KIMA: The Wheel – Voice turned into Vision

A participatory, immersive visual soundscape installation

Oliver Gingrich (Visiting Research Fellow, National Centre for Computer Animation), Sean Soraghan (Sound Developer), Alain Renaud (Visiting Research Fellow, Centre for Digital Entertainment), Evgenia Emets (Artist), Dario Villanueva-Ablanedo (Visual Programmer)

analema.group@gmail.com

Abstract

Over the last five years, KIMA, an art and research project on sound and vision, has investigated visual properties of sound. Previous iterations of KIMA focused on digital representations of cymatics - physical sound patterns - as medium for performance. The current development incorporates neural networks and machine learning strategies to explore visual expressions of sound in participatory music creation. The project, displayed on a 360 degree canvas at the London Roundhouse, prompted the audience to explore their own voice as intelligent, real-time visual representation. Machine learning algorithms played a key role in meaningful interpretation of sound as visual form. The resulting immersive performance turned the audience into co-creators of the piece.

Fig. 1. Performance of KIMA - The Wheel by Evgenia Emets (© Analema Group 2016)

KIMA – The Wheel

Six microphones suspended in a round space, faced a circular projection piece by Ron Arad. Whenever a sound was made, the 360 degree projection came to life and soundwaves started to travel across the circular canvas (Fig.1). KIMA - The Wheel is based on the concept of six Buddhist realms. The audience explored their voice across these six different worlds, each representing the human voice in different ways. Performers guided the audience. Visitors became part of the audio-visual composition. In KIMA – The Wheel, nuances and differences in the human voice were represented in a multitude of ways, a constantly shifting flow of visual currents, of waves travelling across the 360 degree canvas (Fig. 2). In the middle of the space, two microphones invited the public to experience this wave in its entirety, immersing the audience within – visually and sonically. KIMA is the outcome of a five year-long investigation into visual sound by the collective Analema Group.

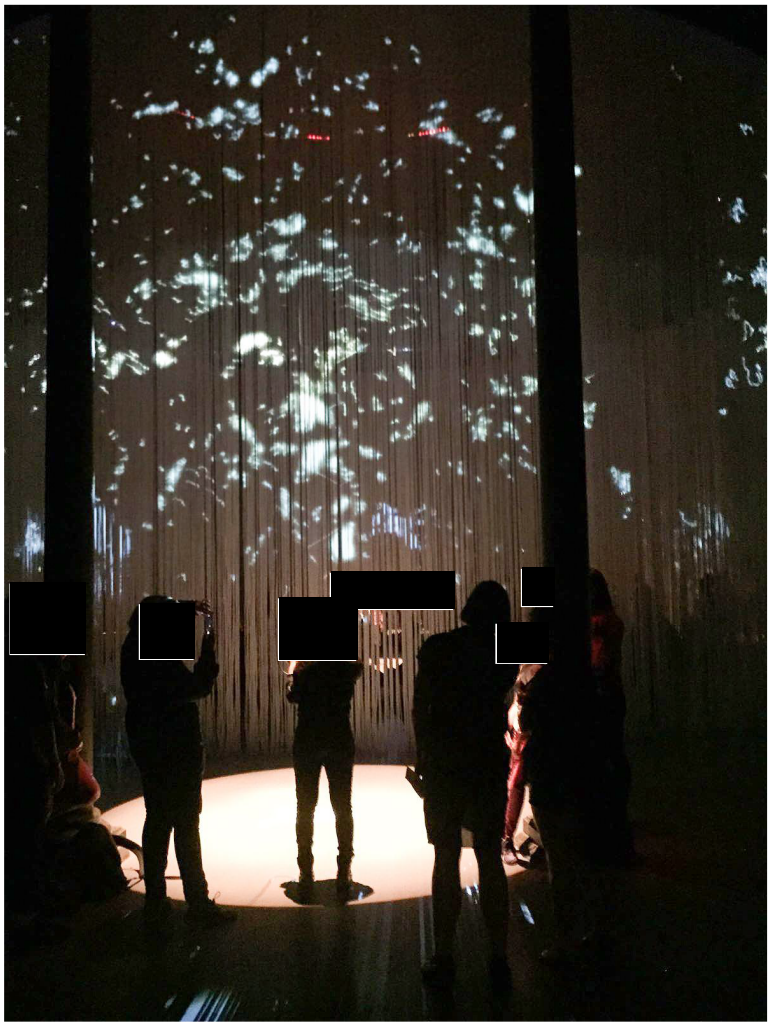

Fig. 2. Audience Participation in KIMA - The Wheel. (© Analema Group 2016)

KIMA - History of a Collaborative Art Piece

Over the last five years, Analema Group created transformative, multi-sensory experiences with a focus on the audience's perception. KIMA was previously shown at Union Chapel London (Fig.3) , Royal College of Music, Kinetica Museum (Fig. 4) and others. From smaller scale installations -representations of physical characteristics of sound waves- KIMA evolved into large-scale immersive environments. The focus on the human voice stems from an interest to empower the public. Beyond dichotomies between artist and audience, KIMA dissolves artificial boundaries of stage and screen. The voice, as medium of emotional expression, acts as instrument accessible to all of us. KIMA provides an interface for the playful re-discovery of the voice. The project itself is routed in a historic discourse on audio-visual art and performance.

