Hologlyphics: volumetric image synthesis performance system

Walter Funk*
Hologlyphics, San Francisco, USA

ABSTRACT

This paper describes a novel volumetric image synthesis system and artistic technique, which generate moving volumetric images in real-time, integrated with music. The system, called the Hologlyphic Funkalizer, is performance based, wherein the images and sound are controlled by a live performer, for the purposes of entertaining a live audience and creating a performance art form unique to volumetric and autostereoscopic images. While currently configured for a specific parallax barrier display, the Hologlyphic Funkalizer's architecture is completely adaptable to various volumetric and autostereoscopic display technologies. Sound is distributed through a multi-channel audio system; currently a quadraphonic speaker setup is implemented. The system controls volumetric image synthesis, production of music and spatial sound via acoustic analysis and human gestural control, using a dedicated control panel, motion sensors, and multiple musical keyboards. Music can be produced by external acoustic instruments, pre-recorded sounds or custom audio synthesis integrated with the volumetric image synthesis. Aspects of the sound can control the evolution of images and visa versa. Sounds can be associated and interact with images, for example voice synthesis can be combined with an animated volumetric mouth, where nuances of generated speech modulate the mouth's expressiveness. Different images can be sent to up to 4 separate displays. The system applies many novel volumetric special effects, and extends several film and video special effects into the volumetric realm. Extensive and various content has been developed and shown to live audiences by a live performer. Real world applications will be explored, with feedback on the human factors.

Keywords: volumetric display, autostereoscopic, music, spatial image processing, animation, video synthesis, spatialization, special effects

1. INTRODUCTION

A system for the live performance of volumetric and autostereoscopic images along with music has been developed. All images are automultiscopic, ie. multiple true 3D views are seen by observers with no 3D glasses. Work has been carried out on both a parallactiscope and an experimental swept volume display. Most work and live performances to date have been carried out on a parallactiscope. When referring to visuals and performances in this paper, most images were displayed on the parallactiscope. The parallactiscope[1,2] is an automultiscopic dynamic parallax barrier CRT display first reported by Homer B. Tilton in 1971. It displays monochromatic 3D images with full horizontal parallax within a 90° viewing range; all views are present within that viewing range. This is possible by using an electro-mechanical moving virtual slit filter. There are no pseudo-scopic zones when viewing the parallactiscope. It is based on the Stereoptiplexer[3,4] by Robert Collender, which used a rotating drum with a vertical slit to display autostereoscopic movies.

2. HISTORY

The project started out in 1994 as a vision to merge holography with music. The idea was to have moving holograms follow and be synchronized with live music, by creating powerful artistic tools that would inspire creativity. Performances would have structure with generous amounts of room for improvisation. After much research, I abandoned the idea of using holography and settled on using the parallactiscope as a display medium. The first performance system consisted of a digitally controlled analog image synthesis network. The network was based around two digitally controlled 16 by 16 analog cross point matrix devices. The cross point matrices would route image and control signals in a modular fashion.

Further author information: (Send correspondence to Walter Funk)
*walter@hologlyphics.com; hologlyphics.com

Stereoscopic Displays and Applications XIX, edited by Andrew J. Woods, Nicolas S. Holliman, John O. Merritt,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6803, 680308, © 2008 SPIE-IS&T · 0277-786X/08/$18

With digital control over the analog signal paths of the various image generators and processors, audio generators and processors, plus separate control of how the control voltages connect to the generators and processors, the Hologlyphic Funkalizer[5] allowed for an enormous amount of of 3D visuals and sound to be created, altered and animated in real-time under control of a live performer, via a modular synthesis architecture.

These cross point matrices allow digital control of routing each of the 16 analog signal inputs to any of the 16 outputs. Connected to the inputs of the first cross point matrix were the outputs of various analog voltage controlled 3D image generators, 3D visual processors, audio oscillators, and audio signal processors. The outputs of the first cross point matrix fed back into the 3D image processors, went to the inputs of multiple audio speakers and fed into the inputs of the parallactiscope. Connected to the inputs of the second cross point matrix were the outputs of various control voltage generators. The outputs of the second cross-point matrix connected to the control voltage inputs of the voltage controlled 3D image generators, audio oscillators, audio signal processors and 3D visual processors. One of the control voltage generators was an analog pitch/loudness follower, allowing an acoustic musical instrument to control image synthesis parameters in real-time.

The 3D visual generators consisted of hyper-chaos pattern generators and a Helix Generator. The 3D visual processors were a Hyperbolic Paraboid Generator, voltage controlled phase shifting networks, analog multipliers, signal delays, and waveform filters. The modulation sources consisted of low frequency oscillators, chaotic waveform generators, and pitch/loudness following for acoustic instruments. The hyper-chaos pattern generators were cross-coupled frequency modulated waveform generators, outputting chaotic x, y, and z signals. The Helix Generator and Hyperbolic Paraboid Generator were originally designed by Homer B. Tilton for general use with the parallactiscope. The helix generator was a 3-phase waveform generator, with a manually controlled analog sin/cosine oscillator, output X was sin, output Y was cosine, sweepable from 1k to 20k, plus an output Z with a 1k sawtooth waveform. The ratio of sin/cos waveform frequency to the stationary 1k sawtooth waveform determined how many coils/winds/turns the helix had. The Hyperbolic Paraboid Generator was an analog spatial image processor able to transform the topology of an image and synthesize new images.

3. NEW ANALOG/DIGITAL IMPLIMENTATION

3.1 Basic design

The system has been redesigned to make use of digital signal processing. A computer uses real-time digital spatial image processing to emulate all the previous visual synthesis functions, audio generation, signal switching and modulation functions. The original 3D visual synthesis functions were greatly expanded and many new image generation and processing functions were developed. True automultiscopic images can be routed to as many as 4 different 19-inch displays, each with different content. The displays are vector addressed. I created a family of real-time spatial image synthesis and processing algorithms for artistic use.

