

# Computer Generated Audiovisuals

*Real Sound in Virtual Space*

February 7, 2006

Maarten Brinkerink

0211834

[m.e.h.brinkerink@students.uu.nl](mailto:m.e.h.brinkerink@students.uu.nl)

<http://maartenbrinkerink.net>

*Get Real!: The Making and Meaning of Computer Generated Imagery*

Ann-Sophie Lehmann

Utrecht University

## Table of Contents

1. <i>Introduction</i> .....	3
1.1 Audiovisual Culture.....	4
1.2 Realisms.....	6
1.3 Virtual Space.....	7
2. <i>Computer Generated Imagery</i> .....	9
2.1 Digital Imagery.....	9
2.2 A Multitude of Visual Cues.....	10
3. <i>Computer Generated Sound</i> .....	11
3.1 Digital Sound.....	11
3.2 Different Types of Sound.....	12
3.3 Spatial Sound.....	13
3.4 A Multitude of Auditory Cues.....	14
4. <i>Audiovisual Virtual Space</i> .....	16
4.1 An Audiovisual Contract in Virtual Space.....	16
4.2 Realism and Virtual Space.....	17
4.3 A Multitude of Audiovisual Cues.....	18
5. <i>Conclusion</i> .....	19
6. <i>Bibliography</i> .....	21

*Our lives are full of an ever-present collage of audio cues that we take for granted.*

(Daniel Bernstein, 1997)

## 1 | Introduction

Today we see computers progressing further and further in creating sophisticated three-dimensional (3D) virtual worlds on our – mostly still just – two-dimensional (2D) screens<sup>1</sup>. These 3D visual models are being created using a multitude of algorithms, which simulate our reality. As described in The Reconfigured Eye – by William J. Mitchell – there are algorithms that simulate various types of perspectives, lighting, shading etc. This *computer generated imagery* (CGI) becomes progressively realistic when different of these simulations are combined creating a multitude of consistent *visual cues*. These cues are based on both laws of nature and cultural convention (Mitchell, 1992).

As early as 1977 computer scientists Alan Kay and Adele Goldberg wrote about the computer as being a *metamedium* in their 'classic' article "Personal Dynamic Media". Because computers work with a digital code they can process a broad spectrum of information, which includes both imagery and sound (Kay and Goldberg, 1977). Thus computers cannot only generate visual cues, but also – among others – *auditory cues*. Today when we speak of 'multimedia computers' we generally mean (consumer) computers that process both imagery and sound.

Computer generated *virtual space* is growing more and more important in our present society. Our lives are progressively becoming 'digital' and are transported into (online) virtual spaces. Therefore, we need to broaden our understanding of the production of these spaces. Continuing to mainly focus on the visual aspects of these spaces narrows the view down to only one of its fundamental aspects. With today's computers capacity to utilize both visual and auditory cues for the production of virtual space this paper aims to answer the following question:

### ***How does sound contribute to the creation of a realistic – computer generated – virtual space?***

This paper will try to answer this question using the following sub-questions:

1. *How do computers generate realistic 3D imagery?*
2. *How do computers generate realistic 3D sound?*
3. *What constitutes a realistic – computer generated – virtual space?*

---

<sup>1</sup> Philips Research is currently experimenting with 3D screens for future consumer applications <[http://www.research.philips.com/technologies/display/ov\\_3ddisp.html](http://www.research.philips.com/technologies/display/ov_3ddisp.html)>.

## 1.1 | *Audiovisual Culture*

Cultural Studies has a tradition of characterizing certain periods of human history in terms of dominance of a specific type of communication, used for the exchange of information. When man first started to communicate, this happened in an – exclusively – *oral culture* (or tradition), where information was handed down verbally from person to person. With the coming of writing systems, society gradually changed into a *literal culture*. Literacy made it possible to record information and for that information to be stored into a (collective) memory. The printing press is considered to be the latest 'turning point'. This invention stimulated that information was increasingly translated into images, creating a *visual culture* in which sight is the dominant sense.

