Developing multimodal virtual experiences based on low-cost games mechanisms, under the perspective of artistic creation

Dr. STAVROULA ZOI[1]

[1]Greek-French Master «Art, virtual reality and multiuser systems of artistic expression», Athens School of Fine Arts – Paris-8 University; Athens, Greece; voula.zoi@gmail.com

Abstract

This paper investigates the use of low-cost games mechanisms for prototyping and developing multimodal virtual experiences, under the perspective of artistic creation. The artistic project "Rowing courses" (Nikos Papadopoulos, 2016, http://www.nikospapadopoulos.eu/), a multimodal navigation experience in a virtual utopian space, is used as a case study. In this multimodal navigation experience, the participant resists to a perpetual fall through voice, while can navigate horizontally through a physical paddle. The virtual experience can also become collective and participatory, through the help of a performer. The implementation of the "Rowing courses" project was based exclusively on games mechanisms that were adapted to the needs of a virtual installation extended in the physical space. The study focuses on design issues and an emphasis is given on an open system architecture enabling the development of different modes of artistic actions and real time parameter configuration for adjusting the virtual experience.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Human-centered computing]: Virtual Reality

1. Introduction

The prevalence of low-cost digital interface mechanisms (e.g. body tracking devices, virtual reality headsets), most of which have evolved from within the computer games industry (e.g. Microsoft Kinect, Oculus Rift, Leap motion, Razer Hydra, Emotiv) has set new conditions in the field of creative technologies. The use of these mechanisms combined with powerful games engines (e.g. Unity, Unreal) enable the design and prototyping, with low cost and in short time, of almost any virtual and even multimodal interactive experience.

Specifically in the field of artistic creation, following the experimentations of important artists of the 90s, such as Jeffrey Shaw (see "Legible City, 1989, and "ConFIGURING the CAVE", 1996) [SHAW], a recent trend is developed that investigates new virtual experiences (sensory, kinetic, cognitive) through the "extension" of games mechanisms as to their functionalities (e.g. their adaptation to non-play scenarios) and conceptual range (e.g. metaphors they support). Their combination with techniques (current or past) from other fields (e.g. scenography, theatrical lighting, artistic installations, multi-channel sound, multiple projections, physical objects and constructions, performers and participation of the audience) helps smooth out weaknesses of those mechanisms, as to their aesthetics and functionality in non games scenarios. Some indicative, recent examples of artistic use of low-cost games mechanisms are mentioned below.

In the work "The Treachery of Sanctuary" [Mil12], 2012, Chris Milk, creates an interaction space through a triple projection. Three Microsoft Kinect body tracking mechanisms, hidden behind each one of the projections, help transfer visitors movements to a Unity based virtual space, enabling them to interact with 3D models of moving birds.

In her installation "Eunoia II" [Par14], 2014, Lisa Park, uses the Emotiv EEG mechanism to control, through her mind, 48 vibration pools of water. Distinct emotional values, such as frustration, excitement, engagement, and meditation, as retrieved through the device, are translated into sound waves that create vibrations in the pools of water placed atop speakers.

In their work “The Styx” [SV14], 2014, the creators (Filip Sterckx and Antoon Verbeeck) are developing a multi-sensory boat navigation experience, to a virtual representation of river Styx, which, based on the ancient Greek Myth, was the passage from the world of the living to the world of the dead. The virtual reality headset Oculus Rift is used in conjunction with 3D mapping techniques and a custom 360-degree virtual environment. The experience is enhanced with stereoscopic sound, smell, touch, as well as subtle real-life effects (e.g. the visitor is wetted with real water).

In “Επιλογή in Crisis” [San14], 2014, the participant is immersed, through the Oculus Rift headset, to a system in crisis, which is metaphorically and artistically represented as a structure of choices and interconnections. This crisis can be handled only through participants view (perspective) and choices. The choices are expressed as questions which the participant is asked to answer in real-time, through the direction of viewing as tracked by the device, thus deciding his own path of solving the crisis. The immersion which extends the concept of “being inside the crisis” and not a spectator, in conjunction with the direction of viewing tracking (head tracking) as provided by the device, create a system which not only develops a sensory or kinetic experience, but a more complex cognitive function, that of “choice”.