Real-time manual control has been greatly expanded as well. The animations are controlled with various interfaces, all sending their respective information to the digital signal processing network. These interfaces are:

- Musical Keyboard- standard Midi keyboards
- Control Panel- connected to USB port
- Motion Sensors- connected to USB port
- External Signal Input- for use with acoustic instruments and microphones
- Computer Keyboard- keys execute changes and commands

3.2 Parallactiscope hardware

The parallactiscope is an automultiscopic display using 4 main components: an electrostatically deflected CRT, a Parallax Scanner, a Parallax Computer, and a Scanner Driver. The Parallax Scanner is an electronically controlled direction-sensitive spatial filter placed in front of the CRT. The Parallax Scanner is constructed with two crossed linear polarizers and a vertical sliver of half-wave retarder in the center. The crossed polarizers block the light coming from the CRT, but, with the .10 inch sliver of half-wave retarder in the center, after the light passes through the first linear polarizer and the half-wave retarder, the light is rotated 90° and then is able to pass through the subsequent linear polarizer. This forms a virtual slit which allows only particular light rays to pass into observer space while blocking all others. The sliver of half-wave retarder is mounted on a high-Q pendulum connected on its other end to a small audio loudspeaker. The speaker acts as a linear drive motor, allowing the virtual slit to scan horizontally in front of the CRT. If the slit moves side to side fast enough, it is no longer perceived as a slit.

The Parallax Scanner slit's horizontal motion is driven by the Scanner Driver with a sinusoidal signal. A second sinusoidal signal with the same frequency, but variable phase, feeds into the Parallax Computer. The Parallax Computer is a special purpose analog computer, which accepts x, y, and z deflection signals and processes them to provide the proper horizontal and vertical deflection signals for the CRT. Since the Parallax Computer receives a signal from the Parallax Scanner, it is able to keep track of the slit's instantaneous horizontal position at all times. As the CRT screen is viewed through the slit, each narrow vertical zone on the CRT screen is keyed to a unique horizontal viewing direction. This configuration allows controlled parallax: the Parallax Computer processes the x, y, and z signals in relation to the slit's instantaneous position, displaying the proper image on the CRT, which, when viewed through the Parallax Scanner, produces automultiscopic images. Autostereopsis is produced, along with multiple views allowing 90° horizontal parallax, with no pseudo-scopic zones. The displayed 3D image represents the x, y, and z deflection signals fed into the Parallax Computer. All images are directly drawn onto the display; there is no raster. The CRT must be electrostatically deflected to obtain a sufficient redraw speed, and must have a high intensity, short persistence phosphor. High intensity is required due to some loss of brightness with the slit; the short persistence helps avoid smearing. Fig. 1 shows the basic large screen parallactiscope hardware.

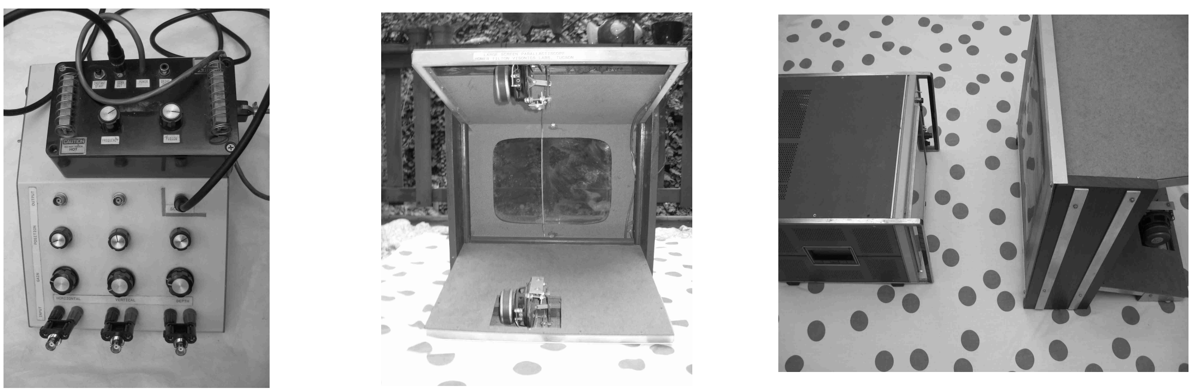

Fig. 1. Parallax Computer with Scanner Driver on top, large Parallax Scanner with one linear polarizer removed to show slit and pendulum, and a large screen electrostatically deflected CRT in front of Parallax Scanner.

3.3 Hardware for image generation, audio, and performer interfaces

To display moving images on a parallactiscope, synthesized via digital signal processing on a computer, the x, y, and z inputs of the Parallax Computer are fed from a multi-channel digital-to-analog converter connected to the computer. The converter is a multi-channel M-Audio soundcard running at 192 khz. A blanking output signal from the DAC controls the intensity modulation of the CRT. Various real-time signal processing oriented programming languages such as CSound, PD, and MSP are used to run spatial image processing and sound synthesis algorithms, defining 3D images by XYZ signals. The XYZ signals of the analog image generators are fed into a 360 Systems 16x16 switching matrix, which has a bandwidth of 0-60 khz, then feed into the soundcard's ADC. Those external XYZ signals can then be routed within the computer, processed and output to the Parallax Computer. Audio is also output from the computer via the multi-channel soundcard and routed to four speakers. Using multiple outputs of the DAC, up to four sets of XYZ and intensity signals can be routed to as many as four parallactiscopes.

The midi keyboard is connected to the computer through an M-Audio midi converter. The sensors are three motion based sensors, feeding into a Teleo sensor board connected to the computer's USB port. The control panel connects to the USB port as well, and external audio, to be analyzed for controlling 3D image synthesis parameters, is run into the ADC input of the multi-channel soundcard. Fig. 2 shows some of the interface devices.

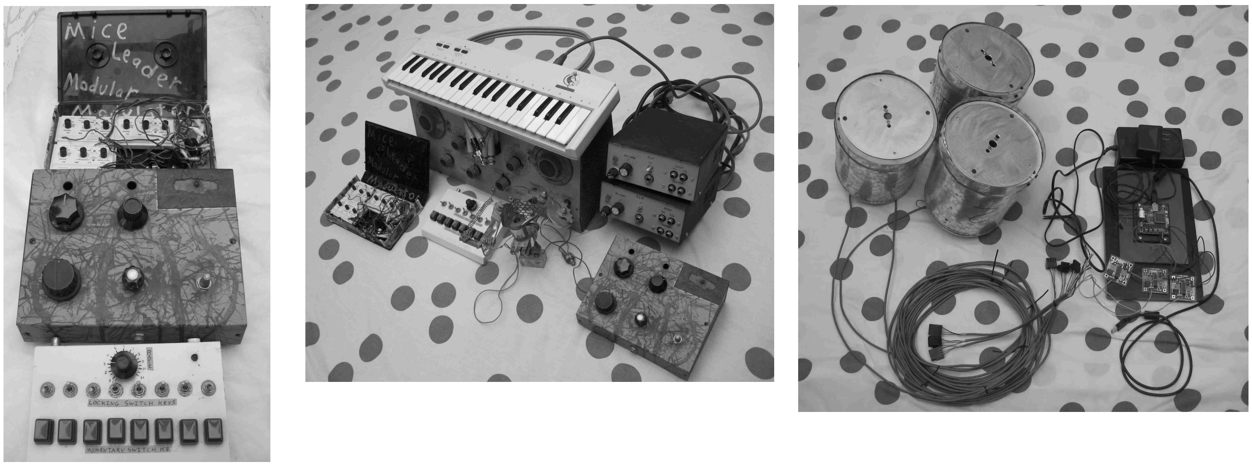

Fig. 2. Control panels closeup, control panel-image generators-midi keyboard interfaces, and optical motion sensors inside buckets with Teleo sensor interface for computer.