Thinking about culture as being a visual culture – and sight as the superior sense – is still the dominant paradigm in Cultural Studies. Although having arrived in the so-called *digital age*, scholars still apply (or translate) this notion of a visual culture to our current situation. For example, in his article "An Archeology of a Computer Screen" (1995) theorist/artist Lev Manovich argues that we live in a "society of a screen". Although this tradition is based on older art traditions (mainly painting), Manovich sees the emergence of the digital computer as responsible for the current importance of the screen in our society: "Today, coupled with a computer, the screen is rapidly becoming the main means of accessing any kind of information, be it still images, moving images or text." (Manovich, 1995: 1).

Recently there has been a growing resistance – be it on a relatively small scale – to the dominance of the visual within Cultural Studies. Part of this resistance has manifested itself in an *auditory turn*. One of the products of this turning point is [The Auditory Culture Reader](#) (2004), edited by Michael Bull and Les Back. This reader provides various articles that aim to offer an alternative to the notion of visual culture. Their effort is based on the claim that society cannot solely be understood in visual terms: "[...] a visually based epistemology is both insufficient and often erroneous in its descriptor, analyses and thus understanding of the social world." (Bull and Back, 2004: 3).

This alternative *auditory culture* emphasizes the fact that hearing adds a great deal to our perception. It wants to bring into attention the important role of sound that is mostly neglected within Cultural Studies. We don't only see the world around us but we also hear it, hence the superiority of sight ought to be rethought. Both sight and hearing have their specific properties, which provide a unique contribution to our perception. A more nuanced approach to culture should also acknowledge the unique contributions of sound (Bull and Black).

Because the computer is a metamedia, it is perfectly suited for (re)producing both visual and auditory cues and hence audiovisual content. In the current digital age we are progressively being confronted with digital media like (portable) computers, cellular phones, mp3-players and navigation systems in our daily lives. These media are constantly evolving in terms of; smaller size, better performance, bigger memory capacity etc.

These so-called 'new media' follow a trend of being *multimedia* (or *transmedia*). This means that these new forms of media are capable to incorporate – because they utilize digital code – many different forms of media content. As part of this trend most digital new media can address both sight and hearing, they are audiovisual multimedia. The increasing integration of digital audiovisual media in our society undoubtedly has consequences for the nature of our culture. Today's cultural production isn't purely visual, nor auditory, it is an *audiovisual culture*.

In his book Audio-Vision (1994) the French film theorist/composer Michel Chion describes how an audiovisual media expression can be analyzed as an "audiovisual contract". This concept doesn't view image and sound as two independent elements, but both as intermingled in the media expression as a whole. Although Chion specifically applies this concept to film, nevertheless today's society is filled with audiovisual contracts (Chion, 1994).

The (digital) audiovisual media that fill today's society combine the unique properties of both imagery and sound in numerous different ways, for numerous different applications. In these combinations both elements are rarely deployed equally. Sound can be inferior to the image, or the other way around. Nevertheless, in the current (technological) developments and innovation in our cultural production 'pure' – solely visual or auditory media – are becoming exceptional and audiovisual media the norm.

Recently W.J.T. Mitchell<sup>2</sup> has published a rather provocative article in the Journal of Visual Culture titled "There Are No Visual Media" (2005). In this article he argues for a more a more nuanced approach to the specific nature of media. He claims that while the term "visual media" is used to specify many popular media, this is incorrect because all media – according to Mitchell – also involve other senses:

"'Visual media' is a colloquial expression used to designate things such as television, film, photography and painting, etc. But it is highly inexact and misleading. On closer inspection, all the so-called visual media turn out to involve

---

<sup>2</sup> W.J.T. Mitchell is an influential author on the subject of visual culture, not to be confused with the writer of The Reconfigured Eye William J. Mitchell.

other senses (especially touch and hearing). All media are, from the standpoint of sensory modality, 'mixed media'."

(Mitchell, 2005: 257)

Mitchell uses the concept of *mixed media* to call attention to the multi-sensory nature of all media. But he also explicitly states that this nature – that all media have in common – doesn't make it impossible to distinguish between different media. He argues that the specificity of media doesn't depend on what single sense a medium appeals to, but rather the typical mixture of senses and the relation between these senses, what he calls "specific sensory ratios" (Mitchell, 2005: 260-61).