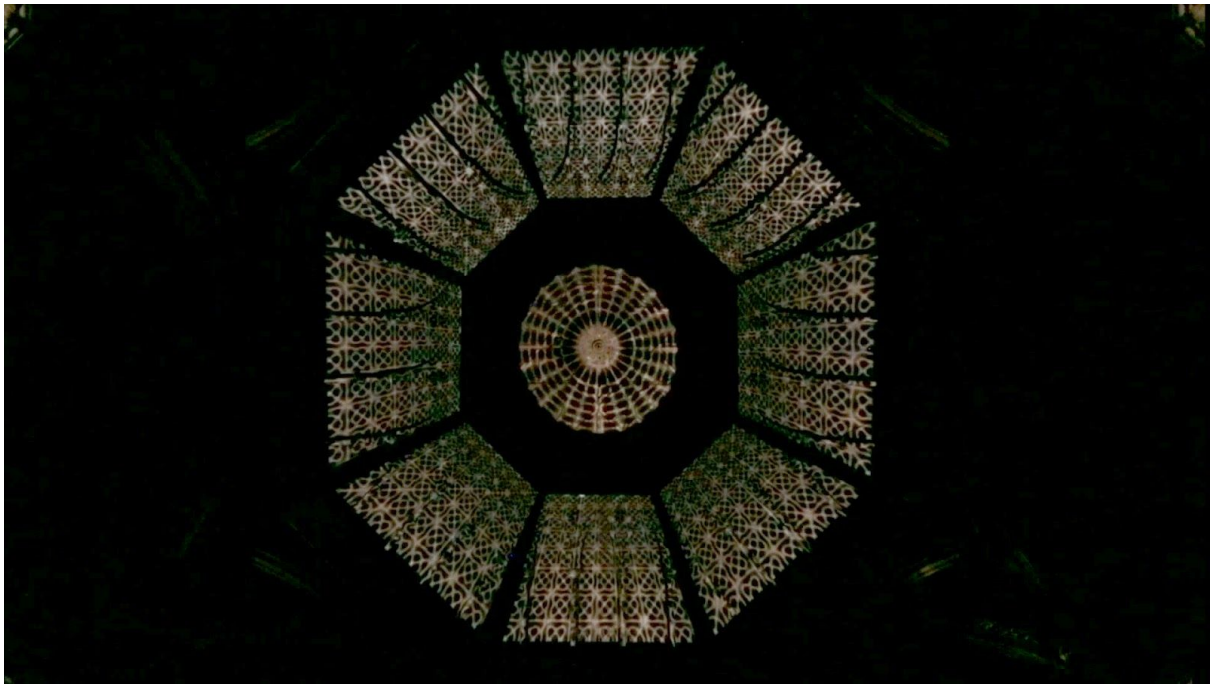

Fig. 3. KIMA - Union Chapel (© Analema Group 2015)

KIMA – Sound and Vision

The history of visual sound can be traced back to the Stone Age [i]. Yet KIMA is inspired by an on-going quest in the 20th and 21st century to combine the realms of sound and vision: From the early 1900s, artists created audiovisual instruments, such as the Futurist Luigi Russolo's Intonarumori[ii]. In the 1950s and 1960s, Wassily Kandinsky pioneered visual expressions for music on canvas. Conversely, the Russian composer Scriabin investigated links between colour and

sound musically, using small lights attached to a piano. With the advent of computer animation and electronic arts, new artistic possibilities to correlate sound and vision emerged. During the Fluxus movement, Nam June Paik coined the term 'cracked media'[iii], to describe art beyond the limitations of genre. In the 1960s, the Whitney's animations amalgamated sound and vision into a "Gesamtkunstwerk" - a mix of different art forms[iv]. In the Seventies, Stan VanderBeek and Neil Youngblut[v] experimented with light and sound to create new, immersive audiovisual experiences. With important retrospectives at the Whitney ('Dreamlands Seminars'[vi]) and the Whitechapel Gallery ('Electronic Superhighway'[vii]), hybrid art forms are encountering a renaissance. Whereas in art genre specific traditions, sound tends to dominate over vision or vice versa, KIMA chooses a third path - a synergy between the two, with one organically emerging from the other.

Fig. 4. Evgenia Emets,Lani Rocillo performing KIMA at Kinetica 2014 (© Analema Group 2014)

KIMA – From Cymatic to Semantic Sound

A mathematical, rule based system connects sound and vision on a physical level. In the 1860s, the ”father of acoustics”, Ernst Chladni, monitored this behaviour as representations of sound on rigid surfaces – so called cymatics [viii]. The Swiss scientist Hans Jenny continued these experiments in the 1960s [ix] and created a systematic catalogue of waveform functions - patterns generated by sound on matter. Inspired by this physical connection between sound and vision, our initial artistic development looked at cymatic waveform equations. Research points to an existing connection between language and phonetics - a cross-cultural, gender and age independent, descriptive reading of sound. Furthermore, a large body of evidence suggests that many of the descriptors used for timbre are visual, textural, and material in nature [x,xi,xii]. For Analema Group, language serves as inspiration for a visual sound interpretation. A purpose-built ‘Feature Extractor’ facilitates the analysis of different aspects of sound. KIMA then links these sound features to meaningful visual expressions through machine learning algorithms. The development is guided by principles of intuitive visual interpretation. The audience explores their voice in visually distinctive, always intuitive ways - resulting in a low barrier of entry to audience participation.