3.4 Experimental display hardware

An alternate display, used on an experimental basis, is a varifocal mirror display. The display is addressed completely by the computer; there are no circuits for generating or processing signals. A mylar membrane is fitted over a loudspeaker, and the speaker is directly scanned from the output of the computer's DAC with a sinusoidal signal. The computer also sends deflection signals to the CRT, displaying images in sync with the mirror's vibration. The CRT screen is reflected by the mylar membrane and, as the mirror moves from concave to convex due to the sinusoidal scanning, image depth is produced. Image correction before output to the CRT is necessary because the deeper the mirror's scan is, the smaller the image appears. The displayed 3D images are interesting, but, at this time, lack the detailed controllability available with the parallactiscope. Further work will be done to improve this.

3.5 Volumetric animation/sound synthesis architecture

The Hologlyphic Funkalizer has a distinct architecture for generating volumetric images with sound. Both visuals and sound can be grouped into three main processes: synthesis, processing and control. Synthesis involves the creation of a visual or sound as a digital signal, wherein processing takes the generated signal as an input and transforms it. Both forms of synthesis and processing can be controlled by the interfaces and internal modulators.

The 3D scene generation process is divided into two parts, synthesis and processing. 3D shapes, scenes and patterns can be generated using either 3D vector graphics or digital oscillators. Then many combinations of spatial image processors can subsequently process the images. The 3D vector images, which are vector list based, can also be processed through a list processor, before being sent to the spatial image processing network. Sound synthesis is also divided into two main parts, generation and processing. Sounds can be generated via digital oscillators, sampling, and multi-channel sound file playback. Then the audio can be processed by many combinations of signal processors, using a different set of algorithms than the visuals, with the same overall structure, then spatially distributed through multiple loudspeakers.

For both 3D images and sound, the two synthesis processes, generation and processing, can be dynamically modulated. Basic types of modulation parameters for generated images include shape changes, shape morphing, pattern sequences, and scaling. Basic types of modulation parameters for the 3D images signal processors include scaling , rotations, xyz coordinate motion, kaleidoscopic transformations, 3D video wipes, and multi-image morphing. Basic types of modulation parameters for generated sound include timbre, amplitude, frequency, enveloping, and morphing. General types of modulation parameters for the audio signal processors include filter parameters, custom spectral effects and spatial placement/position. There are many more modulation parameters and transformations available, too many to describe in detail in this paper, but a few interesting ones will be described in detail below.

Control signals can be low frequency oscillators of many waveshapes, random and chaotic functions, inputs from external interfaces such as motion sensors, perceptual features from acoustic instruments, musical keyboards, plus knobs and sliders. The parameters of the modulators themselves can also be controlled and modulated, either by other modulators or any of the external interfaces.

3D visuals and sound are integrated, and can be linked in unique and interesting ways. They can share modulation sources or variations and derivatives of specific modulators. When specific images appear, they can have their own associated sound triggering at the same time. If that same image is modulated in size, rotation, or shape, the associated sound can be modulated by multi-channel spatial audio effects at the same rate of change as the visual. This creates a synchronicity between the visuals and sounds. A basic setup performance setup is shown in Fig. 3.

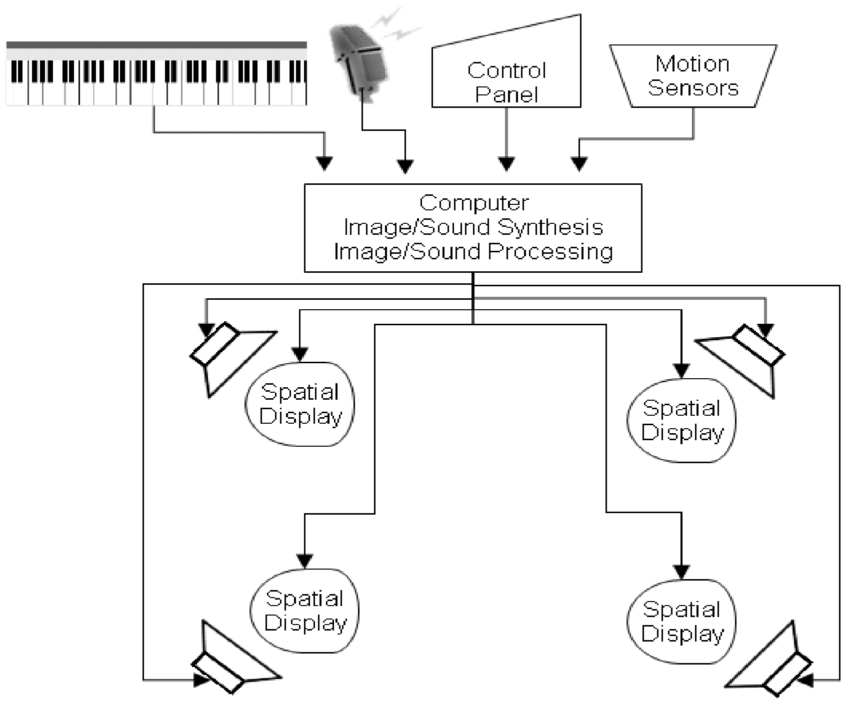

Fig. 3. Diagram shows a basic setup configuration. Musical keyboards, microphone, the control panel and motion sensors connect to the computer. 3D Audio/Visual scene generation is controlled by the interfaces, then output to four spatial displays and speakers. The performer and audience are in the middle.