One approach that has been proposed many years ago, before the appearance of low-cost tracking devices, regarding the navigation and interaction in virtual spaces, is that of transferring functions from participants’ everyday experience (e.g. body interaction instead of using a specific controller). In this case, it is possible to incorporate, as an interface, a physical object which helps not only develop familiar conceptual metaphors [Eri95], but also normalize technology intervention. One of the most representative examples is Jeffrey Shaw’s “Legible City” [SHAW], 1989, in which the artist uses bike riding, a daily and favorite experience, as a navigation mechanism. Also, in the artwork of Char Davies “Osmose” [DAV95], 1995, the human functions of breathing and balancing compose the main navigation mechanism. These two human functions are tracked through a specifically adapted motion tracking vest.

The above two artworks, and other of similar scale developed mainly during the 90s by important artists [Pau03], define a distinct generation of artistic virtual reality prototyping and production, based on specialized and expensive equipment. In most cases, artists had to collaborate with major centers or festivals. Today, anyone is able to prototype a multimodal interactive system, by using low cost games interfaces and free software.

Nevertheless, this ease of technological prototyping doesn’t necessarily lead to a successful system. A number of critical issues must always be considered, concerning both the selection of the suitable mechanisms (e.g. devices, software), and the development of an appropriate architecture (e.g. protocols of data flowing) allowing a) the synchronization of different modalities of interaction (parallel or sequential), as in real world b) system configuration, even in real time, in order to enable adaptation of the virtual experience based on the objectives of the system [SMTM15], and c) the recording and exploitation of the generated data as semantic findings (e.g. logging of users behaviors as filtered through different virtual experiences) for scientific uses.

Moreover, human factors (e.g. perceived quality based on the desired objectives, technological intervention, mental states such as cognitive load or stress, required physical skills) play the most important role when developing a multimodal system [BRC96]. The investigation of those issues requires multidisciplinary approaches, with synergies between sciences such as cognitive psychology or medicine, and other fields such as design and arts.

The present study, examines design issues concerning the selection and adaptation of games mechanisms for developing multimodal virtual artistic experiences. Emphasis is given on an open architecture enabling the development of different modes of artistic actions and real time parameter configuration for adjusting the virtual experience. Throughout the study, the artistic project “Rowing courses”, of Nikos Papadopoulos (http://www.nikospapadopoulos.eu/), 2016, is used as a case study. It is a multimodal navigation experience in a virtual utopian space, developed through a combination of visual immersion, rowing motion and use of participant’s voice for balancing a perpetual fall due to gravity. The implementation was based exclusively on games mechanisms that were adapted to the needs of a virtual installation extended in the physical space.

The scientific part of the research has been developed within the framework of the Greek-French Master “Art, virtual reality and multiuser systems of artistic expression” (Athens School of Fine Arts, Paris-8 University) (www.eumaster.asfa.gr).

2. A multimodal virtual reality artistic system as a technological structure

The requirements of a multimodal virtual reality artistic system should be examined under a different perspective than that of a multimodal game or an application that leverages game mechanisms.

Art often pursues “the unpredictable”, in contrast with a usable application in which interaction should be predictable in order not to cause uncertainty to the user. Even most computer games (although the limits between a digital game and a digital artwork are indistinct [San2006]), are designed for a comfortable seated position of the player, good lighting conditions and clear rules of play. It is not coincidence that a wide range of gaming interaction mechanisms are still designed for a desktop computer, while body interaction mechanisms (e.g. Microsoft Kinect) require good lighting conditions and line of sight. On the contrary, artistic installations are usually set up in dark spaces and the participation of the visitors may involve standing position and physical movement. Moreover, the “enlargement” of a concept that an artwork may intend (e.g. fear, pain, thought) may conflict with the concept of “user friendliness” of a usable application (e.g. pain feedback in Painstation, 2006, of the “/////////fur//// art entertainment interfaces” artistic group).

The intervention of the technological elements that help produce the final illusion (e.g. controllers, cables, software) between the human and the system may in some cases need to be covered by specific adaptations, while in other cases it could serve the substance of the artwork (see[Man01], Chapter 4, “The Illusions”, subsection “Illusion, Narrative and Interactivity”).

Ultimately, a multimodal virtual artwork is a complex technological structure, that produces compositions of interactions between human and computer at different levels, and these compositions significantly define the aesthetics of the artwork. This technological structure should be easily configurable, and capable of operating seamlessly and adapting to difficult and unpredictable situations.