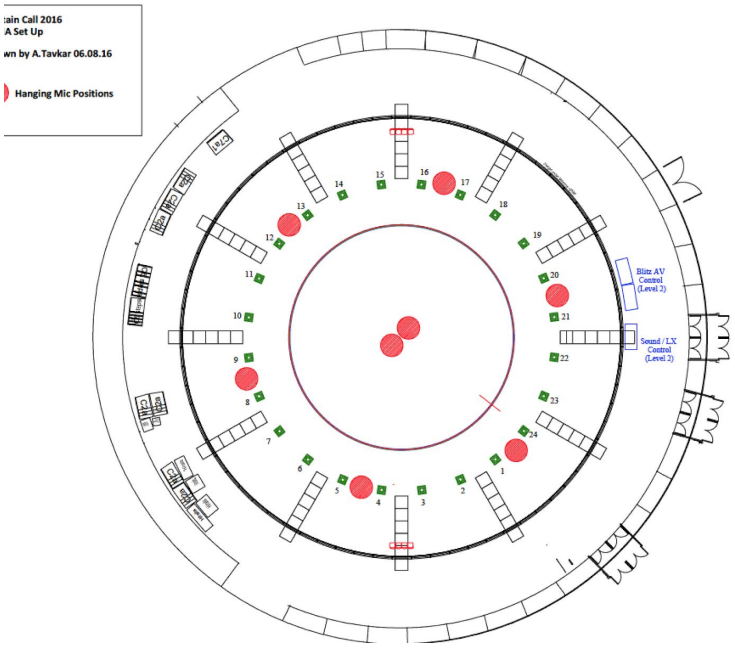

Fig. 5. Technical Layout of KIMA: The Wheel at the London Roundhouse (© Analema Group 2014)

KIMA – a spatial, visual sound environment

"KIMA: The Wheel" was presented on Ron Arad's epochal Curtain Call in August 2016 for the 50th jubilee of the London Roundhouse. The 360 degree installation measured eighteen meters in diameter and eight meters in height, surrounding the audience entirely. The "Curtain Call" consisted of 5,600 silicon rods, acting as permeable projection environment. The audience walked through the Curtain, appreciating the projection from both sides (Fig. 5). Visual immersion was complemented by ambisonic surround sound. Alain Renaud's bespoke design distributed the audience's sound in the round as continuous sound wave, circling through the iconic architecture of the Roundhouse. The result - a total multisensory experience - turned the audience into performer and recipient of the piece.

A poem written by artist Evgenia Emets provided conceptual structure to the artwork. Six microphones, situated on the outside of the Curtain, captured poetry as well as non-verbal sounds. The human voice was transformed into six unique visual abstractions relating to six conceptual zones. A choir of six people guided participants on this artistic journey. At the centre of the installation, two microphones invited the audience to further interaction, facing a constantly changing diorama of sound. Unlike Robert Barker's original panorama, where the canvas was static[xiii], the scenic spectacle of KIMA was ethereal, fleeting and dynamic - created by its spectators.

KIMA – The Wheel: Turning Sound into Vision

The inherent relationship between sound and its visual properties lies at the heart of the project. KIMA - The Wheel pushed the envelope in presenting more complex sound characteristics such as timbre in an abstracted and intelligent ways. According to Richard Bellmann[xiv], pioneer in dynamic programming theory, learning complexity grows exponentially with linear increase in dimensionality of data. For KIMA - The Wheel, a multitude of sound parameters were correlated to visual parameters of particle systems. Up to twelve sound feature traits were mapped onto ten particle system parameters, presenting users with a large range of audio-visual real-time possibilities. Sound features were extracted using the custom-built Feature-Extractor application (Fig. 6), developed using the JUCE cross-platform C++ library. The Wekinator architecture by Rebecca Fiebrink and Perry R. Cook[xv] linked sound features to visual parameters. The Wekinator, an open source machine learning software, provided a framework for iterative training,

using multiple learning algorithms. Wekinator assigned feature vectors of the Sound Feature Extractor to visual parameters using an Open Sound Control (OSC) protocol. Correlating sound and visual parameters through "training", the Wekinator facilitated a higher level of visual complexity. Machine learning thus assisted the creation of *mapping strategies* - from complex audio data to nuanced, versatile visual representations.