3.6 Volumetric image synthesis (VOLSYN)

There are two main ways images are digitally synthesized. 3D shapes, scenes and patterns can be generated using either 3D vector graphics or digital oscillators. The oscillators put out x, y, and z signals in the form of multi-phase multi-waveforms, with many different complex waveform shapes to choose from. The VOLSYNs have modulation inputs that alter the images' visual appearance, aspects such as overall topology, non-linear warping, and shape morphing. There are many VOLSYN algorithms that generate their own unique families of images. The individual VOLSYNs can be added together to produce highly controllable and complex shapes. The vector graphics also output x, y, and z signals. Once generated, the three x, y, z outputs of both the digital oscillators and the vector generators are then available to be input into any combination of the spatial image signal processors.

Besides the digitally generated images, several of the analog image generators remain from the first design, especially the hyper-dimensional chaos pattern generators, shown in Fig. 4. They are very hard to emulate digitally due to the non-discrete nature of analog electronics. The generated images are now fed into the analog cross-point switching matrix, then routed into the host computer. The three x, y, z signals are then available to be internally routed into the computers

Volumetric Image Processing(VOLPROC) system.

Some of the basic VOLSYNs include Helix generators, UFO generators, Tunnel generators, Crop Sphere generators, Glyph and Musical Note generators. The Crop Sphere generators create geometric spatial weavings reminiscent of crop circles. The Glyph and Musical Note generators create unique spatial invented alphanumeric characters, glyphs, musical notations. When displayed sequentially in time with music, these invented glyphs help create a unique musical syntax and personality of the Hologlyphics system. Many more specialty VOLSYNs have been developed, this is just a small overview. Some basic images are shown in Fig. 5.

Fig. 4. Analog hyper-dimensional chaos pattern generator.

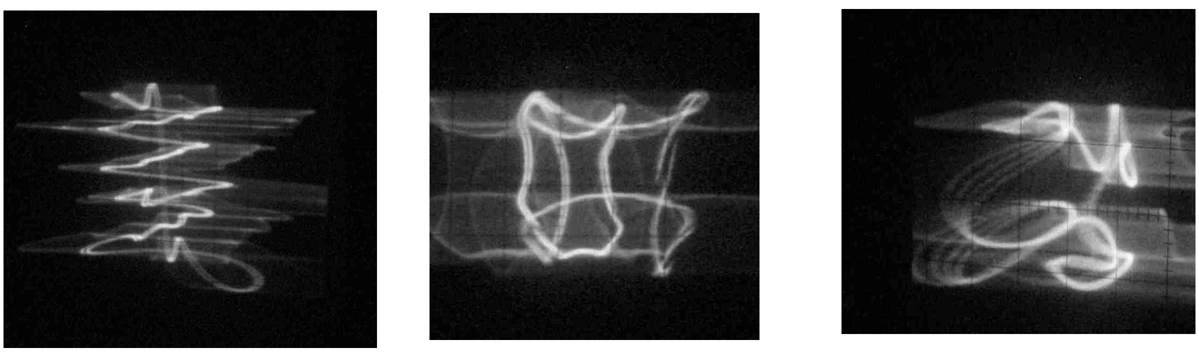

Fig. 5. A few simple glyphs. Complex images and scenes are difficult to capture with photography.

3.7 Volumetric image processing (VOLPROC)

The spatial image signal processors accept three signal inputs, x, y and z, which describe a moving volumetric image. The 3D image signal processors then transform the volumetric image in artistically useful ways. They can be used in a modular fashion and patched in either a serial or parallel manner. Many modulation parameters are available to alter visual aspects of the transformation process. The basic VOLPROCs provide scaling, movement of the images' xyz coordinates, and rotation around the 3-axes.

The Scaling VOLPROC accepts a modulation input that can dynamically alter the size of the object. The Movement VOLPROC accepts 3 modulation inputs that can dynamically alter the spatial location of the object, one each for up/down, right/left and forward/backwards. The Rotation VOLPROC accepts 3 modulation inputs that rotate the image 360° around each axis. Besides the basic VOLPROC algorithms there are many artistic effect processors, including those extended from traditional film and video effects as well as new and more exotic types.

Video wipes are a process of graphically combining two video signals. They were originally developed in the television industry along with other video effects to be the electronic equivalents of film wipes, fades, superimpositions, and traveling mattes. The original film techniques were a combination of optical, chemical and mechanical processes. Many video artists later developed devices that implemented video wipes in creative, unusual and exotic ways. A video wipe takes two video signals, signal A and signal B, and allows a transition from one to another. A vertical wipe starts with image A, then the screen turns into a vertically split screen as image B slowly wipes across the screen from one side to another. At midway point the split screen is half image A on one side and half image B on the other. At the extreme end of the wipe all you see is image B on the screen. Wipes can also happen on the horizontal and diagonal planes, or be diamonds, ovals, even more complex geometric shapes. The wipe and its shape can be dynamically controlled by manual or electronic means.

Volumetric image wipes work on the same principal as video wipes, except they work in the spatial realm, by combining two volumetric images, volumetric image A and volumetric image B, and splitting along a whole plane, not just a line. If the wipe were to be a vertical volumetric wipe, you would see a split 3D scene, with one scene on the left and the other on the right. The split divides the 3D scene into two different 3D sections, just as the video wipe divides the video screen into two different 2D sections. With the volumetric video wipe, you can also rotate the splitting plane 360° along any of the 3 axes. You can rotate it to create a horizontal wipe, forwards/backwards wipe, all kinds of diagonal wipes, and you can rotate the splitting plane while wiping, to create a spinning wipe effect. An example of a forwards/backwards wipe is given in Fig. 6. Besides using a 2D splitting plane, volumetric wipes can be cylinders, cubes, spheres or arbitrary 3D shapes. Depending on the two spatial images used, most wipes have a bit of spatial distortion along the wipe plane. In the spirit of the unique video synthesizers invented by video artists, this spatial distortion can be controlled and emphasized. I have developed special wipe algorithms with spatial interplay between both images such as Morphing VOLPROCs that create a seamless transition between two images by interpolating and distorting their corresponding similarities.