Although Mitchell moves even beyond the notion of an audiovisual understanding of all media because he argues that all media cannot only mix sight and hearing, but all senses (he especially draws attention to touch). Still an audiovisual approach is sufficient for computer generated cultural production, since today's digital media still mainly (re)produce imagery and sound.

Today's audiovisual culture media expressions cannot solely be analyzed from a visual or auditory perspective, because the expression as a whole depends on the way these elements are combined. Cultural Studies should acknowledge the importance of sound and make use of the fact that the auditory aspects of (popular) culture are already being analyzed in disciplines like music theory. An interdisciplinary approach to media expressions could greatly contribute to an audiovisual understanding of culture within Cultural Studies. Nevertheless this still rarely happens.

## 1.2 | *Realisms*

"The term 'realism' is therefore a useful one in this context, no least because it highlights the argument that any representation, however technologically advanced, is a cultural construction and not the 'real' itself. That is to say, a critical notion of realism foregrounds not the 'capture' of the real but its articulation or constitution in representations."

(Lister, 2003: 138)

In an early article called "Computer Imagery" (1989) that addresses the question of realism in CGI, Beverly Jones stresses the fact that these kinds of representations are simulations. Because of this nature Jones connects these representations with art as imitation. Two classic traditions are most commonly associated with this type of art; *Idealism* and *Realism*. Idealism attempts to avoid a "literal representation of physical reality" while Realism does the opposite and seeks "essential or scientific correspondence with physical reality" (Jones, 1989: 32).

In his book New Media (2003) Martin Lister argues that there are several “realisms” that are to be considered when it comes to computer generated representations; *verisimilitude*, *photo-realism* and *hyperrealism*. These three realisms differ in their perspective on reality and hence affect the type of representations based on them.

Verisimilitude imitates visual appearance “as it appears to the human eye” to provide and “unmediated” view of the world”. Photo-realism means representations that look like they have been produced by photographic techniques, but in fact aren’t. Paradoxically this ‘faking’ of a certain technique creates a “reality effect” because the representation because “it conforms to prevailing or emergent realist notions of screen spectacle and fantasy, not the ‘real world’” (Lister, 2003: 140).

Hyperrealism - according to Lister - is used “to identify a distinct and dominant aesthetic in popular animation” (often related to Walt Disney Corporation). This aesthetic is characterized by a “remediation of the codes of live action cinema” but with “exceeded verisimilitude”. Some typical elements are; caricature, exaggeration of the forces of the physical world and plasmaticness. Hyperrealism can also be applied to photo-realism in CGI (Lister, 2003: 141-43).

### 1.3 | *Virtual Space*

“In a virtual world, we are inside an environment of pure information that we can see, hear, and touch. The technology itself is invisible, and carefully adapted to human activity so that we can behave naturally in this artificial world. We can create any imaginable environment and we can experience entirely new perspectives and capabilities within it.”

(Bricken, 1991: 1)

CGI technology is not only used to attempt at creating convincing simulations of our reality, but also to create alternative realities. According to Manovich these *virtual realities* (VR) allow us “to travel through non-existent three-dimensional spaces” (Manovich, 1995: 134). And according to Mitchell VR technology “take us beyond the boundaries of the real world and insert our disembodied viewing presence into modeled fictional worlds” (Mitchell, 1992: 134).

Manovich and Mitchell both emphasize the “non-existent” or “fictional nature” of this alternative space. According to Manovich the opposition of a ‘normal’ space of our reality and a *virtual space* is typical in modern visual culture:

“Visual culture of the modern period, from painting to cinema, is characterized by an intriguing phenomenon: the existence of another virtual space, another three-dimensional world enclosed by a frame and situated inside our normal space.”

(Manovich, 1995: 1)

This paper offers an overview of various techniques used for the production of audiovisual – computer generated – virtual spaces, based on literary research. With this overview this paper tries to bring to attention the mostly neglected role of sound for the creation of a realistic virtual space within our audiovisual culture. In the following “Computer Generated Imagery” offers a short overview of some basic CGI techniques. In “Computer Generated Sound” the basic techniques for (re)producing sound digitally and some important categories of digital sound will be discussed. Finally “Audiovisual Virtual Space” offers an insight in how visual and auditory cues are used to create a realistic – computer generated – virtual space.

