Exploring Heritage Through Time and Space
Supporting community reflection on the highland clearances

J. McCaffery, A. Miller, S. Kennedy,
T. Dawson & C. Allison
University of St Andrews
alan.miller@standrews.ac.uk

A. Vermehren & C. Lefley
Timespan Heritage Centre
director@timespan.org

K. Strickland
University of Highlands and Islands
Keir.Strickland@uhi.ac.uk

Abstract—On the two hundredth anniversary of the Kildonan clearances, when people were forcibly removed from their homes, the Timespan Heritage centre has created a program of community centred work aimed at challenging pre conceptions and encouraging reflection on this important historical process. This paper explores the innovative ways in which virtual world technology has facilitated community engagement, enhanced visualisation and encouraged reflection as part of this program.

An installation where users navigate through a reconstruction of pre clearance Caen township is controlled through natural gestures and presented on a 300 inch six megapixel screen. This environment allows users to experience the past in new ways. The platform has value as an effective way for an educator, artist or hobbyist to create large scale virtual environments using off the shelf hardware and open source software. The result is an exhibit that also serves as a platform for experimentation into innovative ways of community co-creation and co-curation.

I. EXCAVATIONS, ARTISTS AND DIGITAL MODELS

Two hundred years ago the Strath of Kildonan was part of one of the most traumatic periods of Scottish history, the Highland Clearances. Areas of the countryside which had seen human habitation for thousands of years were cleared to make way for sheep farming. The repercussions of this change are still felt to this day and are a fundamental part of the local identity. The two hundredth anniversary of the clearances is an opportunity for local descendants, members of the diaspora and the wider public to critically reflect upon the significance and heritage of the clearance years.

Timespan Museum and Arts Centre is in the village of Helmsdale, at the mouth of the river Helsmdale which flows through the Strath of Kildonan in the county of Sutherland. Established in 1986, Timespan developed from a small local heritage centre to an award-winning museum and the only public contemporary art gallery in Sutherland. Timespan's focus is on relationships with people and long-term creative development; with an emphasis on new ideas and creative exchange. It aims to challenge perceptions of what can be delivered by a cultural organisation rooted in the distinctive, but fragile, socio-economic ecology of a large, rural and remote area.

To commemorate the Bicentenary of the Clearances Timespan has developed the Translocation program, a year long series of events and installations. This included an archaeological excavation of the Caen Township (one of the townships involved in the clearances) which then informed a photographic art project and a museum exhibit based around a 3D reconstruction of the site.
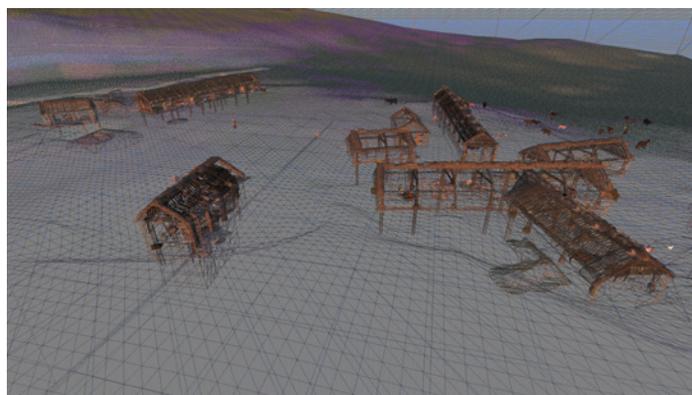

Fig. 1. Pre Clearance Wireframe Render of Caen Sutherland Township.

The community excavation of the longhouse was carried out (8th to 23rd June 2013) by Dr. Keir Strickland and Mr. Richard Barton of the University of the Highlands and Islands. They worked with the community to unearth new finds and facts about living in the Strath of Kildonan over 200 years ago. The excavation focused upon a single longhouse and associated building. The primary aim was to identify and describe the final domestic phase of occupation at Caen. Highlights of the dig included a fragmented copper pot-still (likely used for illicit whisky production), seen in Figure 2, and a shattered ceramic bowl, the deposition of which could be interpreted as evidence of an extremely sudden eviction.

An integrated art project, that explores notions of Home, complements the excavation project and the virtual reconstruction. Carolyn Lefley, invited as Artist in Residence at Timespan during the excavation, has created new work in response to these exciting events. As an artist who uses photography in her practice Lefley is interested in the parallels of the process of excavation, with peeling back the layers of earth to reveal evidence of the past and the indexical quality of a photograph to record reality. The excavation has provided 'images' of the way Kildonan's people lived. Part of Lefley's response has been to collaborate with descendants of Caen, who have been volunteering on the dig, by photographing them within the old longhouse ruin (Figure 3). Other community activities include a photography ramble around the Gartymore part of Helmsdale, where many of the cleared were housed, and the creation of a Clearances story patchwork quilt with local teenagers. Lefley's final output for her residency is an exhibition of photographs printed onto local stone, which tell stories of home, migration and diaspora. 'The Diaspora Stones' are on display in Timespan alongside excavation finds.

To create the exhibit, existing Virtual World architecture

Fig. 2. Still similar to that discovered in Caen.

