Abstract — This paper explores the potential of using Microsoft’s Kinect to create a low-cost and portable system to virtually navigate, through a prototype 3D GIS, the digitally reconstructed ancient Maya city and UNESCO World Heritage Site of Copan in Honduras. The 3D GIS, named QueryArch3D, was developed as part of the MayaArch3D project (http://mayaarch3d.unm.edu), which explores the possibilities of integrating databases and 3D digital tools for research and teaching on ancient architectures and landscapes. The developed system, based on the Flexible Action and Articulated Skeleton Toolkit (FAAST), controls in a remote and touchless mode the movements in the 3D environment in order to create a sense of spatial awareness and embodiment. A user can thus use gestures to interact with information stored in the spatial database, calling up photos, videos, textual descriptions as he/she moves through the virtual space of the ancient Maya city.

Keywords- 3D Modeling, Virtual Reality, 3D GIS, Kinect, Archaeology

I. INTRODUCTION

The latest improvements in data capturing technologies, computing, and the distribution of digital contents offer a wide range of possibilities for innovative applications in many different fields. In particular, the Cultural Heritage field is taking great advantage of the latest developments in 3D modeling, GIS, and visualization to reshape the way Cultural Heritage can be digitized and virtually accessed, even remotely via the web. Indeed, one of the most significant consequences of the introduction of digital 3D models in the heritage field is the opportunity to use them as highly effective and intuitive means of communication or as interfaces to share and visualize information collected, for example, in databases. Given the possibility to link 3D geometries to external data, 3D models can be analyzed, split in sub-components, and organized following proper rules for a range of purposes. For example, in the case of (modern) buildings, where geometries, topologies, and semantic information are organized in a BIM (Building Information Modeling) system, such a system serves as a shared knowledge resource to support decisions. The BIM concept, extended to Cultural Heritage, facilitates the organization, storage, use, web-based visualization, and communication of archaeological data. Thus the heritage community has realized the benefits of using computers, 3D models, and advanced web-based repositories to enrich their studies, and often at lower costs. The field has also received a boost from the entertainment industry where technologies normally used in video games and for movies have been adapted to heritage needs. Additionally, recent developments in teleimmersive technologies have resulted in new teleimmersive and collaborative systems in archaeology, leading to challenging and interesting perspectives for the future [1, 2, 3]. The use of these advanced 3D technologies to visualize, share, and query digital structures, objects or stratigraphic layers, can bring together data of different types and across many scales (ranging from individual excavation units to monuments or entire sites) to assist in the archaeological interpretations of heritage sites.

This paper presents a collaborative project between the MayaArch3D project partners and the HUMlab at Umeå University (Sweden) to control the QueryArch3D — a 3D WebGIS visualization and query tool that deals with multi-resolution 3D models [4] — with the Kinect platform. The work seeks to understand whether enabling users to move through an archaeological site and query 3D models in a Virtual Reality (VR) environment without touching the keyboard or mouse creates a sense of spatial awareness and embodiment that can enrich user experience and facilitate new avenues for archaeological research. After a summary of related and previous research, and an overview of the technical developments made in the project, the initial results exploring the potential uses of linking 3D GIS with the Kinect sensor for archaeological purposes are presented and discussed.

II. RELATED WORK

3D models help to represent and analyze the real world — past and present — in a digital way. Virtual Reality (VR)
environments simulate real or imagined places and contextualize 3D models. VR can serve as a medium not only to visualize and communicate Cultural Heritage information, but it can also serve as an environment to integrate and analyze 3D models, even in online applications [5]. When 3D models in a VR environment are connected to a database, users can query and interact with underlying archaeological data offering them a way to actively engage with archaeological objects to generate new knowledge. Some advantages of using 3D models within advanced GIS environments include:

1. Contextualization of 3D models in larger spatial and environmental context;
2. Extension of the scale of analysis beyond individual objects or buildings;
3. Possibility to dynamically and interactively perform queries or spatial analyses;
4. Option to analyze 3D models from multiple perspectives (e.g., bird’s eye, façade view, interior view);
5. Offers a sense of embodiment in architecture or landscape (sense of place) [6, 7];
6. More intuitive interaction with archaeological data;
7. Increase awareness of spatial relationships between objects and associated data [8, 9].

Powerful 3D management and visualization tools already exist, but they often have restrictions in the geometry (2.5D or low-res 3D) or they implement limited query functionalities for data retrieval, and few are web-based. Queries that are typical functions of current GIS packages are usually not available when dealing with detailed and complex 3D models. Different authors have proposed solutions for 3D data management (store, query, measure, annotate, etc.) and visualization, [10, 11, 12, 5] but, to our knowledge, no unique, reliable and flexible package is commercially available.