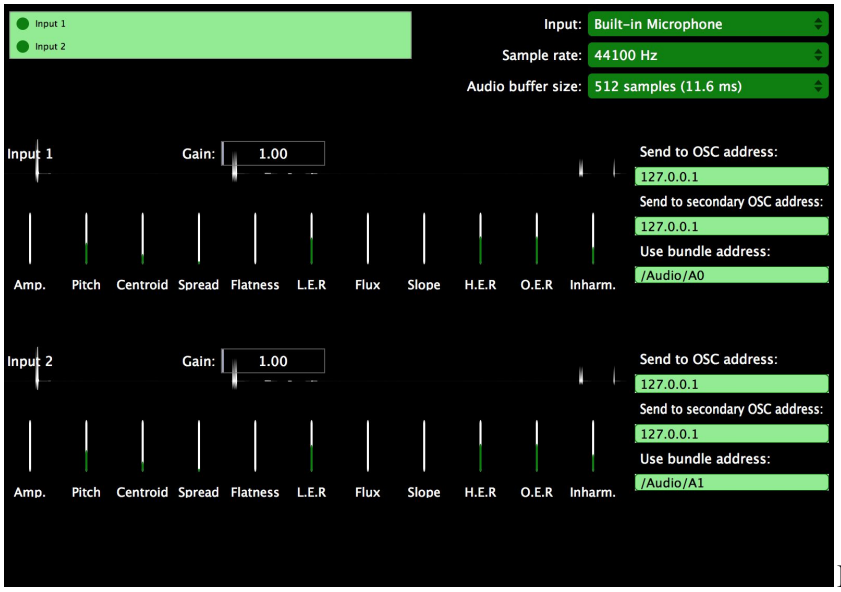

Fig. 6. Sound Feature Extractor (© Sean Soraghan 2016)

The Wheel: A Spatial Pattern Poem

The poem 'The Wheel' is based on six conceptually different, visual and phonetic "realms" - inspired by Buddhist mythology. According to the Buddhist concept of Samsara [xvi], beings travel through six realms during rebirth, experiencing a whole spectrum of emotions, both positive and negative. This idea of transitory realms built the starting point for three sets of performances for six singers, followed by audience participation. Each word was transcribed on a hexagonal piece of paper and positioned in hexagonal formations next to each microphone. Each of the six singers performed a given arrangement of the poetry sequence, while moving through the space. The poem served as initial score for both performers and audience and an inspiration for making sound with their own voices, embodying qualities of these "realms". Sound interpretations varied between arranged melodies and noise. The sound design distributed these soundscapes in the circular space. Live audience participation created an ever-changing soundscape, a continuous arrangement - at times harmonious and at times dissonant.

Fig. 7. KIMA: The Wheel. (© Analema Group 2016)

Each visual ‘realm’ used different representation strategies for audio data by “training” the machine learning software. A given sound, when produced in one of the six zones, was ascribed a different visual interpretation according to its context. The use of perceptual audio features to derive visual form from words, phrases and sounds as opposed to their existing linguistic and social meanings is central to the project. One of the key challenges, therefore, involved the characterisation of sounds in real-time - to translate their fundamental properties into intuitive visual representations (Fig. 7).

KIMA – Meaning of Sounds

The 'character' of a sound is predominantly defined by its timbre. Identification of timbre is often linked to the type of source that produced the sound. The description of timbre has therefore mostly been limited to instrumentation or its sound source. However, in the context of abstract sound sources (e.g. digitally produced sound) research has shown that we often turn to metaphors in order to describe timbre [xvii]. In such semantic timbre spaces', each dimension is labeled using a specific descriptor (e.g. 'bright', 'dense', 'hollow')[xviii,xix, xx, xxi]. A large proportion of

vocabulary describing timbre has a visual context, often using textural, physical or material terms [xxii]. KIMA explored metaphorical timbre descriptions in the context of the human voice. Separating physical sound characteristics of words and vocal utterances from their linguistic or social meaning, KIMA offers a visual interpretation of sound. A key requirement of such a representation paradigm is a mapping strategy from audio characteristics to meaningful visual features.

Research investigating acoustic descriptors of timbre has come from various fields including psychology, composition and performance, as well as music information retrieval (MIR)[xxii,xxiii,xxiv, xv, xvi]. Differences in disciplines lead to difficulties in collecting a global set of sound descriptors. Peeters et al. [xxvii] contributed towards this goal with their extensive list of descriptors. The feature extraction application developed by Sean Soraghan provides access to some of the spectral features in this list. Together, they give an instantaneous ‘snapshot’ of the frequency envelope and timbre of the audio at any one moment. By linking these features to control parameters of the Unity visualisation via Wekinator, each moment in time (frame) acted like a keyframe in an animation. This provided an audio-driven control interface to the generative visualisation algorithm. Studies investigating the alignment of semantic timbre features with perceptual timbre parameters have shown that semantic descriptors often correlate with multiple perceptual dimensions. [xxviii, xxix].

The Wekinator software facilitated the creation of arbitrarily complex mappings from audio features to meaningful visual properties. Through a process of iterative ‘training’, i.e. feeding the software with sound samples and linking them to visual parameters, we created the desired visual response. By repeating this training process for a collection of different audio inputs and visual outputs, we created a visual language for sound. As the regression algorithm interpolated between sounds to produce visual responses, resulting visuals morphed from sound to sound. The visualisation system facilitated the creation of our mapping strategy from visual responses to a specific audio input.