3. Case study: The “Rowing courses” project

The issues discussed in Section 2, are addressed through a case study, the “Rowing courses” project, a multimodal navigation experience in a virtual utopian space.

3.1. The virtual space and the 3D objects

The virtual space of the “Rowing courses” project is structured from elements with references to public works of art of the city of Athens (e.g. the equestrian statue of Karaiskakis in Zappeion) and Paris (e.g. the equestrian sculpture of Jeanne d'Arc at the Place des Pyramides). These public works of art are charged with the concepts of nation, destruction, salvation, death, hope. The public work of art is a screen on which are displayed collective emotions and conflicting standards of a society. They are symbols that attract and generate poles, in society, with conflicting emotions, developing the potential future exercise of collective violence.

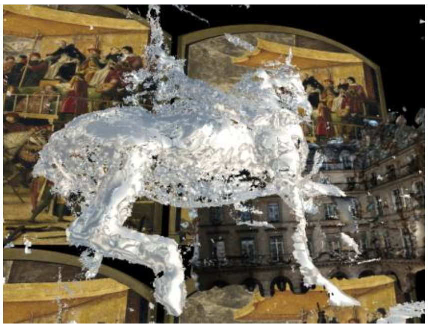


Image 1. A view of one of the levels of the virtual space, illustrating the distortion and the fragmentation of the 3D forms. Copyright © Nikos Papadopoulos 2016.

The selected artworks were deconstructed and reconstructed in digital 3D form. The dominant features are those of distortion and fragmentation of the form. The digital object is now immaterial, the form can be seen from the inside; it is hollow (Image 1).

3.2. Personal immersion or collective participation

The virtual experience can be either personal (a visitor is immersed through a virtual reality headset), or collective. In the second case, through a large scale projection and with the help of a performer, the virtual experience becomes shared, and the audience can participate through their voice. In other words, there are 2 modes of system operation, and the transition from one to the other is possible in real time.

3.3. Navigation through voice and body movement

The virtual space is structured as a vertical construction with different levels (one such level is illustrated in Image 1). Image 2 provides an overall view of the virtual space.

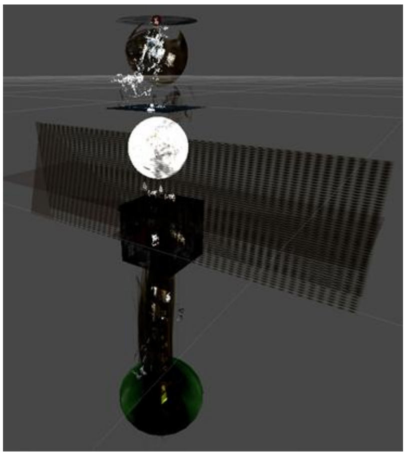


Image 2. A distant view of the virtual space as a vertical structure. Copyright © Nikos Papadopoulos 2016.

Navigation in this space simulates a perpetual fall, a voice controlled rise, and a horizontal rowing navigation across the different levels, in first person. These functions may evolve in parallel or sequentially, based on participant’s will.

Specifically, the perpetual fall in the vertical axis is based on the gravity simulation that the game engine provides, while the participant’s voice loudness captured by the microphone (with the contribution of other sounds from the physical or virtual space) balances that fall by boosting upwards. The voice is a tool of transferring and creating emotions, while the sound of each voice is unique. An experienced performer is able to handle this tool effectively, by speaking, singing, reciting poetry, laughing, or screaming, thus enhancing the collective experience and participation.

By using a paddle (physical object), and the basic technique used for navigating to the water, the visitor or the performer, can cross a world level horizontally and turn left or right. There is no boat; the participant’s body plays this role.

3.4. The selection of interaction mechanisms and their adaptation in the physical space

The virtual reality system was developed using the Unity game engine. The Oculus Rift headset was used for personal immersion. To control microphone input, through Unity, the MicControl 2 compatible software was used (see Unity Asset Store).

In order to simulate the rowing function through a physical paddle, a mechanism should be selected capable of precisely capturing the position and direction of the two hands (and thus the virtual axis connecting them) in physical space (6 DoF), while allowing the projection of this movement in the three dimensional space of Unity. It is noted that the installation took place in the basement garage of the Michael Cacoyannis Foundation (www.mcf.gr), with minimal lighting (Image 3), and therefore, the hands tracking mechanism should be capable of functioning without lighting and line-of-sight requirements. For this reason, the Leap Motion and Microsoft Kinect solutions were rejected, as they require line-of-sight, which makes them sensitive to low light conditions. Leap Motion was also rejected due to limited detection range (25 to 600 millimeters above the device), which is not sufficient for simulating rowing movements.