Fig. 3. Carolyn Lefley photographing a home within a home. Photo credit: Carrie Kitchin

Fig. 4. The recreation of the outhouse where the still was found.

has been extended to allow it to be used as a platform for museum installations. Virtual Worlds are networked 3D environments. An immersive environment is one where the user feels more directly engaged within the environment than with traditional 3D paradigms. A user is surrounded by screens of a size and configuration such that the experience extends into their peripheral vision. This is supported by controller free interaction mechanisms to keep the user directly involved.

The system brings together Virtual World technology, natural user interaction (NUI) and large scale visual environments to create an interactive 3D digital interpretation of a site of significant cultural value. This 3D model is based on traditional scholarship but brought to life using current and developing technologies designed to utilise emergent digital literacies. The reconstruction provides a third dimension of community engagement and feedback. It incorporates both images from the residency and evidence from the excavation, thus helping visitors visualise the township of Caen.

II. FROM STORYTELLING TO DIGITAL NARRATIVES

The Timespan storytelling room is a 5m^2 space with audio and visual equipment which projects community created animations and films of folklore, legends and documentaries. This functionality has been extended by the installation of the TARDIS virtual world system into the room, building on its role as a place for the local community to share its stories with the wider world and adding new digital ways of interacting and exploring historic narratives[1].

The Caen township cluster is made up of the eight individual structures shown in wireframe view in Figure. 1. To the east, set slightly apart, is a longhouse. To the North is a corn drying kiln. In a cluster are five more structures: a barn, two longhouses, and another, smaller structure which is shown in the model as under construction. In the area at the centre of the model are some hay bales. To the south of the cluster of houses animals graze. There are a number of residents positioned around the area.

The layout of the virtual space is based on real world measurements. Initially a survey was done using a digital theodolite and the data from this overlaid onto the contours from the Ordinance Survey maps. The placement of reconstruction buildings (fig. 11) were based upon these initial measurements. The furnishings and fittings inside the eastern longhouse are based on historic accounts of contemporary

longhouses. The byre end features a cow being milked, the centre of the longhouse has a loom, and the living area is quite complete with a fire, box bed, chest of drawers full of crockery and nook in the wall with a bible (fig. 7). The initial reconstruction helped provide context for excavation volunteers.

Contour data was later supplemented with differential GPS readings taken as part of the excavation, enabling the landscape and placement of buildings to be refined. The results of the excavation informed the reconstruction of the second central longhouse and its auxiliary structure. The fireplace was central, the southern end was cobbled and contained a box bed, internal dividing walls were located and the byre end was gravel covered with beaten earth. There were rough paving stones outside the entrance and peat stacks along the external eastern wall. The whisky still was located in the western auxiliary structure (fig. 4).

The excavation which informed the reconstruction was carried out by volunteers. It was local volunteers who found the smashed sherds of the dairyware creamer on the cobbled floor of the longhouse and it was a descendent of a Caen resident who found the first fragment of the whisky still in the ancillary structure. These physical finds are on display in the museum, with their virtual counterparts situated within the reconstruction to give visitors a real sense of the link between the site today and the history of the place.

With no visitor present in the storytelling room the camera follows a simple loop through the virtual space inviting visitors into the exhibit. As soon as the visitor walks into the room a Microsoft Kinect®(see below) picks up their presence and the story of Caen begins. The story is communicated through a guided tour around the reconstruction. It is split into twelve sections, each one highlighting a different feature of the model, and therefore a different aspect of how life was in Caen. The camera movements are synchronised with an audio track, recorded by a local resident.

By default the guided tour plays until the end and then goes to the main menu (fig. 9) but the visitor can choose to skip straight to the main menu at any point. The main menu is split into two sections. On the left are the five stories that made up the original storytelling room. The options on the right allow the visitor to re-play the guided tour, to explore the model with an avatar, to view slideshows about the township life and about the excavation or to explore the model without an avatar.

Choosing to explore with an avatar puts the system back into the virtual world. In the centre of the screen is a girl, in

---

[1]A short video demonstrating the exhibit is at: http://bit.ly/1ag0WCj. The attached Vimeo account features more videos on the topic.

Fig. 5. Local primary school photography workshop at the excavation site. Photo credit: Jamie Brett

Fig. 6. Carolyn Lefley photographing Gerry Wood, a Caen descendant. Credit: Carrie Kitchin

Fig. 7. The kitchen area of the longhouse.

modern clothes as in fig. 11. As the visitor gestures with their upper body, the girl walks through the space at a fairly leisurely pace. If the visitor chooses to explore without an avatar they get free control of the camera. This camera can move forward or backwards, up or down. It can yaw left and right. It cannot tilt and it cannot move laterally.

III. ENABLING EXPLORATION THROUGH MOVEMENT

To interact with the system the user stands in the middle of the room with the three screens taking up approximately $150°$ of their field of vision. The goal is to give the user direct control over exploration. Having a physical device to manipulate would add an extra layer between the visitor and the exhibit, consequently the Kinect was chosen as a control device.