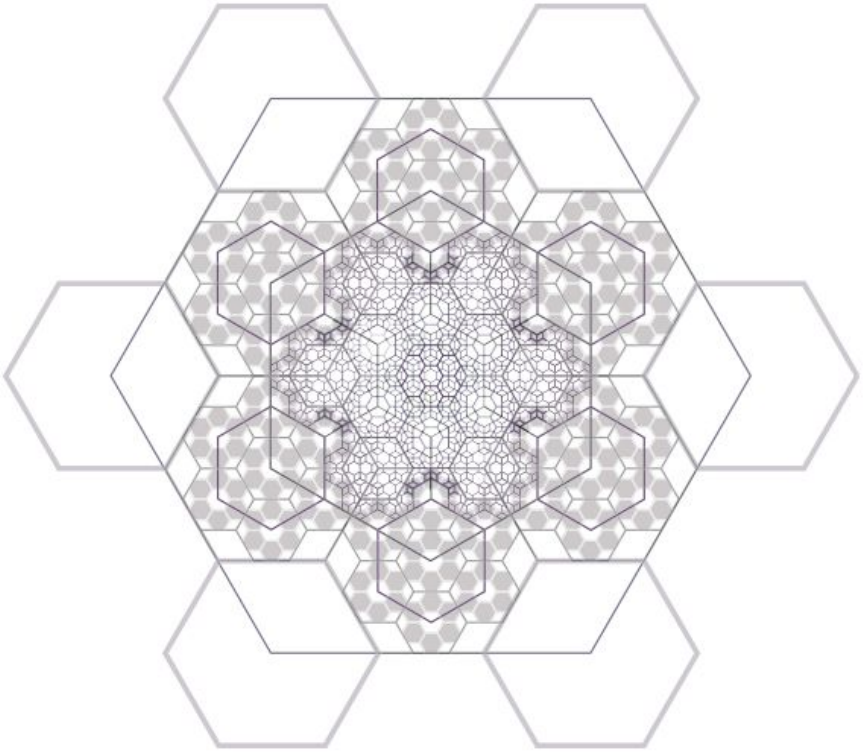

Fig. 8. Sketches for hexagonal forcefield of KIMA (© Analema Group 2016)

Hexagonal Meta-structures

The particle-system-based KIMA visualisation algorithm was developed by Dario Villanueva Ablanedo based on a creative brief by Analema Group. Six 'zones' were arranged on a hexagonal force grid lattice, at which particle system parameters took on different values. These values were assigned according to mapping strategies pertaining to their conceptual context. As visual sound representations travelled throughout the environment, they morphed from zone to zone. All six particle systems were created with a specific visual expression in mind, inspired by the realms referenced in Evgenia Emets' pattern poem. Whereas sound in the "animal" realm was represented as sprawling organism, the same sound resulted in a more ethereal, fluid representation in the "Demi-God-realm." A force field, controlled by central microphones, linked these zones together. Its lattice-grid parameters simulated a perlin-noise based vector grid field (Fig. 8). All six particle systems were connected through this force field as unified waveform representation. Its microcosmic hexagonal meta-structure was combined with macro-cosmic representation of six Buddhist realms. The audience thereby not only turned from spectator into performer, their voices became activator for a holistic, multi-sensory experience.