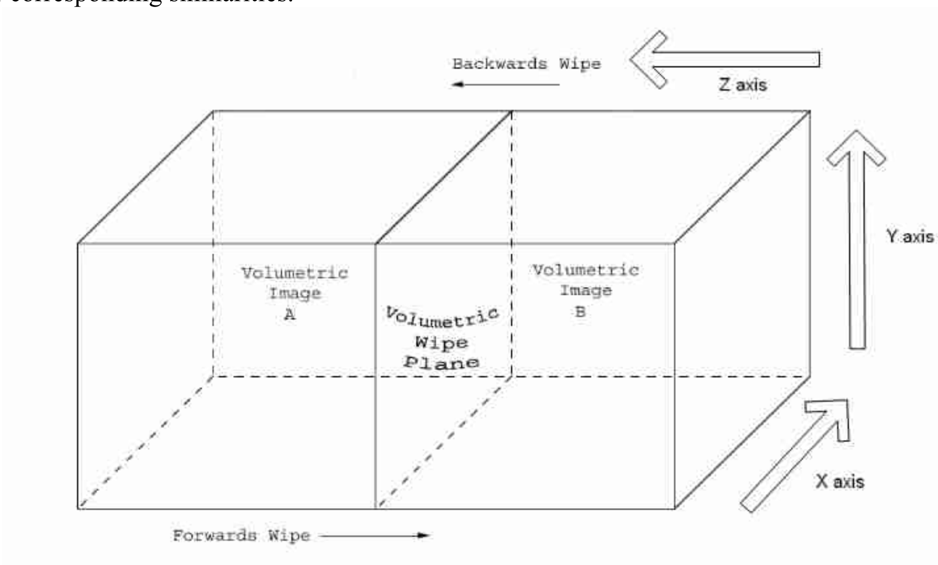

Fig. 6. Volumetric wipe. A side view of a forwards/backwards wipe along the Z axis.

Along with the video wipe, the video keyer is one the most common of video effects. Keying is the video equivalent of cinematic matting, with the purpose of causing one image to be inserted into another image, by controlling the electronic selection of one video signal in favor of another. Only those areas of an image which are above a certain threshold level are inserted into the other video image. For example, a person standing against a green or blue background can be keyed into a scene with a jungle landscape, so that they appear to be walking in the jungle. The true 3D nature of volumetric displays has made it difficult to apply video keying techniques to a volume. Several experimental approaches were taken; while the results were sometimes very interesting, none could be called volumetric keying. They were expanded on and put into the category of special function image processors. Further keying research will be continued when newer rendering techniques and display types become available to the Hologlyphics system.

Scan processing is another early video effect, more esoteric than the common video wipe or keyer, employed by such machines as the Scanimate and the Rutt-Etra Video Synthesizer.[7] It involves removing the synch information from a rasterized video signal and converting the raster information into two XY signals plus a separate intensity signal, then separately processing those three signals to re-shape and warp the raster. The signals can be transformed by stationary values, dynamic manual control, or be modulated by oscillators. Depending on the type of oscillator used and its relation to the X and Y signals' frequencies, i.e. different frequency ratios, different waveshapes, phase relationships, all sorts of scan processing shapes and patterns are possible. The processed XY and intensity signals are then scan converted by being displayed on an oscilloscope and re-scanned via a video camera. The variations possible are so wide, that different scan processing schemes can definitely be classified into many categories of special/video effects.

Volume bending is an extension of scan processing. Since there is no raster used in the Hologlyphics system, the XYZ signals are simply directly processed, and on output there is no need for re-scanning. The volume of an image can be twisted, bent, warped, stretched by a multitude of signals, usually bearing some relation to the image itself. As with scan processing, the specific relation between the modulating oscillator and the XYZ signals, plus the type of volume bending used, creates a specific effect.

Kaleidoscope effects have also been added to the VOLPROC family of spatial signal processors. These effects are based on mirroring and symmetry image processing functions. In the basic Kaleidoscope effect, the scene volume is split into two, four, or eight sections, then only one of those sections is selected and the remaining ones are blanked, using spatial clipping, and then the selected section is mirrored symmetrically around the other blanked sections. With the scene split into two sections, true 3D Rorschach type effects can be made. When it is split into four parts, a tubular Kaleidoscope effect is created, with eight parts a true 3D spherical Kaleidoscope image is generated. Not only can the Kaleidoscope be animated as the volumetric image moves, but if the image is stationary and the section selected to be mirrored is changed, as can be done by a modulation control, the sweeping of the modulation will cause the Kaleidoscope to animate. Besides the basic Kaleidoscope effect, there are special function Kaleidoscope effects that provide expanded spatial image warping effects of the kaleidoscopic image, before, during and after the mirroring process.

More exclusive to spatial imagery, Spatial Warping VOLPROCs also are commonly used. They have some similarities to volume bending, but the processes and results are different because the space itself is being bent . Quadrant/Octant Space Effects can be assignable by spatial region, allowing VOLPROCs to be applied only to a specific spatial portion of the image. Asteroids FX, named after the classic vector video game, allows images to move off the display on any spatial direction, reappearing on the "other side" of the display. Using this technique, images can also be scaled larger than the volume of the display, allowing all sides of the image to reappear on the opposite side, creating a densely overlapping zoom effect. Many other unique specialty VOLPROCs using similar concepts are also used sparingly.

3.8 Control interfaces and internal modulators

The system also provides a method of control over the parameters of the generators and signals. There are two categories of control, manual control and internal modulation. Manual controls consist of knobs, pots, sliders, xy joysticks, midi keyboards, motion sensors, perceptual analysis of acoustic instruments, etc. Internal modulators consist of time varying signals of various multi-phase waveforms, spatial oscillators, algorithmic sequences and chaotic functions. Any of these controls can be mapped to any parameter of the VOLSYN and VOLPROCs, plus the parameters of the audio synthesis and processing algorithms, allowing control of the quality of the synthesized image and sound in an integrated manner. The mapping can be transformed to be scaled and non-linear, to provide even more control over the parameters. A sequence of events can be triggered by a motion sensor, images played with a keyboard, or images guided and re-shaped by knobs, joysticks and sliders. A basic example of how control signals are mapped to the parameters in an integrated manner is shown in Fig. 7.

Using perceptual analysis of acoustic instruments, perceptual features such as pitch, loudness, and timbre can be analyzed and used as control sources. In many cases the instruments are acoustic sound sculptures with transducers attached. Perceptual analysis is also sometimes used with synthesized sounds. In the case of text to speech synthesis, a voice can be synthesized and the features extracted. Then the pitch, loudness, and timbre features can be non-linearly mapped to VOLSYNSs and VOLPROCs. If loudness features are mapped to a Scale VOLPROC, set to scale a Helix VOLSYN on the X and Y axes, the amplitude of the voice can be used to modulate the size of the helix. Using non-linear mapping, we can limit the range of modulation on the Scale VOLPROC, to allow for more flexibility. This in itself makes the helix seem if it is speaking the words. To increase expressiveness and complexity, other forms of image processing can be applied, then the pitch and timbre features of the voice can be non-linearly mapped to the Helix VOLSYN and various VOLPROCs.