Simultaneous to the growing importance of 3D modeling and VR systems are the increasing research developments and investigations on gaming devices. These include multi-purpose sensors (e.g., Microsoft Kinect, Axus Xtion, etc.) that serve as human-computer interfaces allowing touchless remote control and using the human body, rather than a keyboard or mouse, to interact with a computer or video game [13, 14]. Originally designed as motion sensing input devices mainly for playing video games, developers and researchers are now exploring alternative applications for such devices beyond the original entertainment purpose [15, 16, 17]. For example, the Console Yourself project uses the Kinect to enhance an improvisation theater show—the show alternates between human actors and Kinect controlled acting avatar [18]. In terms of Cultural Heritage, a prototype application called “Aboriginal Dance for Kids” uses the Kinect sensor to enable children to be puppeteers controlling and moving puppets to convey information about indigenous dance movements and body paints [19]. In the archaeological field, researchers at the California Institute for Telecommunication and Information Technology (Calit2) have modified the Kinect to create a portable, low-cost 3D scanner called ArKinect to capture 3D data and then display these data in real time on a monitor or in a virtual reality CAVE [20, 21].

III. THE MAYAARCH3D PROJECT

In 2009, the MayaArch3D project was founded to explore the possibilities of 3D digital tools and Geographic Information Systems (GIS) for archaeological and art historical research on ancient Maya architecture and landscapes. Bridging the humanistic and scientific disciplines, the project brings together art historians and archaeologists with computer scientists and specialists in Geomatics. As a case study, the project selected the ancient Maya city of Copan, Honduras — today, a UNESCO World Heritage Site (Fig. 1). The major research goal of the project is to understand how built forms and natural landscape features communicated information and structured social experience during the reigns of Copan’s 13th and 16th rulers (AD 695-820). Towards this end, traditional fieldwork methods coupled with digital 3D recording and modeling tools are being used to collect, analyze and assist in the interpretation of archaeological data [22, 23, 24, 25]. In order to visualize and query heterogeneous information, a prototype 3D WebGIS tool, QueryArch3D, was developed.

Figure 1. A view of the Temple 22 at Maya site of Copan, Honduras.

IV. TECHNOLOGY DEVELOPMENTS

A. QueryArch3D

QueryArch3D (Fig. 2) is a visualization and query tool that links multi-resolution 3D models to archaeological data that are accessed while navigating in a VR environment [4]. The QueryArch3D is tailored to the needs of researchers working in the heritage field and has the following functionalities:

1. Handles multi-resolution 3D models;
2. Queries both geometries and attributes in the same virtual environment;
3. Supports 3D visualization and navigation of the models;
4. Permits access to the content locally or online.
To satisfy these needs, the tool has two main components: (1) data modeling and storage in a DBMS and (2) visualization. The DBMS uses PostgreSQL with the PostGIS extension to reduce data heterogeneity and allow storage of both non-spatial and spatial data.

For the interactive navigation and 3D visualization, the tool uses Unity3D, a game engine development tool for the creation of 3D interactive contents accessed off-line and on-line using a free web player plugin. Finally, a PHP interface links Unity 3D and PostgreSQL allowing the data retrieval from the database and the (on-line) visualization. The system is organized into four Levels of Detail (LoD) for the different geometric structures. While navigating through the VR system (aerial view, ground-based walkthrough and close-up view mode), the user can perform two types of attribute queries: a) overall inquiry to the whole dataset or b) standard query selecting a 3D object and visualizing the associated attribute values obtained from the database.

The prototype tool currently contains:

1. A virtual landscape of Copan that covers 24 km²;
2. 3D schematic models of over 3,000 ancient structures;
3. A computer graphic model of an hypothetical reconstruction of the main civic-ceremonial complex;
4. Different reality-based 3D models of sculptures and stelae.

B. Kinect – A remote touchless controller

Kinect is a motion sensing input device by Microsoft for the Xbox 360 video game console and Windows PCs. The Microsoft Kinect is a low-cost device composed of different sensors (RGB camera, infrared depth sensor, motorized tilt function, and a microphone array) that enables users to control and interact with video games (or PCs) without the need to touch a game controller. The practical ranging limits for playing video games are 1.2–3.6 m. Third-party communities have quickly grown in the last few months and several software development kits have been released in order to use the Kinect device for other 3D applications rather than only for video game issues. Geometric investigations of the gaming device have been presented in [26, 27].

