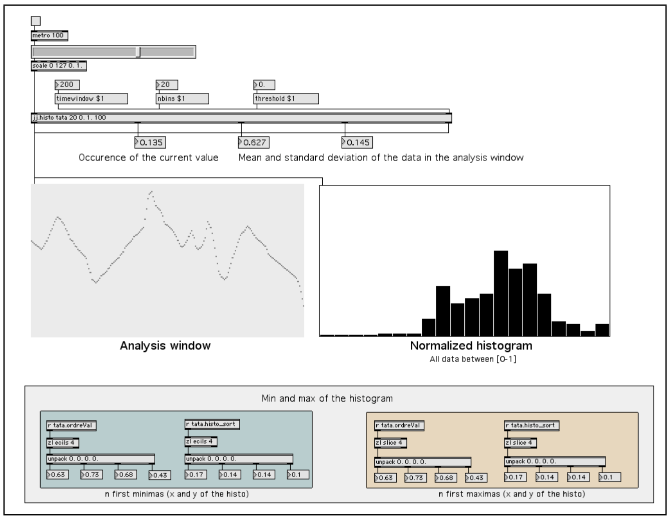

Figure 4: Example of an histogram computed in Max-MSP.

provides for each cell of the grid, corresponding to each bin of the histogram, the number of time the visitor has occupied this position within the analysis window (step 4). We used a display inspired by meteorological cartography to visualize the result of this analysis: the space is visualized as a 2-D plane, and a color is associated to each cell of the grid following the amount of time it has been occupied; a black cell corresponds to a cell which has never been occupied, a red one to a cell occupied for a long period.

This analysis allows to characterize the displacement of a visitor in the installation by providing various information:

- it indicates the visitor if stayed in a small area or visited a large part of the space,
- it provides the locations the more occupied by the visitor,
- it indicates if the current position of a visitor is a 'hot' or 'cold' position, i.e. if it has been or not occupied a lot within the analysis window.

These information complement the description of the scene based on the instantaneous features and have been used for the design of the interaction between the audience and the artwork in *HUM*, as explained in the next section.

3. VISUAL AND SOUND RENDERING

3.1 Visual rendering

The visual rendering in *HUM* is based on simple geometrical shapes: point, lines, circles, curves (Fig.6). To develop this choice in matter of color or movement, we referred to the Kandinsky's theory providing guidelines about the graphical weight of signs [2]. The usage of these basics shapes had two purposes: the first one was to avoid a direct representation of the visitor of the screen, the visitor having to make a small mind effort to understand by which shape he/she is represented. The second purpose was to ensure an

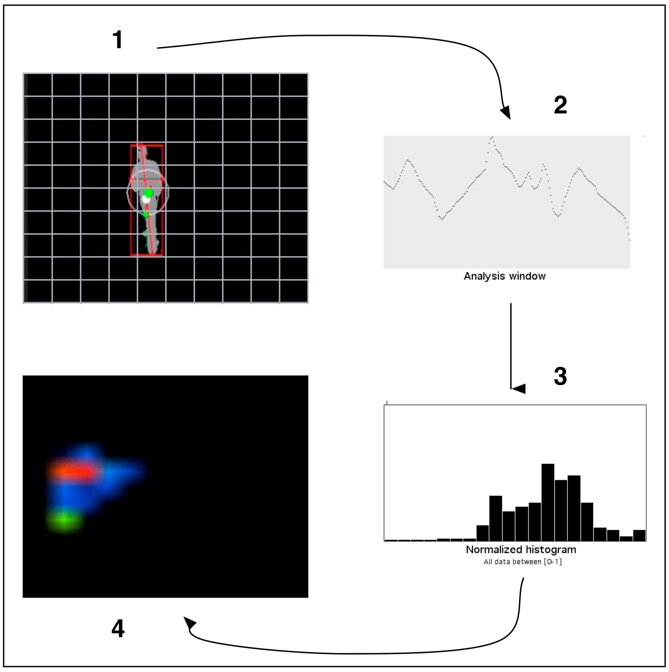

Figure 5: The long-term space occupation analysis.

easy access to any kind of people, independently from their age, ethnicity or education level; children for instance could interact with the installation even if they not fully understand the process. The features extracted by the analysis were used to trigger the apparition and evolution of specific graphical elements around the shape attributed to a visitor. Large displacements in the room made a tree of lines grow and swing whereas fast movement without displacement appended small rotating dots etc... In this way the visitors built together on the screen a blend of visual shapes according to their recent movements. Moreover the evolution of these shapes was periodically saved in a background image, which was the graphical result of all the visitors' interventions.