The Kinect is a peripheral, developed by Microsoft as a controller for the XBox360®console. It works by using infrared light to generate a depth map. The Kinect uses this depth map to get accurate values for the three dimensional Cartesian coordinates of each joint that makes up a user's skeleton. These are supplied by the Kinect in near real time through a USB bus. These coordinates are then parsed to interpret gestures and control software. With no calibration or peripheral, users are able to interact with the computer. This style of input is known as Natural User Input (NUI). NUI input is ideal for our application as the user can simply stand in the centre of the room and interact without ever being presented with a physical device.

There are several different approaches for using the Kinect to control the navigation of 3D spaces. In an approach adapted from touch screen technology, grab and zoom gestures are used to provide a gestural interface to Google Earth® [1], [2]. 'Active areas' combined with arm gestures are another control mechanism. 'Active areas' are areas in space which will cause movement if the user positions their hand to be inside them [3]. They are used for CAD [4] or Archaeology [5] applications. Other approaches are based on games such as 'Dance Dance Revolution', where the user's foot position signals movement intention [6]. A technique similar to the 'active area' mechanism involves twisting or leaning [4]. The final mechanism involves interacting with an onscreen control. The user moves a cursor over a dial icon on the screen [6]. The position of the cursor on the dial is then used to initiate movement. This is adapted from touchscreen applications.

An important influence on the design is the use pattern. When the visitor gets to the storytelling room they should have an informative easily accessible experience within a few minutes. Because of this the initial presentation is automated. If visitors want to dig deeper they can follow simple instructions which enable menu navigation and free form exploration of the reconstruction.

As mainstream access to NUI devices is a recent phenomenon there is not yet a base of shared experience and design from which to draw gestures for common tasks. In other fields, for instance touch screen interface design, their use is widespread enough that a standard set of gestures has started to emerge. There is not yet this level of digital literacy for NUI devices and in the exhibit no assumptions could be made about what gestures the visitor would already know.

The space in which the system is installed is limited and closely matches the limits of the Kinect so a good field of view is achieved. The gestures were designed such that the Kinect would have the best possible chance of interpreting them correctly. The design also takes into consideration the technical limits of the Kinect.

In the storytelling room the gestures used are active area or angle based and arrived at through extensive user testing carried out in a number of environments. The prototype was taken into a secondary school and tested with groups of pupils. It was taken in to the science museum in Dundee for two days of public demos, it was tested in the School of Computer Science in St Andrews with various researchers and it was trialled in the Timespan museum with members of the local community. Evaluation was carried out on these trials in the form of observations and unstructured interviews.

Walking forward is achieved through the 'push' gesture. The user extends one or both hands out in front of them. Once the hand reaches a threshold distance away from their chest the avatar starts to move forward. In flycam mode the further forward the hand is pushed the faster the camera moves. This is cumulative so using two hands results in faster movement than one. Yaw is mapped to three separate gestures: Twisting the shoulders, leaning (head relative to navel) and swinging one or both arms right and left (whilst the arm is extended forward). Flying is based around an 'aeroplane' or T gesture. The user extends their arms as a child would when pretending to be an aeroplane. Once extended, if the hand is raised above the shoulder the avatar will fly up, and dropping the hand below the shoulder means the avatar flies down.

Evaluation showed the system to be effective. Users are able to understand the push and twist gestures almost imme-

Fig. 8. Timespan Heritage Centre.

Fig. 9. Menu greeting visitor on entry to storyroom.

Fig. 10. Descendants Exploring the Township.

diately, which allowes them to navigate freely through space. Some calibration was necessary to make sure that a neutral gesture (hands by the side) produces no movement. False positives, where the user moves the camera or avatar without intending to, tend to be more disruptive to user experience than false negatives, where they attempt to generate movement but are not able to initiate the gesture. When observing there was a marked tendency for users to 'panic' and end up moving through the point they were aiming for. Emphasising that dropping their hands to their side would cause all movement to cease often solved this problem.

Another observation was the difference between flycam and avatar control. When controlling an avatar some basic limits are imposed on the user. The avatar cannot move through walls so objects represent obstacles. Also the camera is anchored directly behind the avatar's head. This means that as the avatar turns the camera swings round behind the avatar in an arc. By contrast the flycam mode is freer. The camera can move through any object, with the exception of the underlying terrain. In avatar mode movement speed is binary, the avatar is either walking or it is not, flying or not. In flycam mode all axes of movement are analogue in feel. The further forward the user pushes their hand, the faster the camera moves. The same goes for flying.

The two modes split users. The original hypothesis was that users would prefer avatar mode as it is more controlled, however this was not always the case. Some users preferred the precision allowed by the greater freedom/analogue control of the flycam, others preferred the more obvious metaphor of the avatar.

IV. OPENSIMULATOR: 3D APPLICATION SERVER

In the past two decades there has been a considerable rise in the use of 3D visualisation technology in cultural heritage [7]. More recently Petridis et. al [8] have covered the use of traditional video game 3D engines to render cultural heritage reconstructions. Often when 3D technology is applied to museums it takes the form of virtual models of physical artefacts [9]–[11], sometimes located within Virtual Museums [12]. Museum installations based around exploring 3D content go back to the turn of the century [13], [14]. Virtual World recreations of cultural heritage sites is fairly rare [15].

The OpenSim project is a virtual world application server which integrates with the World Wide Web. At the core of an OpenSim application is an extensible, mutable 3D environment users inhabit through the proxy of an avatar. Other users are able to see the avatar's actions and engage in synchronous

communication through movement, gestures, text chat, and voice. This projection of presence enables both collaborative and competitive exploration.

