Node Kara: An Audiovisual Mixed Reality Installation

Anıl Çamcı* and Angus Graeme Forbes†
Electronic Visualization Laboratory
University of Illinois at Chicago

ABSTRACT

Node Kara is an audiovisual mixed reality installation by the artist Anıl Çamcı. It offers a body-based interaction using a 3D imaging of the exhibition space. With a blurred audiovisual scene in its steady-state, it invites visitors to interact with the multiplicity of layers it holds within. By moving in the exhibition space, the visitors can deblur both the sound and the visuals as they get immersed in the piece. Rather than a detection of discrete events or gestures, *Node Kara* relies on fluid interactions with continuous visual layers and stochastic audio. This way, the piece reacts to each visitor in a unique way. A video preview of *Node Kara* can be seen at **https://vimeo.com/anilcamci/nodekara** (password: nodekara).

Index Terms: H.5.1 [Information interfaces and presentation (e.g., HCI)]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Auditory (non-speech) feedback.

1 INTRODUCTION

Node and *kara* are two words in Japanese that indicate causality. While *node* is used to describe natural cause-and-effect relationships, *kara* is used when a causality is interpreted subjectively. Although a clear causal relationship between actions and reactions is a staple of interaction design, the user's subjective experience of such relationships often trumps the designer's predictions. By adopting *blurring* both as a theme, and a technique, *Node Kara* obfuscates the causal link between an interactive artwork and its audience. The deblurring of the audiovisual scene becomes an attracting force that invites the viewer to unravel the underlying clarity of *Node Kara*. This comes, however, at the cost of losing a broader perspective of the work as the viewer needs to come closer to both see and hear the work in greater detail.

Figure 1: Wide depth-of-field image of the scenery used in *Node Kara*

*e-mail: anilcamci@gmail.com; web: http://anilcamci.com
†e-mail: aforbes@uic.edu; web: http://creativecoding.evl.uic.edu

Node Kara offers a multimodal mixed reality experience through the use of 3d body-based interactions. This is achieved by depth tracking with a Microsoft Kinect sensor. Using Cornock and Edmonds' seminal terminology [3], *Node Kara* can be characterized as a dynamic-interactive art system, which, while similar to kinetic artworks, relies on audience participation for the artwork to become activated.

In its steady-state, *Node Kara* imparts the least amount of information about the experience it encompasses. The blurred image and the low-rumbling noise act as a black box that invites the visitors to explore what is inside. In some sense, the piece is brighter and produces the frequencies that travel the furthest when it is expecting visitors to interact with it. As the visitor comes closer, a wider depth of field is revealed in the image. At the same time, the oscillation sources used by the stochastic audio engine evolve in a way that creates more partial frequencies. The completely deblurred audiovisual experience can be viewed at the end of the video link provided in the abstract. This, however, is merely an ideal that cannot be fully attained even at the farthest levels of interaction. This implies that there will always be more for the visitor to unravel both visually and sonically.

1.1 Related Work

A great number of artworks have so far used computer vision and depth cameras to implement user interaction. One of the earliest works to utilize body-based virtual interaction is the artist Myrion Krueger's *Videoplace* [6]. Considered as a pioneering implementation of artificial reality, this work relies on the use of real-time computer vision to establish an interaction between users and their visual representations in video.

The artist Alessandro Kassinelli's interactive installation *Khronos Projector* allows the visitors to move parts of a video back and forward in time by touching a projection screen. This way the visitor deform not only the physical material that is being projected onto but also the spatiotemporal continuity of the visual content that is being projected [1].

In her interactive installation *Untitled 5*, Camille Utterback uses overhead cameras to track visitors' position in the exhibition space. This data is then translated into fluid paintings projected onto walls facing the visitors. The artist describes this work as creating painting-like, organic and evocative imagery "while still being completely algorithmic" [10].

The artists Christa Sommerer and Laurent Mignonneau's 2015 work, *Portrait on the Fly*, creates interactive portraits of its visitors in the form of silhouettes contoured by fly-like visual units through the use of computer vision. The artists describe their inspiration as the composite heads created by Guiseppe Arcimboldo in the mid 15th century [9]. A similar computer vision technique was used by Javier Villegas in his 2012 installation *Ant Theater* [11, 12], where visitors' contours were drawn in real-time by a colony of ants on an image of soil.

