Proceedings of the 8th International Congress on Archaeology, Computer Graphics, Cultural Heritage and Innovation 'ARQUEOLÓGICA 2.0' in Valencia (Spain), Sept. 5 – 7, 2016

NATURAL GESTURE INTERACTION IN ARCHAEOLOGICAL VIRTUAL ENVIRONMENTS: WORK IN PROGRESS

INTERACCIÓN GESTUAL PARA ENTORNOS DE INMERSIÓN ARQUEOLÓGICOS: TRABAJO EN CURSO

Niccolò Albertini[a,*], Andrea Brogni[a], Baptiste Caramiaux[b], Marco Gillies[b], Riccardo Olivito[a], Emanuele Taccola[c]

[a] Scuola Normale Superiore, Pisa, Italy. niccolo.albertini@sns.it; andrea.brogni@sns.it; riccardo.olivito@sns.it

[b] Goldsmiths University. London, UK. b.caramiaux@gold.ac.uk; m.gillies@gold.ac.uk

[c] Università di Pisa, Pisa, Italy. emanuele.taccola@unipi.it

Abstract:

Archaeological data are heterogeneous (i.e., data-sheets and pictures, stratigraphic data, 3D models), and innovative virtual reconstractuions helps to visualize and study those data. In this short paper, we describe our work in progress in the design of an innovative way to interact with the complexity of a virtual reconstruction, using natural gestures and advanced machine learning, in close collaboration with archeaeologists.

Key words: cyber-archaeology, gesture recognition, virtual reality

Resumen:

Los datos arqueológicos son heterogéneos (por ejemplo, ficha técnica e imágenes, datos estratigráficos y modelos 3D), y las nuevas tecnologías pueden ser capaces de ayudar en la visualizacion y el estudio de dichos datos. En este documento se presenta nuestro trabajo en curso que describe el diseño de una forma innovadora de interactuar con la complejidad de una reconstrucción virtual, mediante gestos naturales y avanzadas técnicas de aprendizaje, en directa colaboración con los arqueólogos.

Palabras clave: cyber-arqueología, reconocimiento de gestos, realidad virtual

1. Introduction

Nowadays, we can see, thanks to the new technologies, an increase in types and amount of data in any research field. Having pictures or 3D models, writing long report or plot statistical tables became normal not only for engineers or architects, but also for humanists and archaeologists. Manipulating and visualizing those kind of data has always been a limiting factor in the research and the scientific history is full of new graphical conventions, coming out just for better representing new data in a new perspective.

Archaeology is a field that more than others may take advantages from this innovation, because a virtual reconstruction, for example, puts the scientists in front at his/her hypothesis, making it real the availability for verifying it, studying archeological materials and their use. The theory of embodied cognition stresses the importance of our interactions with the world in our cognitive understanding of it. The ability to rotate an artefact in our hand helps us to see it better from different perspectives, the ability to hold a tool helps us to understand how it can be used. We propose that one of the great benefits of virtual reality is the ability to enable more natural forms of interaction that will better support archaeologists in understanding their digital artefacts. Our work goes in this direction, with a virtual reality representation of archeological data that includes a gestural interface enabling researchers to interact with data immersively, using their hand movements without the intermediary of desktop interface devices. This work comes straight from the archaeological researchers and their curiosity toward virtual technologies. The idea is to transfer actions from the excavation to the virtual environment, placing the archaeologist in a working replica of the excavation area, where he/she can access and manipulate the data, edit and consult information, moving around the reconstruction and the actual area. A very important issue will be to design such interactions with these virtual reality representations in a way that he/she feels natural and familiar in his/her imaginative concept of acting on the excavation area.

* Corresponding Author: Niccolò Albertini, niccolo.albertini@sns.it

EDITORIAL UNIVERSITAT POLITÈCNICA DE VALÈNCIA

2. The Application

The aim of the application was to contribute to the development of a new approach to the consultation and analysis of stratigraphic archaeological data. The first experiment was in fact based on a leap-sensor integrated on the 3D glasses to be used within a CAVE-like system, which enabled the interaction with the models through an interface consisting of a small set of gestures (Olivito *et al.* 2015). Although activated by hand movements, the gestures looked quite artificial because the user had to continuously interact with a selection menu mimicking a desktop or touch screen interface. In addition to this, since the 3D glasses and the leap-sensor were not wireless, they strongly limited the freedom of movement within the CAVE.