An OpenSim virtual world is made up of three discrete components: the *client*, the *simulations* and *grid services*. A *client* allows the user to connect to a server and provides an interface for viewing and interacting with the content the server provides. It provides control over lighting, view distance and fidelity of the data. There are several clients including: SecondLife, Firestorm, Imprudence and Hippo. *Simulations* are responsible for tracking all the objects and avatars contained within a virtual space. They receive updates from, and distribute updates to, clients that are connected to them. Each simulation makes use of services, to retrieve information about avatars buildings and landscapes. The *grid services* provide the information required by a system to both run the *simulations* and to control access. These may be located on the same machine as the *simulations*, on a single machine or distributed across multiple servers. Services include user authentication, and provide content.

V. TARDIS SYSTEM DESIGN AND ARCHITECTURE

The TARDIS is the name of the installation as a whole, projectors, computer, coordinator software, Viewer software and OpenSim server. The TARDIS enables a user to explore virtual world content in an immersive 3D environment. The panorama they see is created through orchestrating multiple projections at heterogeneous orientations into a single coherent image. The user's movement and interactions are controlled through natural user input.

The TARDIS is made up of the following components as illustrated in fig. 14. The *Virtual World Service* provides textures, shapes and avatar control information which enables cameras to be located within the 3D scene. The Microsoft *Kinect* detects user presence and movements; it provides RGB (red, green, blue) images, depth maps and skeleton data which the system interprets as natural user input, thereby enabling the user to control the system. *NUILib [16]* hides the IO details of the Kinect and turns the raw skeleton data into a series of abstractions which enables *Chimera* to focus on the logic of interpreting the skeleton data. Hiding the IO mechanism keeps higher system levels device agnostic. *Chimera* is the core of the system. It interprets skeletal abstractions and uses these to control camera and cursor. Camera position, orientation and view frustrum are calculated and supplied to the Virtual World Proxies. Commands and cursor movements are interpreted and passed to the *Overlay Windows*. *Overlay Windows* provide a common control interface which enables
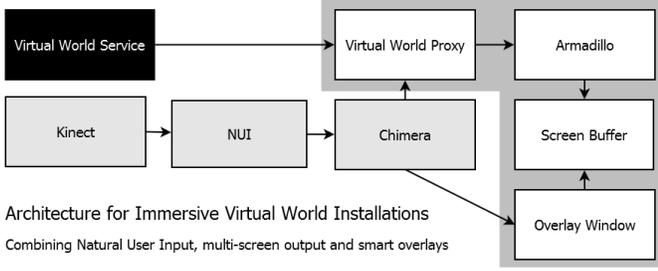
Fig. 14. Architecture of TARDIS immersive installation system.

virtual worlds to be combined with other applications to create integrated installations. *Overlay windows* manifest as a display which sits on top of *Armadillo*. They can mix transparent and non transparent areas to augment the view in the world or they can completely obscure it to provide menu screen functionality.

A *Virtual World Proxy* provides a mechanism for the core Chimera application to control the view in *Armadillo* without being tightly coupled. It acts as a middleman, intercepting all packets between the Virtual World service and its *Armadillo* instance. As it intercepts this packet stream it can add packets and pull data from packets received to control *Armadillo* and react to it. *Armadillo* is a Virtual World client specifically designed to allow external control of both the camera location, independent of any avatar, and the view frustum. This allows multiple instances to work together to create a larger, coherent image. Virtual World clients render the geometry and camera information streamed by the Virtual World service as a 3D space projected onto a 2D screen. *Screen Buffer* is the final step in the graphics rendering pipeline. It represents the 2D array of pixels which can be displayed by a screen or projector. The screen buffer for a frame within the TARDIS is made up of the in world view, as rendered by an *Armadillo* instance with an *Overlay Window* drawn over the top.

A *frame* is associated with a single display device. The goal is for the image rendered on the display device to correspond with the view one would get if one stood at the designated eye location and looked through a window positioned precisely where the display device is located, assuming the window looked onto the Virtual space. The view through each frame is the output of a single Screen Buffer. Figure 12 shows the configuration of the three frames which make up the TARDIS installation. Figure 18 shows a view through the three frames. The view through multiple such frames can be combined into a larger, coherent image to supply a strong feeling of immersion.

Chimera tracks camera position and orientation in virtual space; it provides a mechanism for specifying and dynamically updating an arbitrary number of frames defined relative to one eye position, which together render a large viewing area; it provides two plugin mechanisms; one allowing an implementation of an output interface to be tied to each frame, the other a generic plugin mechanism allowing external code to access and control the core functionality. These are implemented as two classes and two interfaces. For any given instantiation there is only one core object, plugins, and a number of frames.