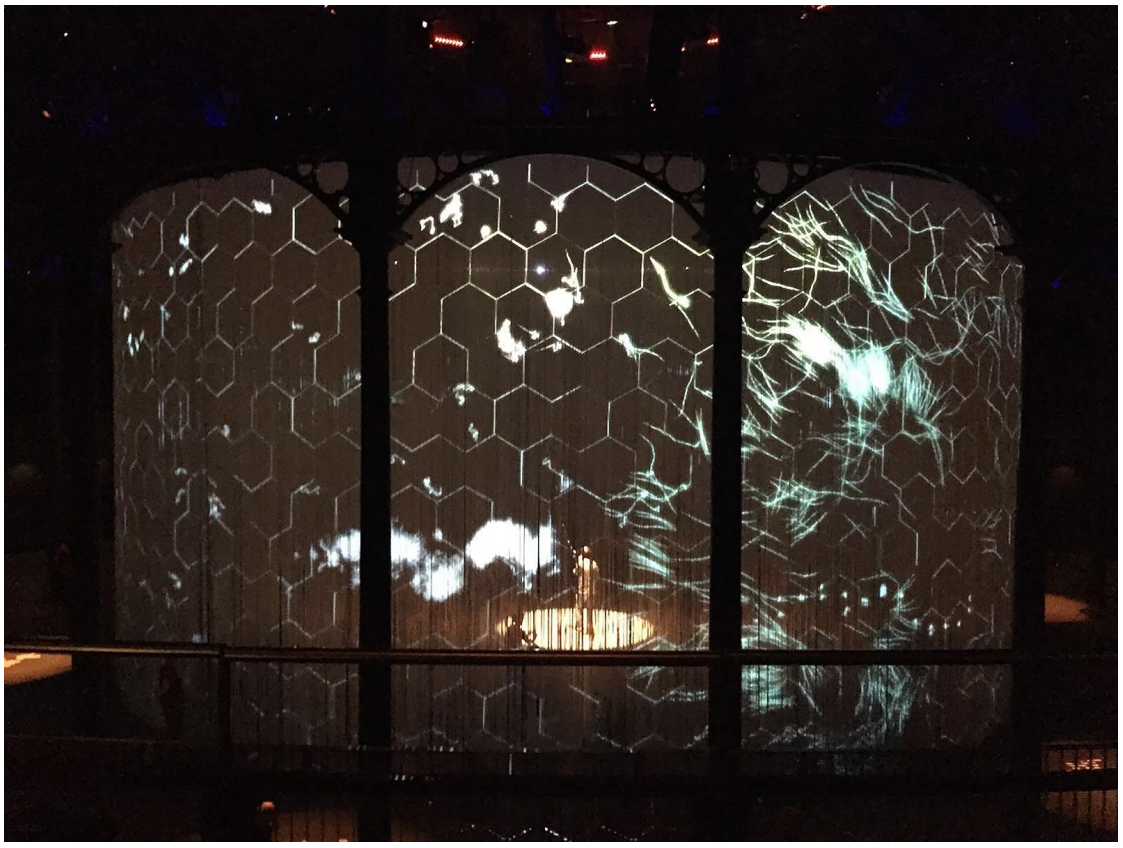

Fig. 9: Performer interacting with KIMA: The Wheel (© Analema Group 2016)

Real-time multi-channel Networking

Ron Arad's Curtain Call had never been staged using real-time generated, interactive visuals. A total of twelve projectors projected onto thousands of fibre rods suspended from a circular rig. Two 7th Sense Delta Infinity Media servers provided a total of twelve image outputs. Both servers were running the KIMA - Unity application developed by Dario Villanueva Ablanedo *(see: figure 11)*. Unity's real-time engine provided the simulated 3D-environment with real-time OSC-input. Servers were synchronised and controlled remotely, using a purpose-built controller Spout application *(see: Fig. 11)*. All audio streams were analysed by Sean Soraghan's cross-platform Feature Extractor on a remote machine *(see: Fig. 11)*. Audio features from each microphone were then sent to individual Wekinator instances, running in parallel on a second server*(see: Fig. 11)*. Sound input was spatially distributed depending on their azimuth value, which was shared with the Unity app, synchronising visual and sonic representation across the space. Audience perceived the resulting visual representation in front of each microphone, as well as in its entirety in the centre of the circular canvas.

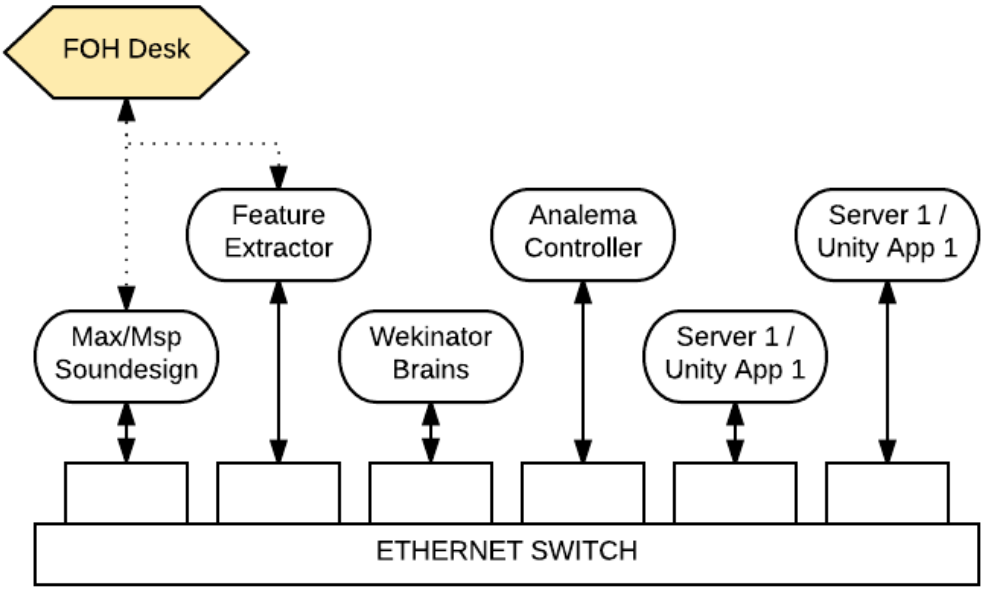

Fig. 10. Server Setup for KIMA: The Wheel. (© Analema Group 2017)

Participatory voice exploration

For the past two years, the collective focused on various forms of audience participation. Whether through music instruments or human voice input, the audience experienced a visual response to sound in real-time. Iterations exhibited at 'Incloodu Deaf Arts Festival' – the largest arts festival for the Deaf Community in the UK - and 'Kinetica Art Fair' paved the way to this large-scale participatory installation. At the Roundhouse, the collective brought the audience into the center of performance through spatial arrangements of sound as well as interactive visual elements, facilitated by performers, but enacted by the audience. The piece was conceived as participatory, in the tradition of an 'open work'[xxx]. Such an 'Open work' suggests various open ended possibilities for the actual final composition, where every participant, performer or member of the audience, contributes and therefore changes the course of the performance. At the Roundhouse, the great number of sonic responses from the audience encouraged further development of participatory experiences of KIMA.