In his project *Falling Girl*, the artist Scott Snibbe injects silhouettes of viewers of the artwork into an existing visual narrative that recounts the story of a young girl falling from a skyscraper [8]. Based on a similar technique, in their interactive installation titled *Ghost* [4], the artists Thomas Eberwein and Tim Gfrerer, capture

skeletal images of visitors using a Kinect sensor to render an interactive image of a snow storm surrounding the virtual representations of the participants.

In a recent project, titled *Diving into Infinity*, the artists Bill Manaris, David Johnson, and Mallory Rourk use Kinect-based depth body-tracking to transform a static painting by M.C. Escher's into "a dynamic-interactive piece of art" [7] through the interactive distortion of images in a way that reveals self-similarities in it.

Finally, in a 2015 implementation of their mixed reality installation series *Artificial Nature*, the artists Haru Ji and Graham Wakefield use a Kinect sensor to integrate a depth image of the participant within a complex generative system presented in the form of a virtual reality environment. The artists describe that the visitors become part of the biologically inspired ecosystem through their embodied interactions with the piece [5].

2 Image

On Febraury 20, 2015, four different high resolution videos of a scenery in Istanbul, Turkey were captured with minute differences in focal length. For filming, a Canon 5D Mark III DSLR camera with a Canon EF 24-105 f/4L IS Lens was used. The selected scenery consists of a great variety of both dynamic and static layers. A wide depth-of-field image of the scene can be viewed in Fig 1. This image represents the deepest level the visitors can reach through interaction.

A virtual representation of the visitor is cast in the scene as a lens that articulates various depth layers of the visuals as seen in Fig 2. The real-time visual processing is achieved using a software developed with the C++ multimedia toolkit *openFrameworks*[1]. A custom OpenGL fragment shader utilizes the depth image from a Kinect sensor as a mask to cast various layers of video on top of the steady-state image based on the visitor's proximity to the sensor, which is positioned beneath the projection screen. This effectively creates the illusion of the work deblurring gradually in the visitor's image.

Figure 2: Four layers of *Node Kara* superimposed onto each other with a representation of the visitor's depth image highlighted

As the visitor interacts with the piece, multiple layers of the same scene blends together. This enables the visitor not only to explore the scene at different focal lengths but also to shift focus from nearby figures to distant ones. Since each video was shot from the same viewpoint, they were filmed consecutively and not simultaneously. The resulting time shifting between the layers adds a temporal depth to the work through the interplay between different shots of the same dynamic elements at different times. This creates an uncanny motion blur-like effect that adds to the fluid aesthetics of the work.

[1] http://www.openframeworks.cc

3 Sound

Sounds of *Node Kara* are generated in real-time using a custom stochastic synthesis engine titled *Temas* (seen in Fig 3), which was designed by Anıl Çamcı in 2014 [2]. Using such a system for the audio component of the piece assures that, while there is a clearly recognizable relationship between the visitor's movements and the their sonic output, no one sound is repeated during the course of the installation. Rather than using the visitor's actions to trigger sound events, the stochastic synthesis evolves gradually based on the visitor's interactions. This way, *Node Kara* does not react in terms of discrete events but rather in the form of a sonic continuum. The piece is therefore experienced less like a musical instrument and more like a fluid and immersive audiovisual experience.

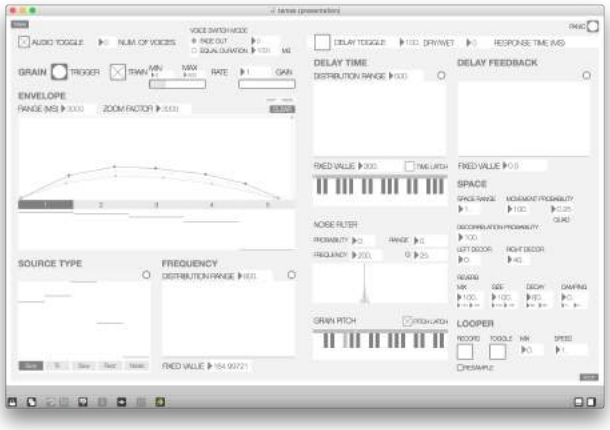

Figure 3: Anıl Çamcı's stochastic audio synthesis engine *Temas*, which was used for the interactive sound in *Node Kara*