Frames enable visuals to be presented on more than one screen. This allows for a larger viewing area in multiple planes. Splitting the viewing area onto multiple planes allows for a

synchronised screen that wraps around the visitor and fills their peripheral vision. Figure 12 shows a 3D scale model of the storytelling room with three screens. A frame is represented as an Object, which records its position, size and orientation and performs all the calculations necessary to create virtual camera set ups that allow the virtual view rendered by each viewer to align correctly. Frames calculate three values; $\vec{n}$, an $\vec{u}$ and a projection matrix. $\vec{n}$ represents the normal for the plane the frame is on. $\vec{u}$ (up) specifies the third axis of rotation (roll) that cannot be specified by $\vec{n}$. $\vec{n}$ and $\vec{u}$ are specified in virtual world coordinates and take into account both the orientation of the physical frame and the orientation of the virtual camera. The projection matrix controls the field of view so that the rendered view corresponds to what the visitor would see through an equivalent physical window. This means adjusting for width and height of the physical screen and offsetting the view correctly in the case where the direction the camera is facing does not align with where the virtual screen needs to be. Equation 1 calculates $\vec{n}$, Equation 2 $\vec{u}$. The projection matrix is shown in Equation 3.

$$\vec{n} = q_c * \vec{n_p} \tag{1}$$
$$\vec{u} = \vec{n_p} x q_c * |\vec{X}| \tag{2}$$

q_c is a quaternion representing the orientation of the virtual camera. $\vec{n_p}$ is the normal vector for the physical screen on which the frame will be projected.

A custom projection matrix is needed because the views must cover specific areas. Projection matrices would normally be calculated based around the resolution of the screen and be centred; in the TARDIS configuration the projection matrix is a function of the physical dimensions of the surface on which it is being displayed. The projection matrix accounts for the possible skew of the view, caused by displacement between projector and eye positions. This skew is shown in fig. 13 which shows the three screens in storytelling room and includes four normal vectors. For each screen its normal vector is presented starting at the origin and again, starting at the centre of the screen. Because the left and right screens are arranged such that their centres do not fall on the normal as drawn from the origin their view frustum has to be offset. The projection matrix for each frame is show in Equation 3.

$$\begin{matrix} \frac{W}{2*(\vec{P}\cdot\vec{n_p})}, & 0 & \frac{2*O_h}{W} & 0 \\ 0, & \frac{H}{2*(\vec{P}\cdot\vec{n_p})}, & \frac{2*O_v}{H} & 0 \\ 0, & 0, & \frac{-D_f+D_n}{D_f-D_n}, & \frac{-2D_f*D_n}{D_f-D_n} \\ 0, & 0, & -1, & 0 \end{matrix} \tag{3}$$

In Equation 3 W and H refer to the width and the height of the physical screen which the frame represents. $\vec{P}$ is the difference between the centre of the physical screen and the position of the visitor's eye (which is generally taken to be at the origin). $\vec{P} \cdot \vec{n_p}$ is the distance between the eye and the physical screen *along* $\vec{n_p}$, calculated using scalar projection. O_h and O_v are the horizontal and vertical offsets between the centre of the physical screen and where $\vec{n_p}$, starting at the eye position, intercepts the plane of the frame. O_h is marked as 'Offset' in Figure 13. D_f and D_n are the near and far clipping planes.

Plugins provide controlled access to core Chimera components. Plugins range from taking Kinect input, to helping position projectors to make frames line up correctly, to recording statistics on how the system is being used to querying
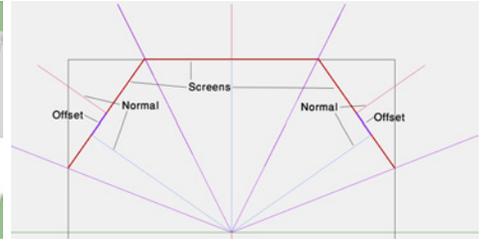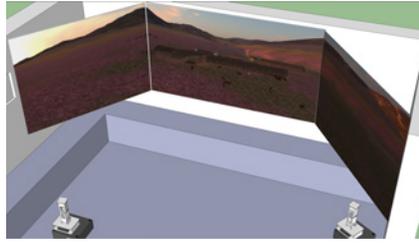
Fig. 11. An avatar standing by the longhouse. Fig. 12. The layout of the storytelling room.

Fig. 13. Top down view of the layout of the screens in the storytelling room.

the OpenSim service for heightmap data and using that to ensure the camera does not go through the ground. The three most visible plugins are discussed below. The *Kinect Plugin* is built on NuiLib. NuiLib allows easy access to Nui devices. It defines four basic Components (Scalars, Vectors and Conditions) which are then automatically updated as the values from the Nui device change. Operator overloading and a series of utility functions makes it easy to set up relationship between raw input from the Kinect and processed values. Listeners can be added to the Kinect or to Components to be notified when something has changed. NuiLib also provides events for when a new Skeleton has been detected, when a Skeleton has been lost or when the Skeleton that has been detected changes (i.e. a different person takes control of the system).

The *Overlay Plugin* is the Plugin most visible to the user. When the guided tour starts, the user is seeing the Overlay. It plays the audio track and renders the subtitles. The Overlay is performing two functions. It is providing a framework for drawing text and images over the Virtual World client to add functionality; this is managed by creating separate $OverlayWindow$ forms, which are borderless and support transparency. The Overlay also adds state to the system. Both of these functions are designed to be flexible. Everything is defined in XML and the framework used to create them is extensible so new overlay features can be integrated into the system as necessary.