## **2 | Computer Generated Imagery**

### *2.1 | Digital Imagery*

Today's sophisticated (3D) CGI has been made possible by the development and combination of several techniques for simulating our visual reality. These techniques are constantly being refined and enhanced with new and more advanced simulation techniques, resulting in even more detailed representations. Jones sums up the essential techniques for the depiction of what she calls "illusory three-dimensional 'reality'", these will briefly be introduced below:

"In the 1960s three-dimensional wire frame graphics were developed. With the consequent development of hidden line algorithms, solid modeling, and lighting and texturing techniques it was possible to attempt depiction of illusory three-dimensional 'reality' in computer graphics."

(Jones, 1989: 34)

Wire frame projection is a technique that translates – mostly rather basic – 3D models into a 2D representation of this model. The depth of space is reduced to the planarity of the image surface, thus creating an illusion of depth. There is an infinite amount of possibilities to achieve this effect, by using various algorithms that have different outcomes and implications. Standard linear perspective projection is the technique that seems most 'natural' to Westerners because of cultural conventions and because of its correspondence with natural optical perception (Mitchell, 1992: 126).

Because pure wire frame projections only simulate the skeleton of a geometric model, further steps are needed to approach a more realistic depiction. A first step in this direction is to portray the opacity of the surfaces of the models, thus making them appear as solid objects. This can be achieved with hidden line removal techniques, algorithms removing the hidden lines that are blocked by the surfaces in front of them in the depth of the picture plane (Mitchell, 1992: 133-134; for more details about this technique see Heger).

To increase the detail of these solid 3D models, simulations of surface detail are being applied. In reality these details convey information about the material nature of objects. For this type of simulation, techniques are being used that describe surfaces in geometrics. This makes it possible to render what a surface of a model looks like using digital technology. This technique is called "texture-rendering". As Mitchell mentions this technique "adds verisimilitude, but it does not report reality" (Mitchell, 1992: 148).

The next step from these – now already rather detailed – solid and materialized models, is applying lighting and shading effects that simulate the light and its effect on the object in a scene. According to Mitchell highlights have several important graphic effects. They provide a scene with “additional dept cues” and “additional information about surface curvature” (Mitchell, 1992: 145). When the simulation of light is combined with synthetic shading Mitchell claims this can lead to “perspective views into closer approximations to – even simulations of – photographs.” (Mitchell, 1992, page 137)

The basic techniques for visually simulating our reality presented above form the fundamentals of today’s sophisticated techniques. Although the practice of creating the current state of the art visual simulations is much more complex then just the sum of these basic techniques, the above does offer an insight in how these simulations are put together. Crucial is the understanding that this involves a multitude of corresponding visual cues – created using algorithms that simulate reality – that add detail to the representations.

Although the techniques mentioned can be used to (re)produce both still and moving images, please note that today’s CGI increasingly are in motion – take for example the breathtaking real-time action of computer games – creating not only a illusion of space, but also of time (and thus movement). And most of all, not only do these images move in space and time, they are often also audiovisual!

## *2.2 | A Multitude of Visual Cues*

Today’s sophisticated (3D) CGI are essentially made possible by the combination of some basic techniques for simulation our reality. Wire frame projections provide models with an illusion of depth. By using hidden line removal these models become solid. By adding surface detail with textures these solid models are further materialized. To further add detail to these images lighting and shading effects can be used. These techniques can be combined (and enhanced with further processing) to form a rather complex image by using a multitude of corresponding visual cues – generated with algorithms that simulate our visual reality – to create a sophisticated visual representation of (virtual) space.

### **3 | Computer Generated Sound**

#### 3.1 | *Digital Sound*

“Sound is produced by vibrations that disturb the air, causing pressure waves that travel out in all directions from the source of the sound. When the waves reach someone’s ear, they set up vibrations that cause electrical signals to be sent to the brain. These electrical signals are perceived as sound.”