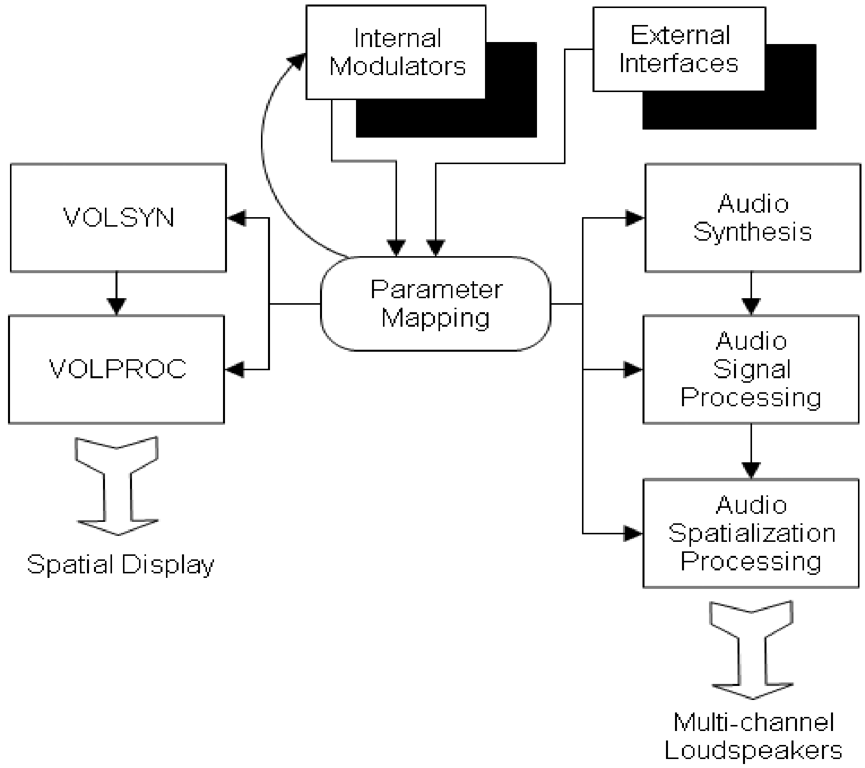

Fig. 7. Volumetric animation/sound synthesis structure. Spatial visuals and sound are integrated by sharing the same control sources.

3.9 Spatial sound integration

Sound is routed to spatial sound processors before being output into a 4 channel sound system. Just as sound quality can be integrated with the image synthesis, the spatial qualities can be integrated as well.

There are many types of 3D sound techniques that have been developed by researchers, including simple panning, Head-Related Transfer Functions (HRTF), Ambisonics, auralization, perceptual based models, and decorrelated multi-channel audio.[8,9,10] HRTF models are used with either headphones or loudspeakers with a very small 'sweet spot'. Since the images are true 3D, displayed without any special glasses, and are intended for multiple viewers in arbitrary public situations, systems requiring headphones or loudspeaker based systems requiring a "sweet spot" would be against the artistic and technical spirit of the project. Ambisonics techniques are more often used for the recording of live music through microphones, and therefore those exact techniques are not incorporated, although there have been a few conceptual elements regarding sound field rotation adapted to the Hologlyphics spatialization algorithms. Auralization refers to the processing of sounds to emulate the acoustic features of a specific room or space. This technique is not used in the Hologlyphics system as a signal processing option since it has no relation to the abstract or impressionistic images displayed. This would be applicable if a volumetric display showing a scene in which a real or fictional space was represented.

I adapted a combination of perceptual based models and decorrelated multi-channel audio concepts. With multi-channel panning and perceptual based spatialization, we have control over the movement of sound,[8,10] and with decorrelated multi-channel audio we have control over envelopment and spaciousness.[11,12] This corresponds to the visual qualities of motion and spaciousness that are typical of the spatial images in the Hologlyphics system. A correlation/uncorrelation parameter allows the perceived quality of spaciousness to be modulated by a control signal. With non-linear parameter mapping, we can also integrate the spatial modulations of visuals and sound, enabling their spatial cues reinforce each other, creating multisensory fusion. For example, image rotation linked with the spacial movement of sound, as in Fig. 8.

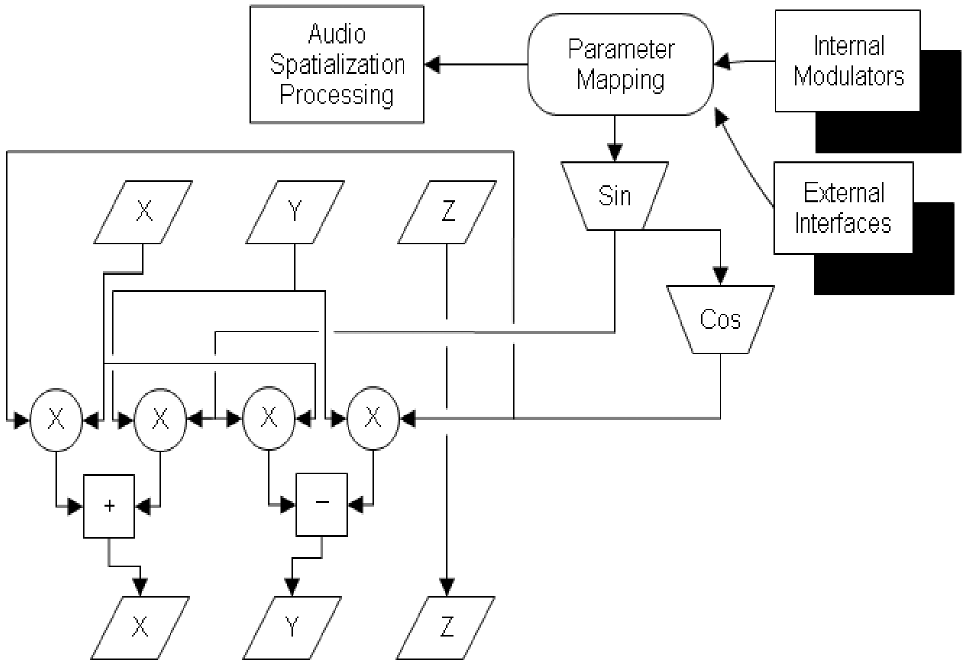

Fig. 8. XYZ signals feeding into a rotation matrix. Mutual control signals are mapped to image rotation and audio spatialization. Rotation is around the Z axis only in this case. Sound/visuals can be equally rotated in space or creatively linked other ways.