The *flythrough plugin* allows a series of camera moves to be recorded in XML and played back. This is the mechanism that implements the guided tour. It comes with Overlay components to control how the flythroughs are played. Flythroughs are built from three general types of events; Position events, which control where the camera is, Orientation events which control where the camera is looking and Combo events which contain a sequence of Orientation events and sequence of Position events. Each event has a specified length and information specific to its type. Fly-throughs are broken down into steps. Each step is one Combo event. This break down allows playback to be segmented, which is helpful when creating flythroughs.

In TARDIS output is through an OpenSim server connected to a modified Firestorm viewer. The system is controlled using the GridProxy library which is part of the libopenmetaverse[2] project. This proxy sits between the client and the server so that all packets which are transmitted between them go via the proxy. This allows information to be extracted from the packets as they pass through and for extra packets to be injected into the stream. To allow Chimera full control over a Virtual World
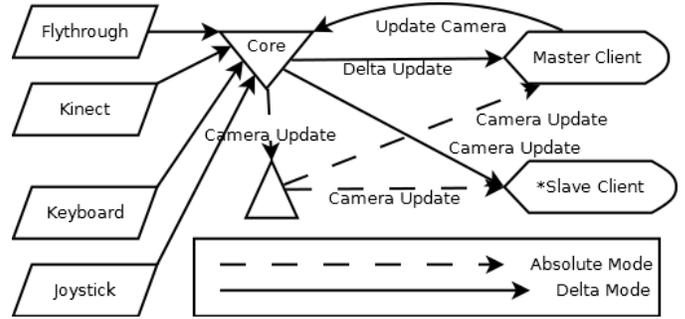

Fig. 15. Sequence of events to synchronize viewers based on camera updates.

viewer the Armadillo viewer adds support for three new packet types: camera packets control the position and orientation of the camera; frustum packets control the view frustum of the viewer; remote control packets supply deltas which are used to move the avatar.

In order for avatar control to work correctly one screen must display the avatar, as per usual, whilst the others show a view that synchs up with the avatar's position. There must also be one viewer where the avatar is being remotely controlled (this is called the master viewer and normally also functions as the screen which is displaying the avatar, all other viewers are slave viewers). The core of the system is designed so this functionality can be created without requiring any direct modifications to how it functions. To do this the core supports two separate modes for updating the camera, Delta and Absolute. Delta mode supplies deltas to move the camera based on its current position. Absolute specifies precisely where the camera should be in global coordinates. In Delta mode, when an update is sent to the core from the Kinect plugin the core generates a delta update event. At this point the master viewer sends a RemoteControl packet to its viewer. The viewer then processes this update, moves the avatar and sends an AgentUpdate packet back through the proxy. The master controller intercepts this packet and parses it for camera position and orientation. The master controller then passes this information to the core as a CameraUpdate. This then triggers a CameraUpdate event which all slave controllers pick up. They then update their viewers using Camera packets. This process is shown in Figure 15.

VI. EVALUATION

This section discusses system measurements, user feedback for both the usability of the system and its value as well as reflection on use from the curator's perspective.

Evaluation of the installation system's performance was undertaken, focussing primarily on frames per second (FPS).
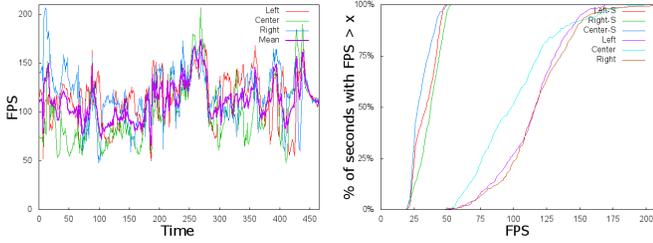
Fig. 16. FPS vs. time w/o shadows. Fig. 17. Cumulative Frequency Distribution of FPS w/ and w/o shadows.

Fig. 19. Interface (SUS) Questions' Fig. 20. Value questions' Likert Adjusted Likert Scores. Scores.

	Question
A	The exhibit helped me imagine what it would have been like to live in Caen.
B	I would now like to find out more about life in the highlands.
C	Exploring the world on my own added more to what I learnt during the guided tour.
D	I enjoyed using gestures to control my exploration.
E	Walking about the reconstruction helped me think about how people used to live.
F	I found the exhibit difficult to understand.
G	I would enjoy exploring other areas of history using a similar system.
H	I felt that I did not see all of the exhibit.
H	Some of the camera movements in the guided tour felt uncomfortable to watch.
J	The big screens made it seem more real.

TABLE I. VALUE QUESTIONS

User tests suggested that FPS of 20 or greater gave both a responsive experience and a smooth display. Several configurations were tested, before settling on a an Intel i7, Haswell quad core, 3.4ghz processor with an Nvidia Geforce 780 graphics card and 16GB of memory costing ~£1150 in August 2013 to drive the exhibit. Measurements were made at high quality graphics settings both with and without shadows. Shadows are a graphical setting that makes a considerable difference to the aesthetic effect of the scene but tends to result in dark interior spaces. The FPS over time for each screen, without shadows, is shown in Figure 16. Figure 17 shows the cumulative frequency distribution (CFD) of FPS with and without shadows for each screen. The CFD shows that without shadows there is a much higher range of FPS but they are almost uniformly considerably above the 25 FPS threshold required for sequential images to be interpreted as moving [17]. With shadows enabled the range of FPS is much narrower but still within a reasonable range for a smooth user experience. Our measurements demonstrate that the chosen commodity hardware is capable of delivering an appropriate user experience.