(Prince, 1997: 1)

Digital technology makes it possible to translate the pressure we perceive as sound into digital files, hence also making it possible to (re)produce sound with digital media. The digital recording of sound is realized by measuring the pressure caused by a certain sound source at very small intervals. This is how sound designer Bobby Prince explains this procedure in his article “Tricks and Techniques for Sound Effect Design” (1997):

“We take snapshots of the amplitudes of a sound at regular intervals and store the values. We are merely storing values that represent the air pressure of a vibration at any one point in time. So a digital sound file is merely a method of storing air pressure!”

(Prince, 1997: 1)

The standard for recording digital audio was set at a sample rate of 44,100 Hz with a 16-bit resolution (standard compact disk quality or Redbook audio). A sample rate of 44,100 Hz means that the air pressure of a sound source is measured 44,100 times per second (compared this to the 24 frames per second refresh rate that video utilizes). A resolution of 16 bit means that the air pressure that is measured per sample can be given a value from 0 to 65535, offering a broad enough dynamic range for a ‘natural’ result. Variations in sample rate and resolution are possible. Using lower sample rates or resolution will result in poorer sound quality, but will also need substantially less space to store the audio files (Prince, 1997: 1-2).

In the past digital sound design has been constrained by limited computational resources and lack of sufficient storage space resulting in having to use smaller sample rates or resolution, thus having to accept a rather limited sound quality. Today’s more advanced technology however is very much capable of utilizing (multiple tracks) of high quality digital audio.

Apart from the digital recording of real sound sources, sound can also be synthesized using digital technology. In his article “Video Game Music” Matthew Belinkie mentions and important technology for digital sound synthesis, MIDI. MIDI is a general (digital)

language used to communicate with synthesizers<sup>3</sup>, for instructions about the sounds they have to simulate. A big advantage of using synthetic sound is the relative small storage space (compared to Redbook audio) that is needed, because only the instructions for the synthesizers need to be stored (Belinkie, 1999: 3).

Another digital technology for representing sound mentioned by Belinkie is the use of MOD (Digital Module). This technology uses samples of digital audio (small segments of music or sound) placed in a sound library, which can be recalled for playback “on the fly”. This technique takes up more memory and processing power than synthesized sound, but allows for complex *interactive* or *adaptive* sound experiences of high quality (Belinkie, 1999: 3).

Digital technology also makes possible the digital editing and manipulation of various types of digital sound. Even with shareware programs like Cool Edit it is fairly easy to – for example – change the length, speed or pitch of a sound file or apply effects like a reverb or delay to them. Today’s more advanced computer systems can even apply this editing and manipulating of digital sounds in real-time. These options further increase the possibilities for digital sound design and the digital representation of sound. For example, by applying editing and manipulation the sonic properties of an auditory cue, it can be altered to correspond with a certain auditory state needed for the representation as a whole.

### 3.2 | *Different Types of Sound*

Digital sound comes in many different forms that serve different purposes in the digital representations of virtual space. Typically distinctions are made between; *sound effects* (or active/direct sound), *background sound* (or environmental/ambient/zone sound), music and *spatial sound* (Belinkie, 1999; Bernstein, 1997; Bricken, 1991; Eckel; Prince, 1997; Stockburger 2003).

Sound effects serve to denote actions and events within the space of representation. These are often linked to visual objects or events in a virtual space, thus creating a correspondence between the visual and auditory cues in the representation. It can also be used to situate an object outside the field of visual perception, which exists in the auditory space of representation (Stockburger, 2003: 5).

---

<sup>3</sup> Popular synthesizers used in digital representations are – for example – AdLib and SoundBlaster FM by Creative Labs and DirectMusic and DirectSound by Microsoft (Brandon, 2004).

Background sound encompasses all sound that characterizes the specific scene that is represented. These sounds can be constant (think for example of a blowing vent) or at random (a plane breaking the sound barrier). Although we are rarely consciously aware of these background sounds, adding them to a digital representation adds detail and atmosphere to a scene (Prince, 1997). Sound designer Daniel Bernstein describes the 'added value' of background sounds to a specific scene tellingly in his article "Creating an Interactive Audio Environment" (1997):

"Our lives are full of an ever-present collage of audio cues that we take for granted [...] All of these cues, though subtle and seemingly unimportant, create the ambiance of a particular scene, imbuing it with identity and significance. Without these background sounds, or ambiances, our lives would sonically resemble a lunar landscape."