C. QueryArch3D controlled with the Kinect

The initial development of QueryArch3D focused on developing the LoD structure and the 3D WebGIS while movements in the VR were programmed with mouse navigation. Therefore, in order to allow a Kinect-based control and navigation of the QueryArch3D system, the mouse navigation commands needed to be reprogrammed, based on keyboard commands and gesture recognition features. First, the keys displayed on both desktop and laptop keyboards were selected and mapped to the English language keyboard. Second, using the Flexible Action and Articulated Skeleton Toolkit (FAAST) [13, 28], the keyboard input triggered by body posture and specific gestures are emulated and thus used to control and navigate the VR environment. Skeleton actions require the user to initially perform a skeleton calibration while then the body-based control mechanism works automatically. Different body movements were programmed to certain distances or degrees to trigger specific keyboard commands. FAAST was programmed and tested for two skeleton modes: (1) Full Body and (2) Upper Body. As an example, Table 1 lists the keyboard commands and gestures for the Upper Body mode.

### Table I. Keyboard commands and Kinect gestures for the remote touchless control of the QueryArch3D system (* LOD4)

<table>
<thead>
<tr>
<th>Keyboard Command</th>
<th>Onscreen movement</th>
<th>Kinect gestures</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>Move Forward</td>
<td>Right arm forward</td>
</tr>
<tr>
<td>S</td>
<td>Move Backward</td>
<td>Left arm forward</td>
</tr>
<tr>
<td>Q</td>
<td>Left Turn</td>
<td>Rotate body left</td>
</tr>
<tr>
<td>E</td>
<td>Right Turn</td>
<td>Rotate body right</td>
</tr>
<tr>
<td>Y</td>
<td>Look Up</td>
<td>Lean backwards</td>
</tr>
<tr>
<td>H</td>
<td>Look Down</td>
<td>Lean Forwards</td>
</tr>
<tr>
<td>T</td>
<td>Start walkthrough or enter LOD4</td>
<td>Right arm down</td>
</tr>
<tr>
<td>O</td>
<td>Open information (access database)</td>
<td>Right arm up</td>
</tr>
<tr>
<td>I</td>
<td>Close information</td>
<td>Left arm up</td>
</tr>
</tbody>
</table>

*Keyboard Command* | Queries/Interaction | Kinect gestures |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>Zoom in</td>
<td>Right arm forward</td>
</tr>
<tr>
<td>X</td>
<td>Zoom out</td>
<td>Left arm forward</td>
</tr>
<tr>
<td>←</td>
<td>Rotate object left</td>
<td>Left arm out</td>
</tr>
<tr>
<td>↑</td>
<td>Rotate object right</td>
<td>Right arm out</td>
</tr>
<tr>
<td>↓</td>
<td>Rotate object down</td>
<td>Lean backwards</td>
</tr>
<tr>
<td>O</td>
<td>Open information</td>
<td>Lean forwards</td>
</tr>
<tr>
<td>I</td>
<td>Close information</td>
<td>Left arm up</td>
</tr>
</tbody>
</table>

D. The system setup

In order to use the QueryArch3D tool with a touchless remote controller (Fig. 3 and 4), the Kinect device is connected to a PC via a USB adapter. To run FAAST and emulate the keyboard input, a skeleton tracker must be loaded.
onto the PC. (FAAST is currently available for Windows only.) FAAST 1.0, the latest version of the software (released on March 30th, 2012), now supports skeleton trackers for both OpenNI (an open source framework requiring the PrimeSense Package and Microsoft Kinect Driver) and Microsoft Kinect for Windows (requiring Kinect for Windows SDK). Our set-up used an earlier version, FAAST 0.10, and employed OpenNI [29]. The upper or entire skeleton identification and gesture recognition are fully automated procedures; therefore, the illumination conditions of the environment are crucial and important in order to allow correct gesture identification and command transmission.

![Image of system setup with Kinect and FAAST](image)

**Figure 3.** The system set-up, with the Kinect device recording the human movements and FAAST emulating the keyboard input to control the QueryArch3D tool.