To evaluate users' experience a sample of fifteen visitors completed a questionnaire split into questions about the usability of the system and the value the exhibit adds. Usability was assessed with the standard System Usability Scale questions (SUS) [18], value with a set of ten Likert scored questions. The SUS scores were interpreted using the method outlined by Sauro [19], yielding a total score of 76.8. This is nearly ten points above the average SUS score. Given that the Kinect is a new input mechanism and that the SUS score was originally designed to evaluate keyboard based input this result is positive. The per question breakdowns in Figure 19 show high scores on questions measuring how comfortable and easy users found the system to use.

The value of the experience the exhibit supplies was evaluated by 10 Likert scored questions shown in Table I with the mean score for each shown in Figure 20. Users' responses indicate that the exhibit as a whole helped visitors to be involved with the history of the clearances and that the three screens added to the immersiveness. Two thirds of visitors strongly agreed that they would like to see other topics presented this way.

Direct observations of user interaction with the model have led the curator to reflect on the way that users connect with the virtual worlds. These observations have been drawn from a variety of workshops with people from all backgrounds and ages.

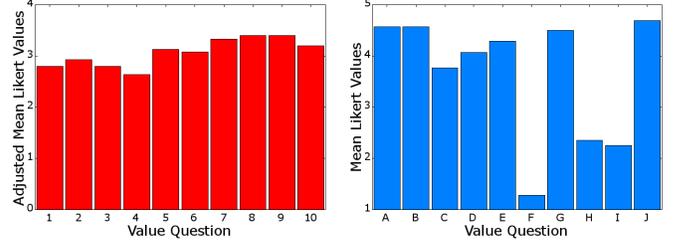To describe those interactions we apply concepts developed for the performing arts. In Timespan's storytelling room the model is presented on three screens. The three-dimensionality of the model suggests a similar set-up to a stage that the user is entering via bodily gestures while the body itself remains in the same position in the darkened room, the movement in the virtual world is hence half physical and half imagined. Dependent on who else is in the room the user might have an audience, or alternatively the user will meet other avatars in the virtual world who access the model from elsewhere.

Entering the virtual world does not resemble walking through a physical landscape. The user is aware of the virtuality when exploring, flying, and walking. The movements are abrupt, the senses are minimised to the visual, the user can walk through walls without hurting himself or fly high above the site, observing it with birds' eyes. The user can enter the virtual world with an avatar or without an avatar. Some people find it easier to navigate with a human substitute; others prefer to pretend to be themselves. Purposefully the avatar is designed as a modern person, in a T-shirt, wearing jeans. In contrast the people working on site are shown in costume. They are responsive and can explain their activities, or tell their stories. When talking to them they enact the same response each time.

Although the model is built to be historically accurate, the virtual reality creates a disruption for the user. In the virtual world the realms of an imagined past and a suggested future merge. The term de-familiarisation suggests that common habits, conventional perceptions and visual aesthetics are presented in a unfamiliar ways [20]. The user has to learn new gestures to navigate around the model. Automated movements such as walking are suddenly unfamiliar, real people are hidden in the representation of other avatars walking around, and although the virtual world aims to resemble reality, it differs from the physical realities we are familiar with.

These disruptions make the user aware of the perceptions and of the limitations of our body. The virtual world is a stage, a theatrical performance, in which we play the main part. From user observations in individual sessions and workshops it has

Fig. 18. A view generated from three screens arranged as half a hexagon.

become noticeable that most users are critical of the content they are viewing. This might be caused by the defamiliarisation but also by the different set of aesthetics that remind us of common realities but do not resemble our physical world. This estrangement from the familiar can also be described with the similar term of *Verfremdungseffekt* (commonly translated as alienation effect or estrangement effect). *Verfremdung* is a set of techniques that prevent the audience from losing itself passively in the narrative of the characters. Consequently, this leads the audience to become a consciously critical observer [21] and is in line with our goal of encouraging reflection on the clearances. In his theatre Brecht uses methods of transposing characters into the past and into the third person. Both techniques are used in the virtual world where the user can teleport into the past, and can become an avatar.

VII. CONCLUSION

This paper has demonstrated how emerging technologies such as Natural User Interaction, accessible 3D modelling and immersive environments can be brought together to create engaging and thought provoking narratives. We have a developed a system where visitors, with no training at all, are able to use a NUI devices to navigate in a 3D space. A system that combines traditional exposition mechanisms, and allows virtual world content to be integrated into a wider program of artistic and archaeological investigation. It provides a flexible architecture for Cave environments and for engaging with local history.

REFERENCES

[1] M. N. K. Boulos, B. J. Blanchard, C. Walker, J. Montero, A. Tripathy, and R. Gutierrez-Osuna, "Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation." *International Journal of Health Geographics*, vol. 10, no. 1, p. 45, Jan. 2011.