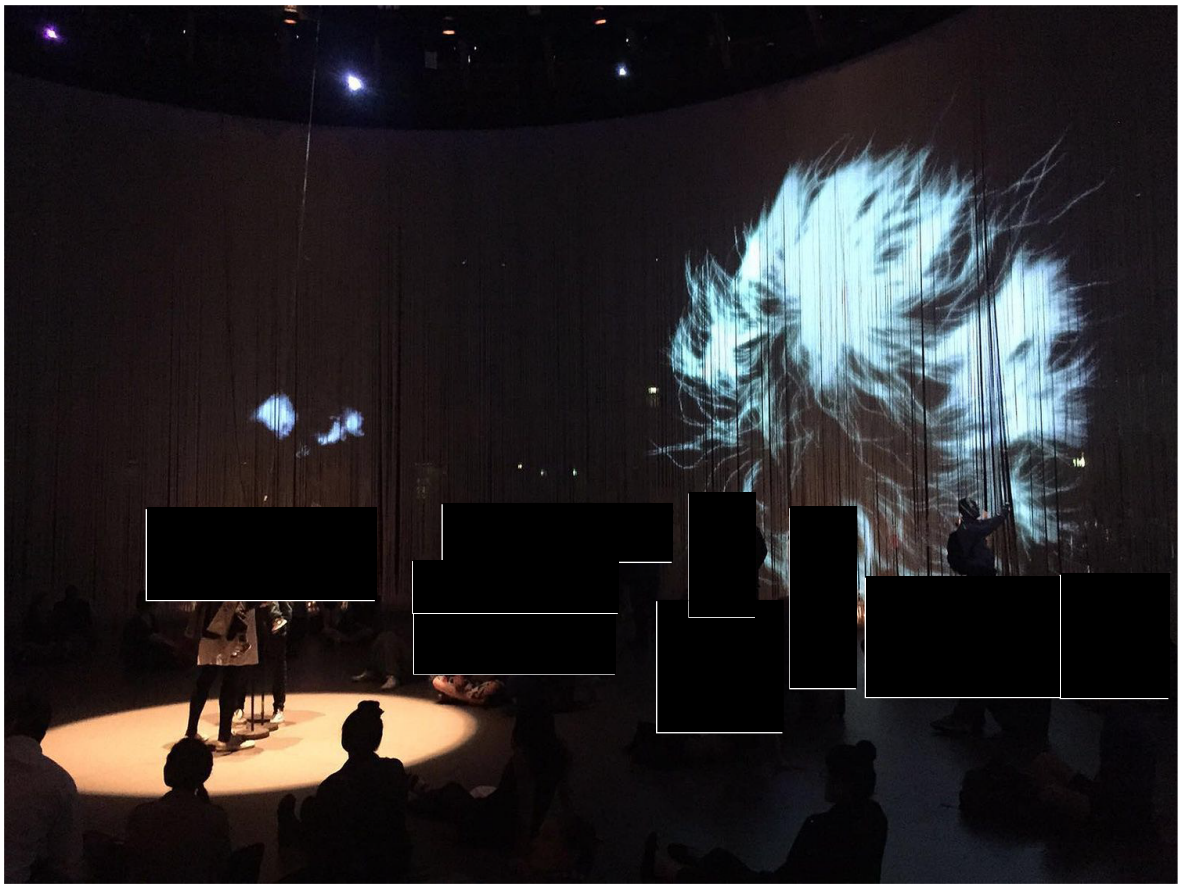

Fig.11. Sound representation of KIMA (© Analema Group 2016)

Discussion & Outlook :

Since the development of KIMA: The Wheel, KIMA was presented at Oxford and Bournemouth University and exhibited abroad.. At Fondation L'Abri, a centre for young art in Geneva, Switzerland, the collective encouraged musical participation with Tibetan singing bowls and the voice. The audience was confronted with real-time interference pattern emerging between different sound sources. Future development of KIMA will explore this space in-between sounds, tones and instruments, the ideas of resonance and dissonance, harmony and atonality. The focus on audience participation furthered the project substantially, with a renewed emphasis on the role of the audience as active performer. Augmented by the integration of written word into the narrative arch, the future will also see a concentration on colour, mathematical wave equations, and a continued effort to mirror sound within the architecture of the spaces we work with. As an environment for participatory sound exploration, KIMA builds a framework for audiences and artists to experience sound differently - as geometry, light, colour or form.

Fig. 12. KIMA: The Wheel (© Analema Group 2016)

Acknowledgments: *Our gratitude goes to the Roundhouse London, Arts Council England, Blitz, 7th Sense, Katherina Loschinina and Victor Vorski, Dan Cox and Zoe Leonard at the Centre for Digital Entertainment.*

References & Notes

---

[i] Waller, Steven, "Rock Art Acoustics in the Past, Present and Future," *1999 IRAC Proceedings,* Volume 2, 11--20 (1999).