3.10 Pulling it all together- volumetric scene synthesis (VOLSCENE)

A number of spatial images are first generated simultaneously with synthesized or sampled sound. Then both the generated visuals and sound are added together, processed, and sequenced to generate a 3D scene. The visual 3D scene is displayed on one or more automultiscopic displays and the audio is spatialized through a multi-channel sound system.

We have seen how we can use VOLSYNs and VOLPROCs to gain artistic control of moving spatial images. After a suitable number of images and animations have been created, they can be pulled together to create a sequence of scenes that comprise a whole performance, using serial and parallel audio/image synthesis and processing techniques, mapping, scene switching, and interface devices.

After creating groups of images controllable by the control panel, musical keyboard, and acoustic instruments, these groups of images can be selected using the computers keyboard, enabling the performer to select any of the image groups, each with their own unique character, sound relation and control scheme. Instant visual feedback is available to the performer.

An example of scene creation will be described to provide a better understanding of how controllable 3D audio/visual scenes can be created from the various elements. We can use the simple Helix VOLSYN as the basis for a more complex scene. The Helix VOLSYN is processed in a serial manner through four VOLPROCs, in order, the Scale VOLPROC, Movement VOLPROC, Kaleidoscope VOLPROC, and finally the Asteroids FX VOLPROC. The Scale and Movement VOLPROCs are used to control the volume and position of the Helix before it is processed by the Kaleidoscope VOLPROC, allowing greater variability. The Kaleidoscope VOLPROC is set to slice the image into four regions. With the Helix in the center position, the kaleidoscope effect is not very strong. Once Modulation is applied to the Scale VOLPROC and Movement VOLPROCs, the region being mirrored by the Kaleidoscope VOLPROC becomes less symmetrical, adding complexity. Modulating the Kaleidoscope VOLPROC to select different quadrants of the Helix increases this even more. At extreme settings of the Scale VOLPROC and Movement VOLPROCs, since the Asteroids

FX VOLPROC is at the final stage of the processing, the Helix can be smoothly transformed into a diffuse field of double helixes and dancing stretched zig-zags. Using non-linear mapping, we can create a control parameter for the smooth transition from the simple, centered Helix image to the visually diffuse field, and by synching that mapped control parameter to the correlation/uncorrelation parameter of the audio spatialization processing, the spatial properties of the audio/visual scene can be fused. As the visual scene becomes more diffuse, the audio becomes more spacious and enveloping. This scene synthesis configuration is showing in Fig. 9.

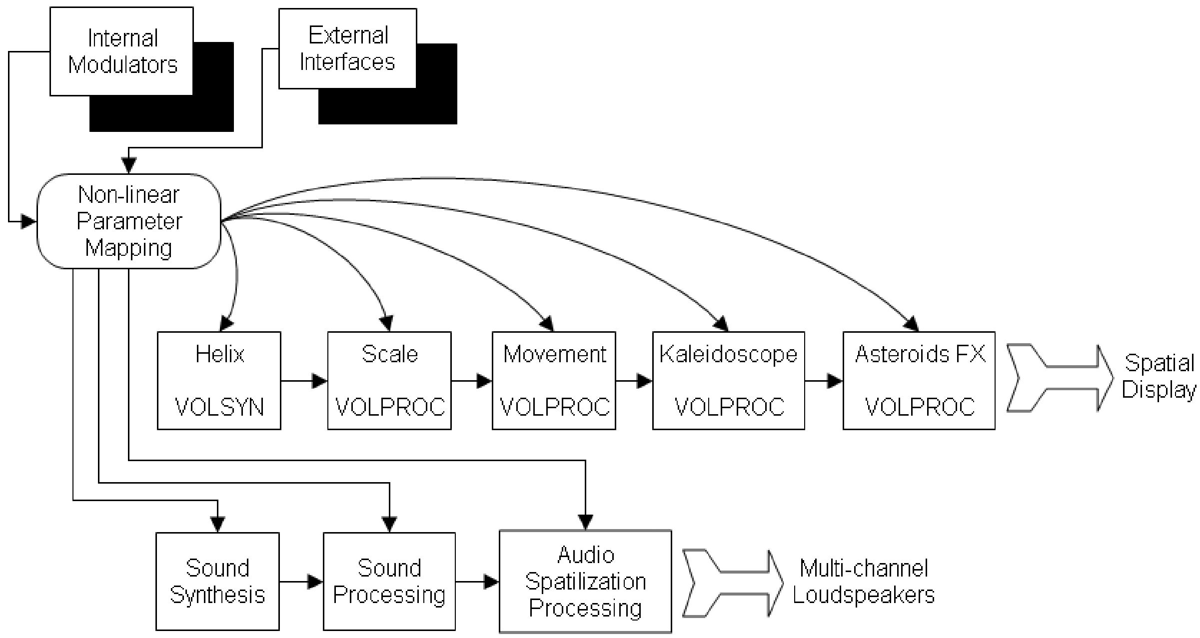

Fig. 9. Complex scene synthesis. Non-linear mapping allows the control of a whole scene. Scenes can then be linked together sequentially.

4. REAL WORLD USAGE: PERFORMANCES AND INSIGHTS

I have given numerous live performances in solo and ensemble settings. Resulting feedback and insights follow below.

Audience feedback has been extremely positive, and the correlations between the sounds and images have been perceived as clear, especially when live instruments, sound sculptures or motion sensors are used. The gestures of the musician/video-artist help bring together not only the audio/visual 3D scene, but also deepen a sense of the interconnectivity of it all. This is evident most with solo performances.

In terms of performance environments, darker is better; it does not need to be pitch black. Depending on the setting, the audio/visual experience is always intimate and sometimes immersive. In the most complex setup, there are four 19-inch parallactiscopes and four speakers, one in each corner of the room. This is the most immersive situation, and the immersion does not come from a sense that the audience is in an environment, but from the immersive feeling of the spatial ‘objects’ around them.