V. RESULTS

A. Technical aspects

The creation of a low-cost and portable touchless control system to navigate and query the QueryArch3D tool was successfully achieved. From a controller point of view, the success of the gesture recognition and keyboard / mouse emulation was environment and user dependent. The initial tests have identified six critical factors that influence the system’s performance. These are:

- **Light/reflectivity:** the system’s performance decreases with larger amounts of ambient light (i.e., natural light) and it also decreases in the presence of highly reflective objects such as glass furniture.
- **Distance:** the system’s performance increases the closer the user is to the Kinect although the small FOV of the Kinect might not entirely capture a user too close to the active sensor.
- **Direction:** the system’s performance increases when the user directly faces the Kinect; vice versa, when the Kinect is positioned off-center, maybe due to room-size constraints, the system’s performance decreases.
- **Angle:** the system’s performance increases when the Kinect is orthogonally facing the user (i.e., the Kinect itself is not angled).
- **User’s height:** the system’s performance depends on user’s height. This factor corresponds directly to the angle of the Kinect and thus, the optimal height placement of the Kinect is at the torso level of the user. Additionally, the actions in FAAST are programmed to be triggered at specific distances, for example, an information table can be programmed to be opened when a user raises his/her arm 20 inches above the head; however, if a child or person of small stature is using the system, then the action will not be triggered if he/she cannot raise his/her arms high enough above the head. The latest version of FAAST (v.1.0) includes a manual pitch (up/down angle) function for tracking users in order to account for people of different heights.
- **User motion speed:** the system’s performance increases when user gestures are relatively slow and well-defined.

These factors are not mutually exclusive. In other words, they work together to increase or decrease the system’s overall performance. Some of these factors could be minimized, reprogramming FAAST from the Full Body to the Upper Body skeleton mode and thus improving the system’s performances. Indeed in the first tests, the Full Body mode was employed in order to better simulate reality in terms of user’s movement and attempting to offer a more embodied experience. For example, in this mode the user moves his/her feet forward or backward to simulate movement in the VR. But the performance of the Full Body mode, however, was more heavily influenced by environmental factors, particularly light and reflectivity. In a large open space located in a windowless basement, the Full Body tests were quite successful. In contrast, in a smaller, glass-furnished and bright office, Full Body mode tests had limited success. Thus a FAAST reprogramming to the Upper Body mode was necessary: users were able to move closer to the Kinect, increasing the system’s overall performance and minimizing the negative impact of environmental factors (such as light and distance).

B. Archaeological application

GIS and 3D tools offer ways to overcome certain
limitations of 2D media (e.g., lack of multi-scalar and multi-perspective analysis). Using spatial databases to overlay and link datasets, a GIS helps people to interact with and understand data and to reveal complex relationships, patterns and trends that are not evident when using traditional, or non-spatial, databases that are dependent on 2D media [30, 31]. 3D models placed in a VR landscape can provide a sense of spatial awareness and embodiment that can help users to learn about indigenous spatial concepts and the organization or use of spaces in ancient cultures [8, 9, 32].

The QueryArch3D tool — being a 3D WebGIS — brings together GIS, 3D models and VR enabling users to access underlying archaeological data as they navigate and click on 3D models in the VR of the ancient Maya city of Copan. The release of the Kinect has made low-cost and portable controller-free interaction for PCs widely available, and recent studies have indicated that the multi-media and multi-sensory capacity of Kinect facilitates and enhances teaching and learning [33]. In the case of QueryArch3D, using FAAST and Kinect, users can perform gestures to interact with information stored in the spatial database calling up photos, videos, and textual descriptions as they move controller-free through the virtual environment. But, what advantages or disadvantages does this controller-free interface offer for user experience and new avenues of archaeological research? Does using the Kinect to interact with the VR’s 3D models and their underlying archaeological data augment the user’s sense of spatial awareness and embodiment, invite new ways to explore and interact with 3D data, and/or result in the creation of new knowledge? While our work is in its initial stages and the results are preliminary, we have begun to investigate these questions by exploring a specific scenario relevant to archaeological and art historical research at Copan.

A test case was performed to simulate the visual and spatial experience that late 8th and early 9th centuries visitors may have had as they walked through the city, climbed up to the top of the 30m high Acropolis and arrived at the East Court (Fig. 5) to enter one of the city’s most important temples — Temple 22. Then, standing at the temple entrance, high on the Acropolis of the city, what could they see? As they stood in the East Court, which sculptures were visible, what texts could they read? Could they see the elaborately carved doorway in the interior of Temple 22?

The spatial arrangement of forms and symbols (e.g., architectural, hieroglyphic) provide important clues for archaeologists to estimate the temple’s intended audiences and messages. Navigating these VR spaces using one’s body instead of a mouse is useful for more accurately simulating how this now half-ruined temple might have been experienced in the 8th and 9th centuries when it still stood. Being able to query information during the experience, without having to leave the environment or switch from the “embodied” experience to using a mouse or keyboard also allows the user to experience an uninterrupted flow of movement and information.