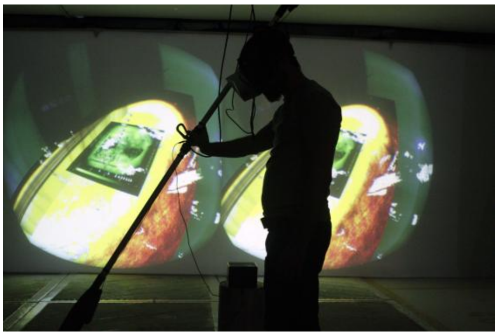


Image 3. The installation in the physical space. Copyright © Nikos Papadopoulos 2016.

Finally, the Razer Hydra gaming controller was selected, which provides among others two cabled 6DoF hand tracking controllers. This mechanism is mostly suitable for gaming in a seated position, in front of a desktop. It does not require line-of-sight to function and it provides an optimum detection range of 2-3 feet around the base station.

In our case, the 2 hand tracking controllers of Razer Hydra were adapted onto the physical paddle (Figure 1), in such a distance (d1) that they are inside the coverage area. The participant who is holding the paddle, does not feel the technological medium, but uses only body energy to produce movement in the virtual space.

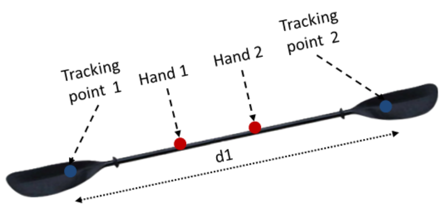


Figure 1. The adaptation of the technological mechanism to the physical paddle.

In fact, what is tracked by the Razer Hydra mechanism is the movement of the two points on the virtual axis connecting the hands (tracking point 1 and tracking point 2). A parametric virtual representation of the hands and the paddle exist in the 3D space, for calibrating and configuring the system, but they are not visible during navigation.

In the installation space, the physical paddle and the Oculus Rift are hanged from rope, so that the movements of the participant are limited in the scope of the tracking range (Figure 2).

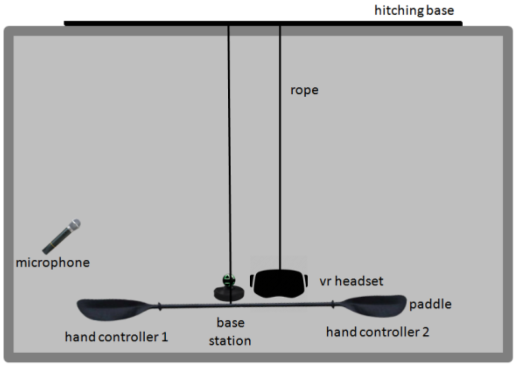


Figure 2. Installation in real space.

The visitor or the performer, who is standing and “wearing” the technological equipment (paddle, Oculus Rift, microphone), are using only their feet to support the body against gravity, while asked to adapt their skills to an unfamiliar space.

4. Simulation of multimodal navigation

Multimodal navigation involves a) the simulation of individual modes and b) their integration under a common architecture which determines the range of different combinations during system operation. In the case of the “Rowing courses” project, the entity that integrates the individual modes of navigation is called *Motion Controller*.

The Motion Controller, developed within the Unity engine, is the main mechanism that drives the First Person Character Controller (https://docs.unity3d.com/) and coordinates three functions a) viewing (personal immersion through Oculus Rift or first person view of the performer), b) balance-rise against gravity through voice and c) rowing navigation. These functions may evolve in parallel or sequentially, depending on participant’s will and may vary according to the mode of system operation (personal immersion or collective experience). In the next sections, the simulation of individual modes of navigation are analyzed.

4.1. Control of viewing

There are two suitably configured virtual cameras defined in the hierarchy of the First Person Character Controller, and the mode of system operation determines which is active (only one can be active at a time).

In the collective mode (projection and performer), the common view is driven by the performer’s movements in virtual space (voice, rowing), which are translated, by the Motion Controller, to movements of the corresponding virtual camera of the First Person Character Controller. Looking around (corresponding to *Mouse Look* in the mouse-keyboard configuration) is not supported in this mode, for conceptual reasons. The movements of the performer and thus the creation of the common view are targeted.