4.1 Artworks on multiview displays and the freedom of parallax

The parallax available to the viewer is indeed a very intimate experience. The whole scene is artistally composed from the beginning with all the views available in mind. Parallax is an intrinsic part of the artform. While there is no ‘right’ pacing for any moving visual artwork, many filmmakers have techniques, rules, frameworks and preferences for the pacing of shots. Some paces just ‘feel more natural’ for certain types of visual images and transitions.

There is a different "natural" pacing preferred for many images in the Hologlyphics system due to parallax. For certain very detailed images, with the viewers capability to ‘look around’ the object, it needs to be taken into consideration how much ‘time you leave’ to look around the object. Imagine a series of shots of a stationery spatial object displayed on a volumetric display; each shot lasts 3 seconds. With a stationery object displayed on a volumetric display, “you don’t see all you can see” for that shot as soon as your mind visualizes the object. This is not a consideration with works using only one point of view as is the case with video, film, and even traditional 3D-Movies where, even though the viewer sees 3D, only one perspective is viewable at any moment. Again, while no ‘right’ pacing exists, such techniques, rules, frameworks and preferences for temporal pacing of multiview shots will need to incorporate that consideration. The pacing considerations have been noted not only by the artist, but by colleagues and audience members as well. Untrained people at public art events have mentioned they did not get too see all of a scene before the shot changed.

There are two other observations about parallax that would be useful to the potential artist or filmmaker wishing to work with the automultiscopic artistic medium, assuming parallax is a wanted, welcome and wonderful capability. Although none of the artworks created to date have been narrative with a storyboard or presented on a very large size display, consideration, research and planning has gone into the potential. The following does not apply to the Hologlyphics system currently, but would need to be taken into consideration in order to produce a narrative work with multiple views. With traditional film, video and 3D movies, the director of a narrative story, through the cameraman, has complete and total control over the perspective of every shot the viewer perceives. This is a very powerful and integral tool in the art of filmmaking. With automultiscopic and volumetric displays, the viewer has a choice and control of his own perspective. This capability/freedom will differ depending on the nature of the display. With a 360° viewable volumetric display, where one has the freedom to walk all the way around, the director could not use the same form of perspective control that so many filmmakers have mastered.

The second consideration regarding multiviews and parallax is related to scaling and logistics. With parallax available, people want to move around to see different views. Technical or financial considerations aside, the implications of showing a highly multiview movie at a large theater on a very large screen and giving the audience full freedom to “look around” the scene brings up many problems. The scaling of the size of the screen to the observer would require a much longer time to move from one view to another. Combined with the large number of viewers, this would make it difficult for people to avoid bumping into each other. If the display were very large but only one viewer was present, one would still loose the intimacy of the experience that one has with smaller displays in regards to parallax. The larger the highly multiview display, the more time it takes to move to different perspectives.

A couple of other observations by audience members have been noted. In some of the more adventurous 3D scenes created, strange optical illusions have been displayed that radiated motion from the center outward, yet seemed to convey a strong impression that the viewer was moving into some sort of terrain. Although the imagery was obviously radiating outward on horizontal plane in equal proportions and was completely non photorealistic, a strong sense of deep viewer motion was experienced.

Many audience members have also pointed out, “This thing is just like when R2D2 projected Princess Leia’s message to Obi Wan Kenobi!”

5. CONCLUSIONS AND FUTURE WORK

I have developed a live performance system that synthesizes true 3D visuals integrated with music and sound. A live performer has geometric, spatial and temporal control over the visual animations by way of musical keyboards, sound sculptures, control panels and acoustic analysis. Visuals and sound are integrated at both the event and spatial levels, enabling multi-sensory fusion of the artwork/performance. Wipe, Kaleidoscope and other effects have been extended into the spatial realm as well as the development of new unique effects exclusive to multiview displays. Parallax achievable with volumetric and automultiscopic displays has proven to be a powerful artistic tool, and new paradigms need to be invented. A new art form has been explored, tying together volumetric animation, music, spatial sound, acoustic analysis, and human gestural control. Live performances have been carried out with great success. Audience feedback has been extremely positive, and the correlations between the image and sound has been perceived as clear. Future work consists of integrating more audio spatialization techniques to volumetric scene synthesis, adaptation of other display types, exploring more interfaces and creating new volumetric image synthesis tools.

ACKNOWLEDGMENTS

The author would like to thank Homer B. Tilton for constructing the first parallactiscope originally used in this project, technical advice on the construction of newer displays, discussions about analog spatial image generation and most of all for his encouragement.

REFERENCES

[1]Tilton, H. B., "Real-Time Direct-Viewed CRT Displays Having Holographic Properties", Proceedings of the Technical Program, Electro-Optical Systems Design Conference - 1971, pp 415-422, 18-20, 1971

[2]Tilton, H. B., "Nineteen-inch parallactiscope", In *Proc. SPIE*, vol. 902, pp 17–23, 1988

[3]Collender, R. B., "The Stereoptiplexer", Information Display, vol. 4, no. 6, pp 27-31, Nov/Dec, 1967

[4]Collender, R.B., "True Stereoscopic Movie System Without Glasses", Part 1, Information Display, vol. 5, no. 4, pp 25-38, July/Aug 1968

[5]Hopkin, B., [Gravikords, Whirlies and Pyrophones], Ellipsis Arts, Roslyn, NY, pp 81, 1996

[6]Youngblood, G., [Expanded Cinema], E.P. Dutton & Co. Inc., New York, pp 256, 1970

[7]Vasulka, W., [Pioneers of Electronic Art], Oberösterreichisches Landmuseum, Linz, 1992

[8]Chowning, J.M., "The Simulation of Moving Sound Sources", J. Audio Eng. Soc., vol. 19, no. 1, pp. 2-6, Jan. 1971

[9]Blesser, B., "An interdisciplinary synthesis of reverberation viewpoints" J. Audio Eng. Soc., vol. 49, no. 10, pp. 867–903, 2001

[10]Roads, C., [The Computer Music Tutorial], MIT Press, Cambridge, MA 1996

[11]Kendall, G., "The Effects of Multi-Channel Signal Decorrelation in Audio Reproduction", Proceedings of the 1994 International Computer Music Conference, San Francisco: ICMA, 1994

[12]Kendall, G., "The Decorrelation of Audio Signals and Its Impact on Spatial Imagery", Computer Music Journal vol. 19, no. 4, pp 71-87, 1995

[13]Lipton, L., [Foundations of The Stereoscopic Cinema], Van Nostrand Reinhold Company, New York , 1982