New tracking technology and new interaction strategy: these two key points are driving our work to improve the quality and effectiveness of the original system. The validation of the functionality of the gestures represents the result of a continuous debate between archaeologists and system developers, whose main goal has been to satisfy two specific requests by the archaeologists: on one hand, the possibility to use a limited gestures set, which easily memorizable because it is natural and, on the other hand, to reproduce through these gestures the movements an archaeologist usually carries out, not only during the field activity, but also in the cognitive process activated during the excavation and the interpretation phases.

The case study is the agora of Segesta, in Sicily, Italy. In 2014, archaeologists started a project of photogrammetric documentation of the digging activity, creating six 3D image-based models of the stratigraphic sequence. The objective is to support the traditional tools used in the archaeological field, so as to enrich the interpretative process due to the use of 3D data and to a simulation of the excavation activity while operating within an immersive virtual environment, that allows for a full embodiment and an interaction with the digital models as natural as possible. A further benefit is represented by the possibility of re-creating, also from an emotional and perceptive side, the mental dynamics an archaeologist processes during the field activity or even after.

Virtual reality and 3D visualization improved the observing process, letting the researcher perceive the data in a more natural way. An immersive visualization adds new feedback in the perception of the data, adding to vision the proprioception of our own body and inducing natural reactions during the exploration of the virtual objects. This will help to recognize shapes and dimensions, distances and spatial relationships between elements of the scene, increasing our space perception at a level more similar to reality. In the case of archaeology, for example, seeing an artefact in a 3D reconstruction and comparing it with others, it is more intuitive and efficient than watching pictures or columns of numbers.

Gestures and hand manipulations are actions that we perform every day without thinking, and in many cases, they can also be the best solutions in specific interactions. Physical gestures and actions are a natural way to interact with our external environment. However, designing gesture interaction still involves important challenges such as defining a gesture vocabulary that is relevant to the application area, as well as designing an accurate gesture recognizer that allows for taking into account expressive components of the performed gestures. Gesture design has been investigated following two distinct approaches. The first was designer-centered, considering ergonomics and technical constraints, while the was user-centered. Long *et al.* (2000) asked users to rate similarity between shape-based gestures to define a vocabulary avoiding ambiguity. Also, there may not be a single gesture that is applicable to all people. Gestures can be highly culturally specific: for this reason, it is often better to allow easy personalisation of gestures by individuals, an approach we take in our work.

A recent promising approach, in gesture recognition field, has been proposed (Caramiaux *et al.* 2014), where input gesture variations are explicitly taken into account in the model and estimated continuously while the gesture is performed. This method is called *gesture variation follower* (GVF) and it is the one we have applied.

The whole design strategy used in this work is also inspired by User Centred Design and in particular Participatory Design (Muller and Kuhn 1993). User Centred Design (Norman 2002) is an approach to designing technology where the needs of users are the key driving force behind any design decisions. Participatory design is key to multidisciplinary endeavors like digital heritage because it ensures that stakeholders like archaeologists are able to determine the design of their tools, even if they are eventually implemented by computer scientists. They key enabling technology has been Interactive Machine Learning (Fiebrink *et al.* 2011), the use of statistical machine learning algorithms to allow users to design interactive systems by giving examples of interaction rather than by programming. This has the advantage that the creation of the system can be much quicker and that it can be done by end users who do not have programming skills.

3. Work in Progress

The aim was to make gesture design and personalization as easy as possible, with users only having to record a single example of each gesture. On th one side, this enables each user to design and perform gestures in a way that is the most natural and comfortable for them and avoids the problem of trying to build gesture vocabularies that generalize across different individuals, professional specialisations, cultures and physical abilities. On the other side, visual feedback allows users to support them in training and using gestures.

The main idea behind the design of the application was to give archaeologists the opportunity to think about the data and not the interface. In particular, the application should allow the researchers to access to the data in a natural way by manipulating them with simple recorded gestures.

The first step was defining tasks and gestures which areintegrated into the working pipeline. The archaeologists identified the tasks (e.g., comparison between layers of excavation, analysis of the finds). They also identified suitable gestures. The first task is selecting the layer of excavation via two gestures to move up and down; a layer of interest is chosen; particular finds are selected by touching them (this was not implemented via the gesture system). Once selected, objects need to be measured and then deselected. There is also a context menu that displays different data types depending on where you are. This needs to be opened and closed.