[ii] Henry, Michel, *Seeing the Invisible* (London: Bloomsbury, 2009)

[iii] Kelly, Caleb: *Cracked Media. The Sound of Malfunction* (Massachusetts: MIT Press, 2009)

[iv] Whitney, John: *Digital Harmony: On the Complementarity of Music and Visual Art, John Whitney Sr.* (New York City: McGraw-Hill Inc, 1980) p.174

[v] Shaw, Jeffrey, *Future Cinema: The Cinematic Imaginary After Film* (Massachusetts: MIT Press, 2003)

[vi] Iles, Chrissie, *Dreamlands: Immersive Cinema and Art* 1905-2016 (Massachusetts: MIT Press, 2016)

[vii] Kholeif, Omar, *Electronic Superhighway. From Experiments in Art and Technology to Art After the Internet.* (London: Whitechapel Gallery, 2016)

[viii] Chladni, Ernst F. F., *Treatise on Acoustics. The first comprehensive English Translation of E.F.F. Chladni's Traité d'Acoustique* (NewYork: Springer 1967)

[ix] Jenny, Hans, Cymatics, *A Study of Wave Phenomena* (Berkeley: MACROmedia Publishing, 1967)

[x] von Bismark, Gottfried,"Timbre of Steady Sounds: A Factorial Investigation of its Verbal Attributes," *Acta Acustica united with Acustica 30(3):146--159* (1974).

[xi] Alluri, Vinoo & Toiviainen, Petri, "Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre," *Music Perception: An Interdisciplinary Journal,* 27(3):223--242 (2010)

[xii] Zacharakis, Asterios, Pastiadis, Konstantinos, Reiss, Joshua D, "An Interlanguage Study of Musical Timbre Semantic Dimensions and Their Acoustic Correlates," *Music Perception: An Interdisciplinary Journal, 31(4):339--358 (2014)*

[xiii] Grau, Oliver, *Virtual Art: From Illusion to Immersion* (Massachusetts: MIT Press, 2003)

[xiv] Bellman, Richard, *Dynamic Programming* (Princeton, NJ: Princeton Univ. Press, 1957)

[xv] Fiebrink, Rebecca, Trueman, Dan and Cook, Perry "A meta- instrument for interactive, on-the-fly learning," *Proceedings NIME,* (2009).

[xvi] McLaughlin, Maya Shara, *Life in Samsara: Torment, Torture and Hell,* (Annandale-on-Hudson, New York: Bard College, 2013)

[xviii] Smalley, Dennis, "Defining timbre—refining timbre," *Contemporary Music Review*, Jan 1;10(2), 35--48 (1994)

[xix]Pratt, R and Doak, Philip, "A subjective rating scale for timbre," *Journal of Sound and Vibration 45.3, 317--328. (1976)*

[xx]Kendall, Roger and Carterette, Edward, "Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's adjectives*," In: Music Perception, An Interdisciplinary Journal,* Vol. 10 No. 4, 445--467 (1993)

[xxi]Disley, Alastair, Howard, David & Hunt, Andy, "Timbral description of musical instruments," *International Conference on Music Perception and Cognition, Proceedings 2006,* 61--68. (2006)

[xxii] Zacharakis, Asterios, Pastiadis, Konstantinos, Reiss, Joshua D, "An Investigation of Musical Timbre: Uncovering Salient Semantic Descriptors and Perceptual Dimensions," *ISMIR Proceedings 2011,* 807 -- 812 (2011)

[xxiii]: Grey, J. M. and J. W. Gordon. Perceptual Effects of Spectral Modifications on Musical Timbres. In: The Journal of the Acoustical Society of America 1978, 63:5, 1493--1500 (1978)
[xiv]: Wessel, David, "Timbre Space as a Musical Control Structure," *Computer Music Journal,* 45--52 (1979)
[xxv]: Wold, Erling, Blum, Thom, Keislar, Douglas, Wheaton, James, "Content-Based Classification, Search and Retrieval of Audio," *MultiMedia, IEEE* 3:3, 27--36 (1996)
[xxvi]: Agostini, Giulio, Longari, Maurizio & Pollastri, Emanuele, "Musical Instrument Timbre Classification With Spectral Features," *EURASIP Journal on Applied Signal Processing* 2003, 5--14 (2003)
[xxvii] Peeters, Geoffrey, Giordano, Bruno, Susini, Patrick, Misdariis, Nicolas, McAdams, Stephen, "The Timbre Toolbox: Extracting audio descriptors from musical signals," *The Journal of the Acoustical Society of America. Nr. 130, 2902 (2011)*
[xxviii] Ethington, Russ & Punch, Bill, "SeaWave: A system for musical timbre description," *Computer Music Journal,* No. 18(1), 30--39(1994)
[xxix] Faure, A., McAdams, S., Nosulenko, V, "Verbal correlates of perceptual dimensions of timbre," *Proceedings of the 4th International Conference on Music Perception and Cognition,* 79--84 (1996)
[xxx] Eco, Umberto, *The Open Work (Massachusetts: Harvard University Press, 1989)*

Biographical Information:

Dr. Oliver Gingrich, is artist, researcher with a focus on presence research olivergingrich.com
Dr. Alain Renaud is researcher and sound designer, based in Geneva http://www.alainrenaud.net/
Dr. Sean Soraghan is researcher, and sound developer https://seansoraghan.github.io/
Evgenia Emets, is artist and creative director at Analema Group http://www.evgeniaemets.vision/
Dario Villanueva-Ablanedo, is a visual developer and data artist http://www.alolo.co/about-me/