3.2 Sound rendering

The purpose of the soundtrack in *HUM* was to create a sonic ambience which could enhance the feeling of immersion of the visitor in the installation, while making sense with the visual rendering. This soundtrack was composed of four superimposed layers, generated by four digital sound synthesis instruments in Max-MSP and controlled by data coming from both the short-term and long-term scene analysis described in the previous section. The first layer of the soundtrack consisted of recorded samples of environmental sounds (rain, fire), which were spatialized over the four loudspeakers around the installation according the long-term space occupation of the visitors. The second sound layer was a pulse whose tempo increased with the number of visitors in the installation, evoking the heartbeat of the installation. The third layer of the soundtrack was based on a physical model of percussive instrument allowing to generate grained sonic textures. The density of the resulting texture was controlled by the amount of energy emitted by the visitors within a few seconds duration analysis window. The last layer of the

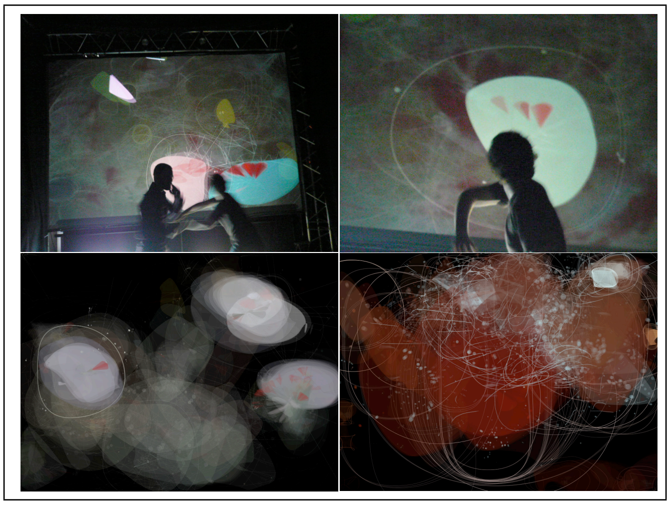

Figure 6: Snapshots of the installation. On top visitors interacting within *HUM*, on bottom background images resulting of a one day visitors' interventions

sound rendering relied on subtractive synthesis whose filters envelope were interpolated according to the position of the visitors in the installation.

4. CONCLUSION

This paper described a recent interactive art installation relying on motion tracking analysis. We presented both the artistic intention of its creators and the technical contributions related to this work. In particular, we developed tools enabling a long-term analysis of the behavior of the visitors in the installation, which can be used to enhance the interaction between the artwork and its audience. In the future we plan to go further in this direction by investigating more complex mapping strategies relying on this analysis approach in the context of interactive installations.

5. ACKNOWLEDGMENTS

The work described in this paper results of a collaboration between the visual artist Francois Zajéga and Numediart[3], a long-term research program centered on Digital Media Arts and funded by the Région Wallonne, Belgium. We also warmly greet the BRASS cultural center and its director Roger Burton for their support in the creation and development of *HUM*.

6. REFERENCES

[1] O. Grau. Virtual art: from illusion to immersion. *Leonardo, The MIT Press*, 2003.

[2] W. Kandinsky. *Point and Line to Plane*. Dober Publications, 1979.

[3] R. Kronenburg. *Flexible: Architecture that responds to change*. Laurence King Publishers, 2007.

[4] S. Wilson. Information arts: intersections of arts, science and technology. *Leonardo, The MIT Press*, 2002.

[3]http://www.numediart.org