For the creation of the scene, we imported 3D models of layers stacked on each other and for each layer we imported 3D models of the most interesting findings, lined up with the excavation. The scene was integrated with a reconstruction of the entire agora, which can be useful for comparison with the current state of excavations.

The set of gestures under test consists of two sets of two coupled gestures, in which one gesture is the inverse of the other: opening and closing the menu and moving up and down levels. There are also two single gestures: initiating measurement of the selected object and deselecting an object.

These gestures exemplify a number of gesture design strategies. The open and close menu gestures are relatively arbitrary, they do not use a particular metaphor, but were designed based on convenience of movement and the constraints of the system. The measurement gesture used a fairly straightforward real world metaphor: stretching out the hands in a clear echo of the movement made when using a measuring tape. The deselection gesture was similar: a sharp downward movement of the hands similar to the movement that would be made to throw down an object that was being held or to shake off an object stuck to the hands. Both of these have clear metaphors from daily life and are generic in the sense that it easy to see them being applicable in many domains, not just archeology. The gestures to move between levels, however, are much more specific to archeologist because they use a metaphor directl from their practice. The gesture to unearth a layer is an upward movement of the hand as if lifting soil off from the layer would reveal the layer below (the archaeologists explicitly described the movement in this way). Conversely, movement to the layer above is a downward movement replacing the earth and covering the current layer (Fig. 1). This is particularly interesting as it contrasts directly with the gestures used by the engineers in the team when doing initial testing on the system. Unlike the archaeologists, they immediately and without thinking about it used an upward gesture to move up a layer. This example clearly shows that, while many gestures are sufficiently generic to be similar and usable across a wide range of domains, certain gestures are specific of archaeologists. They are gestures that have clear metaphors to the archaeologists' physical practice of digging and handling artifacts (in fact, in later discussion the archaeologists raised the possibility of gestures mimicking the use of toolst is therefore important, when designing interfaces for heritage experts (or experts in other domains for that matter) that the experts themselves are participants in the design and as far as possible could design the interface themselves. Otherwise, the result will be generic interfaces that miss the particular physical metaphors that arise from expert practice.

Figure 1: The gesture vocabulary. From left to right and top to bottom: 1) initiating measurement by stretching hands out 2) moving to a lower layer by raising the hand 3) moving to a higher layer by lowering the hand 4) Rotation of the right forearm from inside to outside for opening the contextual menu 5) closing the menu with an inverse movement 6) deselecting the current object with a sharp downwards movement of both hands.

4. Conclusions

The application design process forced a reflection on the interaction modalities that archaeologists could use in virtual environments. This has made possible not only to visualize 3D models, but also to re-design and use natural gestures, in order to interrogate and consult data associated with models.

The gesture recogniser gave promising results, both for the efficiency of the application and its future developments. Archaeologists have emphasized the importance of having the possibility to mentally and virtually reproduce the field activity, so to be able to re-think and/or re-formulate interpretations elaborated during the excavation. Besides the possibility to operate in a completely hand-free interaction, thanks to hand-tracking sensors, the system will allow the use of external devices (i.e. tablets) for a wider consultation of metadata linked to 3D models, which are at present hardly visualizable within the CAVE.

References

CARAMIAUX, B., MONTECCHIO, N., TANAKA, A. and BEVILACQUA, F., 2014. Adaptive gesture recognition with variation estimation for interactive systems. *ACM Transactions on interactive intelligent systems (TiiS)*, **4(4)**, 18. DOI: 10.1145/2643204

FIEBRINK, R., COOK, P.R. and TRUEMAN, D., 2011. Human model evaluation in interactive supervised learning, *Proceedings of the 2011 annual conference on human factors in computing systems*, CHI, 147, 2011. DOI: 10.1145/1978942.1978965

LONG JR, A.C., LANDAY, J.A., ROWE, L.A. and MICHIELS, J., 2000. Visual similarity of pen gestures, *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 360-367. DOI: 10.1145/332040.332458

MULLER, M.J. and KUHN, S., 1993. Participatory design, *Communications of the ACM,* **36(6)**, pp. 24-28. DOI: 10.1145/153571.255960

NORMAN, D.A. 2002. *The design of everyday things*, New York: Basic books, 2002.

OLIVITO, R., TACCOLA, E. and ALBERTINI, N., 2015. A hand-free solution for the interaction in an immersive virtual environment: the case of the agora of Segesta. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* **40-5/W4**, pp. 31-36. DOI:10.5194/isprsarchives-XL-5-W4-31-2015, 2015