VI. CONCLUSIONS & FUTURE WORKS

Recent technological advancements have led to a proliferation of projects seeking to develop applications for 3D GIS or touchless motion control. However, to our knowledge
no project has linked these two technologies together. This paper describes our work to link these two technologies and offers preliminary evaluations of the advantages and disadvantages of using gesture-based interaction to navigate a 3D GIS for education and research in archaeology.

A. Technical aspects

The work successfully created a touchless control system to navigate and query a 3D GIS called QueryArch3D. With this system, users employ gestures to interact and query 3D models set in the VR landscape of the ancient Maya city of Copan (Honduras). Compared to other approaches, the user's interaction is natural-based rather than device-based (with mouse or keyboard). The system has three components: (1) QueryArch3D, the 3D WebGIS based on PostgreSQL and Unity3D, (2) Microsoft’s Kinect, a proprietary multi-purpose sensor offering controller-free interaction with VR applications, and (3) FAAST, a free software (based on the open source OpenNI platform) to link QueryArch3D to the Kinect sensor. The developed system (software plus controller) is low-cost and portable, requiring only a laptop, the Kinect sensor, and external display. However, the work identified some key factors that impact the performance of the system and thus must be kept in mind when setting up the system, particularly for use in public areas for educational reasons, e.g., museum exhibits or classrooms, where environment and diversity in users (particularly user height) affect the system's performance.

B. Application aspects

Given recent investigations into the potential of interactive technologies, such as Kinect, to facilitate teaching and learning [33], the MayaArch3D project is evaluating the advantages and disadvantages of controller-free navigation of 3D GIS for archaeological education and research. The preliminary results indicate that device-based interaction with the mouse/keyboard is more useful than gesture-based interaction for investigating the segmented and high-resolution 3D models (LOD4) because users can more easily zoom in/out, rotate and query the digital models. However, natural-based (controller-free) interaction using gestures seems to create a sense of embodiment and spatial awareness that provides users with a better sense of space and consequently a better sense of place. Moreover, as users move their “bodies” through the VR environment, the ability to click on 3D models and acquire archaeological information about those objects using a simple hand gesture instead of stopping to click on mouse, maintains their flow of movement through the VR. Generally speaking, controller-free navigation offers a different perspective or frame of reference (from device-based navigation) to explore archaeological sites. Coupled with the ability to simultaneously query 3D models, the integration of 3D GIS and controller-free interaction technologies uniquely allows the public to dynamically and actively explore cultural heritage and may prove also to lead to new avenues of archaeological research.

C. Future Work

The future work of the interdisciplinary project will focus on: (1) enhance the QueryArch3D functionalities and contents and (2) improve the controller-free navigation into the VR environment.

To enhance QueryArch3D tool, textures and transparency features will be added to the buildings, vegetation and hydrological features will be inserted in the natural landscape, additional GIS functionalities will be developed and, finally, a sound option will be added to enrich user experience. Currently, the device-based version of QueryArch3D permits users to measure distances, display line-of-sight between ancient buildings, and perform spatial queries to highlight subsets of 3D models within the virtual landscape. The touchless control system linking Queryarch3D to Kinect does not yet have such functionality. Finally, the capability for two active “players” to navigate and interact in QueryArch3D will be instituted in order to facilitate shared experience.

To improve the controller-free navigation based on the Kinect sensor, two possible strategies have been identified. The first option is to continue to use the FAAST toolkit. At the moment, FAAST offers only limited navigation control; however, it has developed a plugin for the Unity3D engine (soon to be released) that may improve navigation. The second option is to develop a custom interface that will allow much more advanced navigation control. For this purpose there are currently three options: (1) use the ZDK Unity3D bindings in the Zigfu Development Kit (works with OpenNI/NITE and the Microsoft Kinect SDK, but it is not free), (2) develop a custom interface in the .NET framework using Kinect for Windows SDK, or (3) develop a custom interface in the OpenNI framework.

Finally, the project will continue to evaluate the advantages and disadvantages of using a touchless control system to navigate a 3D GIS for archaeological education and research purposes, keeping an eye on the latest technological developments and exploiting all the possible solutions to benefit the cultural heritage field.

ACKNOWLEDGMENTS

This project was partly supported by two Digital Humanities Start-Up Grants from the National Endowment for the Humanities (#HD50583, #HD5097910), National Science Foundation (#1064648), Alexander von Humboldt Foundation (Germany) and HUMlab, Umeå University (Sweden). We also thank the Honduran Institute of Anthropology and History (IHAH) for the collaboration and permission to work at Copan.

REFERENCES