[2] S. Stannus, D. Rolf, A. Lucieer, and W. Chinthammit, "Gestural navigation in Google Earth," in *Proceedings of the 23rd Australian Computer-Human Interaction Conference on - OzCHI '11*. New York, New York, USA: ACM Press, Nov. 2011, pp. 269–272.

[3] J.-o. Kim, C. Park, J.-s. Jeong, N. Baek, and K.-h. Yoo, "A Gesture Based Camera Controlling Method in the 3D Virtual Space," *International Journal of Smart Home*, vol. 6, no. 4, pp. 117–126, 2012.

[4] F. Sen, L. Diaz, and T. Horttana, "A novel gesture-based interface for a VR simulation: Re-discovering Vrouw Maria," in *2012 18th International Conference on Virtual Systems and Multimedia*. IEEE, Sep. 2012, pp. 323–330.

[5] H. Richards-Rissetto, F. Remondino, G. Agugiaro, J. von Schwerin, J. Robertsson, and G. Girardi, "Kinect and 3D GIS in archaeology," in *2012 18th International Conference on Virtual Systems and Multimedia*. IEEE, Sep. 2012, pp. 331–337.

[6] P. Dam, P. Braz, and A. Raposo, "A Study of Navigation and Selection Techniques in Virtual Environments Using Microsoft Kinect," in *Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments*, ser. Lecture Notes in Computer Science, R. Shumaker, Ed. Berlin, Heidelberg: Springer, 2013, vol. 8021, pp. 139–148.

[7] E. F. Anderson, L. Mcloughlin, F. Liarokapis, C. Peters, P. Petridis, and S. D. Freitas, "Serious Games in Cultural Heritage," in *VAST'09: 10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST - State of the Art Reports*, 2009, pp. 29–48.

[8] P. Petridis, I. Dunwell, S. de Freitas, and D. Panzoli, "An Engine Selection Methodology for High Fidelity Serious Games," in *2010 Second International Conference on Games and Virtual Worlds for Serious Applications*. IEEE, Mar. 2010, pp. 27–34.

[9] D. Grammenos, X. Zabulis, D. Michel, P. Padeleris, T. Sarmis, G. Georgalis, P. Koutlemanis, K. Tzevanidis, A. Argyros, M. Sifakis *et al.*, "A prototypical interactive exhibition for the archaeological museum of thessaloniki," *International Journal of Heritage in the Digital Era*, vol. 2, no. 1, pp. 75–100, 2013.

[10] F. Bruno, S. Bruno, G. De Sensi, M.-L. Luchi, S. Mancuso, and M. Muzzupappa, "From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition," *Journal of Cultural Heritage*, vol. 11, no. 1, pp. 42–49, Jan. 2010.

[11] D. Monaghan, J. O'Sullivan, N. E. O'Connor, B. Kelly, O. Kazmierczak, and L. Comer, "Low-cost creation of a 3D interactive museum exhibition," in *Proceedings of the 19th ACM international conference on Multimedia - MM '11*. New York, New York, USA: ACM Press, Nov. 2011, p. 823.

[12] D. Tsichritzis and S. J. Gibbs, "Virtual museums and virtual realities," in *Proceedings of the International Conference on Hypermedia and Interactivity in Museums - ICHIM '91*, vol. Technical Report #14. Pittsuburgh, PA: Archives & Museum Informatics, 1991, pp. 17–25.

[13] J. Cremer, J. Severson, S. Gelo, J. Kearney, and M. McDermott, ""This old digital city" one year later: experience gained, lessons learned, and future plans," in *Virtual Systems and Multimedia, 2001. Proceedings. Seventh International Conference on*, 2001, pp. 49–56.

[14] V. DeLeon and R. Berry, "Virtual Florida Everglades," *Proceedings of VSMM98 Future Fusion: Application Realities for the Virtual Age*, vol. 2, pp. 458–463, 1998.

[15] L. M. Sequeira and L. C. Morgado, "Virtual archaeology in second life and opensimulator," *Journal For Virtual Worlds Research*, vol. 6, no. 1, 2013.

[16] J. Mccaffery, A. Miller, and C. Allison, "Gonna Build Me a TARDIS : Virtual Worlds for Immersive Interactive Experience," in *PGNET 2012 - Proceedings of the 13th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting*, 2012.

[17] R. T. Apteker, J. A. Fisher, V. S. Kisimov, and H. Neishlos, "Video acceptability and frame rate," *IEEE MultiMedia*, vol. 2, no. 3, pp. 32–40, Jul. 1995.

[18] J. Brooke, "SUS: A quick and dirty usability scale," in *Usability evaluation in industry*, P. W. Jordan, B. Weerdmeester, A. Thomas, and I. L. Mclelland, Eds. London: Taylor and Francis, 1996.

[19] J. Sauro. (2011, Feb.) Measuring usability with the system usability scale (SUS). [Online]. Available: http://www.measuringusability.com/sus.php

[20] V. Schklovsky, "Art as method," in *Russian Formalist Criticism: Four Essays*, L. Lemon and M. J. Reis, Eds. Lincoln (Nebraska) and London: University of Nebraska Press, 1965, pp. 3–24.

[21] B. Brecht, *Brecht on Theatre: The Development of an Aesthetic*. Hill and Wang, January 1977.