In the personal immersion mode through the Oculus Rift headset, the corresponding virtual camera is activated, and the Motion Controller aligns the rowing movement direction with the direction of viewing (direction of the virtual camera), especially after looking around.

The transition from one mode to the other (switch between virtual cameras) is possible in real time; it is given as a command (e.g. through a button press or gesture) either by a system operator or by the performer, by sending an appropriate message, as will be explained below, in Section 3.4.

4.2. Balance and rise through voice

The gravity parameter of the First Person Character Controller, acquires variable value which is proportional to the voice *loudness* value detected by the microphone (from the visitor, the performer or the environment, as the performer can place the microphone close to sound sources in physical space). When the loudness parameter exceeds a certain threshold (*loudnessThreshold*), then the following function applies:

*customGravity = -1 * loudness * accelFactor;*

where *accelFactor* is a parameter that defines the speed of fall, and which, as will be shown later on (in Section 3.4), can be adjusted during system functioning, thus adapting the virtual experience.

4.3. Navigation through rowing

Rowing supported by a physical paddle, allows the following movements:

Movement 1: alternately paddle on the left and right side, from front to back in order to shift the virtual camera forward.

Movement 2: alternately paddle on the left and right side, from back to front in order to shift the virtual camera backward.

Movement 3: paddle continuously on the left side, from front to back, in order to turn the virtual camera right.

Movement 4: paddle continuously on the right side, from front to back, in order to turn the virtual camera left.

A virtual (sea) level is defined, as a cube collider placed under the virtual paddle level (defined when the paddle is at rest), at a height relevant to the physical paddle's length and position in virtual space (this is easily adjustable).

The velocity of the rowing shift movement and the rotation angle are parameterized (*moveSpeed*, *rotationSpeed* parameters correspondingly), as will be shown, in order to be configured analogously (even in real time).

The Motion Controller has to identify, initially, whether any part of the paddle is under the virtual (sea) level. This is achieved by using the Unity Raycasting mechanism against the cube collider of the virtual (sea) level. Two rays are defined, from the center to each of the two edges of the paddle (with the red and yellow colors), as illustrated in Figure 3.

Figure 3. Raycasting mechanism identifying the paddle colliding with the virtual (sea) level.

The identification of each of the above 4 movements is realized based on the concept of *rowing cycle*. For each of the two rays, a complete rowing cycle lasts as long this ray is returning true (metaphorically: from the moment this part of the paddle enters water until the moment it exits).

The paddle movement, during a whole rowing cycle, can be sampled based on time units. Let's assume we use Unity frames as a time unit and let's assume also that a rowing cycle is completed into k frames (Figure 4).

In each frame i (where for each rowing cycle i=1,…, k), if any of the paddle edges is under the virtual (sea) level:

- Raycast returns true and the current hit point is returned (*current_hit_point*). The hit point of the previous frame is also stored (*previous_hit_point*).
- the delta z position value (dzi) of the paddle edge under water is calculated as:

dzi = current_ hit_point.z – previous_ hit_ point.z

The virtual camera is shifted on z axis, -dzi* *moveSpeed* units per second, and turned around y axis, *rotationSpeed* euler angles per second, right or left depending on the edge of the paddle under water.

Throughout a whole paddling cycle, which corresponds to k Unity frames, the sum of dzi values, multiplied by *moveSpeed* factor determines the whole shift of the virtual camera, forward (if the sum is negative) or backward (if the sum is positive).

$$\textit{total shift} = \textit{moveSpeed} * \sum_{i=1}^{k} dzi$$

The sum of rotationSpeed values determines the angles of rotation of the virtual camera around the Y axis (to the right if left paddle edge is under water, to the left if right paddle edge is under water).

$$\textit{total angles rotation} = \sum_{i=1}^{k} rotationSpeed.$$

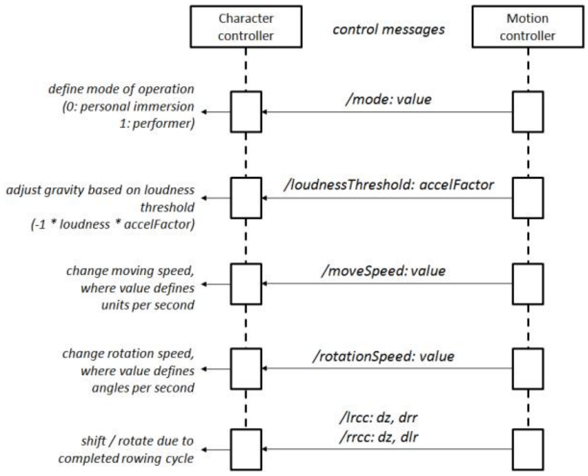
Figure 4. A paddling cycle and the resulting dz.