The visitor's movement in three dimensional space is mapped to various parameters of the stochastic synthesis. Once an agent in the depth image is recognized, a fundamental frequency is assigned to it. As the visitor moves on the horizontal axis parallel to the projection screen, the probability curve that determines the partials around this fundamental frequency sweeps the frequency spectrum according to the visitor's position. These mappings are intended to create an aural counterpart of the continuous interactions in the visual domain. The amplitude envelopes of the sounds are accordingly flat to create long ambient gestures rather than sparse events.

When the visitor moves closer to the projection screen, the distribution that determines the synthesis source evolves from sine waves to square waves. As a wave with a fixed maximum amplitude morphs from a sine wave to a square wave, its root mean square amplitude (RMS) and the number of audible partials increase. In *Node Kara*, this creates the effect of gradually removing a high-pass filter from the audio, which is analogous to the deblurring the same action creates in the visual domain. Furthermore, the area the visitor takes up on the screen is proportional to the width of the probability curve that determines the overtones added to the fundamental frequency.

In terms of the spatialization of the sounds, the panoramic positioning of the sound output from each synthesis unit is linked to the position of an agent detected in the depth image. This ensures that a sound resulting from a particular visitor's interaction with the system will also follow her spatially. Since each agent is assigned a unique sound generator, this sound spatialization applies to all sounds regardless of the number of visitors.

4 IMMERSION

Although the visual component of *Node Kara* is displayed on a 2D display, the visitor's 3D interaction with the visuals, and the addition of interactive sound amount to an immersive experience. A three dimensional depth image of the exhibition space informs *Node Kara* of its visitors. The more a visitor gets immersed in the work, the more layers of the scene become visible.

Figure 4: A visitor interacting with *Node Kara*

By their acoustic nature, low frequency sounds can travel long distances and are hard to localize. The omnipresent quality of such sounds in the steady-state of *Node Kara* are intended to lure the audiences to be immersed in the work. As with the visuals, when a larger area of the tracked space is interacted with, more components of the sonic structure are also revealed. Yet again, as a visitor uncovers a sound element more extensively by getting immersed in the piece, she surrenders her contextual awareness at the same time.

Multiple visitors can collaborate to reveal a larger portion of the audiovisual scene at a given time. An example of this with three visitors can be seen in Fig 5.

Figure 5: Two visitors collaborating to reveal the left edge of the piece while a third visitor towards the right is exploring the piece individually

5 IMPLEMENTATION

The audiovisual system underlying *Node Kara* offers scalability in terms of the implementation of the artwork. In the least extensive setup, the piece can be presented with a full-wall projection, a set of stereo speakers adjacent to each side of the video, and a single depth camera placed beneath the projection field. If the low-frequency driver of the selected speakers utilize a cone smaller than 8", a separate subwoofer becomes necessary for the reproduction of the low frequency rumbles heard in the steady state of the work.

In terms of audio, the system can be scaled up to an arbitrary number of speakers placed on a horizontal array located behind a translucent projection screen. A four-speaker version of such an implementation can be seen in Fig 6. Furthermore, since the visual content is based on regional deblurring of an essentially fog-like texture, the high resolution videos can be repeated horizontally. As the lateral extent of the installation increase, more depth sensors to cover the navigable area of the exhibition space will be required.

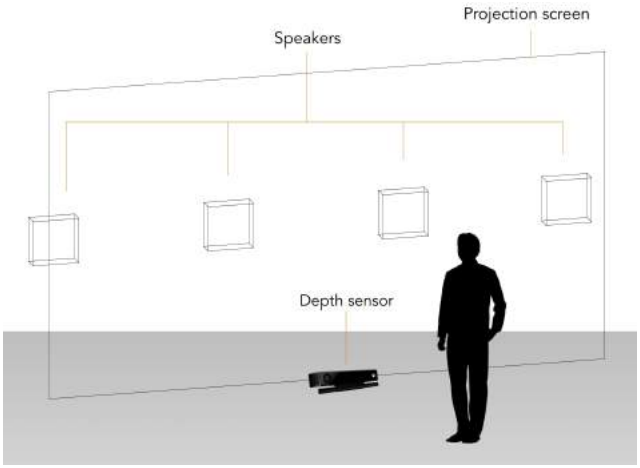

Figure 6: An implementation scenario for *Node Kara* including a four-speaker sound system

In an optimal implementation of *Node Kara*, as seen in Fig 6, a 18(W) by 10(H) translucent screen is placed in front of 4 5 speakers. This setup is supported by a subwoofer that can be placed arbitrarily in the exhibition space. Live visuals are projected onto the screen, while 4-channel spatial audio is played-back through the speakers. The Kinect sensor feeds depth images to a single computer, which processes this data to manipulate the audio and the visuals in real-time.

6 CONCLUSION

The use of mixed reality in artworks offers distinct possibilities. In such works, the physical presence of the audience is augmented with audiovisual artifacts. These augmentations can be used to heighten a visitor's awareness of her physical self, or to completely transport this self into a separate reality. In *Node Kara*, the visitor's presence becomes a tool of interaction that uncovers (i.e. deblurs) the internal narrative of the piece both visually and sonically.

This paper described the theory and implementation of *Node Kara*. As a mixed reality installation, this work relies on body-based audience interaction, where a visitor's motion in 3D space acts as an interface to activate the artwork. The main pillars of *Node Kara* were described as images (i.e. live-processed video content), sounds (i.e. stochastically synthesized audio), and immersion (interaction through depth sensing). Possible implementation scenarios were described for a variety of setups ranging from single-wall projection with stereo speakers to large scale installations with multiple projectors and multichannel speaker arrays.

ACKNOWLEDGEMENTS

Early studies in this project were conducted at the Istanbul Technical University, Center for Advanced Studies in Music, where Anıl Çamcı acted as the founding coordinator of the Sonic Arts program.

REFERENCES

[1] A. Cassinelli and M. Ishikawa. Khronos projector. In *ACM SIGGRAPH*, volume 5, 2005.

[2] A. Çamcı. *The Cognitive Continuum of Electronic Music*. PhD Dissertation. Leiden University, 2014.

[3] S. Cornock and E. Edmonds. The creative process where the artist is amplified or superseded by the computer. *Leonardo*, pages 11–16, 1973.

[4] T. Eberwein. Ghost. Online, 2013. Retrieved January 15, 2016 from http://thomastraum.com/#works/ghost/.

[5] H. Ji and G. Wakefield. Endogenous biologically inspired art of complex systems. *IEEE Computer Graphics and Applications*, 36(1):16–21, January/February 2016.

[6] M. W. Krueger. *Artificial reality II*. Addison-Wesley Professional, 1991.

[7] B. Manaris, D. Johnson, and M. Rourk. Diving into infinity: A motion-based, immersive interface for mc eschers works. In *Proc. of the 21st International Symposium on Electronic Art (ISEA 2015)*.

[8] S. Snibbe. Falling girl. Online, 2008. Retrieved January 15, 2016 from http://www.snibbe.com/digital-art#/projects/interactive/fallinggirl/.

[9] C. Sommerer and L. Mignonneau. Portrait on the fly. Online, 2015. Retrieved January 15, 2016 from http://www.aec.at/center/en/2015/09/07/portrait-on-the-fly/.

[10] C. Utterback. *Artist Statement*. ACM SIGGRAPH Art Gallery, 2005.

[11] J. Villegas and A. G. Forbes. Double-meaning: Interactive animations with simultaneous global and local narrative. *Proceedings of the Renew Digital Arts Festival*, pages 302–306, 2013.

[12] J. Villegas and A. G. Forbes. Analysis/synthesis approaches for creatively processing video signals. In *Proceedings of the ACM International Conference on Multimedia (MM)*, pages 37–46, Orlando, Florida, November 2014.