In this way, different combinations of alternate or continuous paddling on the left and right paddle sides result in each of the 4 movements described above.

5. System architecture

The Motion Controller mechanism operates autonomously, in relation to the control mechanisms of First Person Character Controller (e.g. physics, movements) and the special characteristics of the virtual space (e.g. intervention of complex colliders). In terms of Unity, the two entities run through different executables. The communication between the two entities (Motion Controller, First Person Character Controller) is possible through a messaging mechanism (Figure 5).

This allows the distribution of operations, as well as the real-time (and even remote) control of the system. Moreover, the same Motion Controller is more generic and therefore easily applicable to any virtual world.



lrcc: left rowing cycle completed
rrcc: right rowing cycle completed
dz: delta shift in z axis, due to rowing cycle completed
drr: delta rotation right
drl: delta rotation left

Figure 5. The communication between the Motion Controller and the First Person Character Controller.

The messaging mechanism was implemented based on a client-server architecture and the Open Sound Control protocol (OSC) (http://opensoundcontrol.org/). The OSC protocol allows for the communication between different, homogeneous or heterogeneous sources of data (e.g. multimedia devices, processes), thus enabling open architectures that support complex spatio-temporalities. Furthermore, it enables the definition of different message patterns based on the semantics of the information to be sent, and the sending of lists of values, as illustrated in the examined case in Figure 5.

The buttons of the Razer Hydra controllers were used for triggering, through the OSC protocol, the different real-time configuration commands. In this way, the performer was able to animate a variety of experiences, based on spatial expressiveness and the participation of the audience.

The MIT Licensed *UnityOSC* component (https://github.com/jorgegarcia/UnityOSC/) was used for enabling OSC messaging through Unity.

6. Conclusions and further work

This study focused on the development of multimodal virtual experiences using low-cost games mechanisms, under the prism of artistic creation. The main hypothesis used is that a virtual reality artistic system is a complex technological structure that when suitably parameterized, and organized under a scalable architecture is capable of unfolding hidden aspects that can be examined as artistic values (e.g. aesthetics, participation, complexity that leads to the unpredictable). Games mechanisms, when incorporated in such a structure, may constitute useful tools towards investigating new experiences.

References

[BRC96] BURDEA, G., RICHARD, P., COIFFET, P., Multimodal virtual reality: Input-output devices, system integration, and human factors. Published in International Journal of Human-Computer Interaction - Special issue on human-virtual environment interaction archive, Volume 8 Issue 1, Jan.-March 1996, Pages 5 – 24.

[Dav95] DAVIES, C., Osmose, 1995, http://www.immersence.com/osmose/

[Eri95] ERICKSON, T.D.: Working with interface metaphors. Morgan Kaufmann Publishers Inc. 1995.

[Man01] MANOVICH, L., The language of new media. The MIT Press, 2001.

[Mil12] MILK, C., The Treachery of Sanctuary, 2012 http://milk.co/treachery.

[Par14] PARK, L., Eunoia II”, 2014, http://www.thelisapark.com/#/eunoia/.

[Pau03] PAUL, C.: Digital Art. Thames & Hudson. 2003.

[San06] SANTORINEOS, M., Gaming Realities. Festival MEDIATERRA, Fournos Center for Digital Culture, 2006, Athens.

[San14] SANTORINEOS et al., Επιλογή In Crisis, 2014, https://www.youtube.com/watch?v=ej7wHJ7k4Mk

[Shaw] SHAW, J., compendium web site, http://www.jeffreyshawcompendium.com/.

[SMTM15] SONG, M., MOKHOV, S. A., THOMAS, J., MUDUR, S.P.: Applications of the Illimitable Space System in the context of media technology and on-stage performance: A collaborative interdisciplinary experience. In the Proceedings of the IEEE Entertainment Media Conference (GEM), 2015, Toronto.

[SV14] STERCKX, F., VERBEECK, A., The Styx, 2014, http://skullmapping.com/, 2014.