(Bernstein, 1997: 1)

### 3.3 | *Spatial Sound*

"Sound is not necessarily the first thing that come to mind when we think about spatial representation. Nevertheless, it is a significant factor in the emergence of specific immersive environments [...]"

(Stockburger, 2003: 1)

Spatial sound is especially important for the production of a realistic – computer generated – virtual space. Spatial sound encompasses auditory cues that characterize space and the way sound is affected by it, how space alters sonic qualities. Elizabeth M. Wenzel, Joel D. Miller, and Jonathan S. Abel report a growing interest in what they call the "simulation of acoustic environments" – in their article "A Software-Based System for Interactive Spatial Sound Synthesis" (2000) – that is resulting in many technological innovations:

"Interest in the simulation of acoustic environments has prompted a number of technology development efforts for applications such as auralization of concert halls and listening rooms, virtual reality, spatial information displays in aviation, and better sound effects for video games. Each of these applications imply different task requirements or emphasize different aspects of the listening experience that, in turn, require different approaches in the development of rendering software and hardware."

(Wenzel, Miller, and Abel, 2000: 1)

Surround sound is a popular technique<sup>4</sup> that simulates certain types of spatial sound. As Rob Bridgett describes in his article "Hollywood Sound: Part Two" (2005) the concept of

---

<sup>4</sup> Popular consumer surround sound is mostly developed by Dolby Labs, with Dolby Digital 5.1 being the current standard of HDTV and DVD players. Other popular surround sound

surround sound isn't new. However, recent technological developments (especially the success of the DVD format) have caused surround sound to break through to the mass consumer market, making the technology affordable for a wide audience. This makes spatial sound an important emerging (popular) cultural phenomenon (Bridgett, 2005: 1-2).

According to Gerhard Eckel – as described in his article "Applications of the Cyberstage Spatial Sound Server" – three aspects are essential when it comes to spatial sound; localization, reverberation and the radiation pattern of a sound source:

"The first aspect (localization) is important for navigation and orientation and hence contributes mainly to the illusion of presence. The second aspect (reverberation) conveys important information about the environment shared by the sound source and the observer [...] and therefore enhances essentially the degree of immersion. A third aspect concerns the radiation pattern of a sound source. Many sound sources do not emit sound omni-directional but in a potentially complicated and frequency dependent pattern."

(Eckel: 2)

In his article "The Game Environment from an Auditive Perspective" (2003) Axel Stockburger argues that contemporary sound technology is increasingly capable of simulating "the spatial signature of sounds". Especially with the use of (digital) reverb and echo effects the "experience of hearing the sound in a different type of environment" is facilitated with today's digital technology (Stockburger, 2003: 10-11).

### 3.4 | *A Multitude of Auditory Cues*

Digital technology makes it possible to (re)produce the pressure we perceive as sound digitally. Today's technology is very much capable of utilizing high quality audio in digital representations. Apart from the digital recording of real sound sources, sound can also be synthesized using digital technology allowing for a broad array of possible sounds. Sound can also be sampled in real-time allowing for complex interactive or adaptive sound experiences. Digital technology also facilitates (real-time) editing and manipulation of the sonic properties of a sound. This makes possible the altering of the sonic properties of a sonic cue to correspond a certain auditory state needed for the representation as a whole.

---

technologies are THX by Lucasfilm (for movie theatres) and DTS for both theatres and more sophisticated consumer applications (Bridgett, 2005).

Digital sound comes in many different forms that serve different purposes in digital representations. Sound effects serve to denote actions and events within the space of representation and are often linked to visual objects in a virtual space, thus creating a correspondence between the visual and auditory cues. Background sound characterizes a specific scene, adding detail and atmosphere to it. Spatial sound encompasses auditory cues that characterize space and the way sound is affected by it. Contemporary sound technology is increasingly capable of simulating spatial sound, especially with the use of reverb and echo effects.

## **4 | Audiovisual Virtual Space**

### *4.1 | An Audiovisual Contract in Virtual Space*

“Multi-sensory interaction adds additional layers of perception and meaning to our experience.”

(Bricken, 1991: 2)

In her article “Virtual Worlds” (1991) Meredith Bricken explains the importance of multi-sensory cues in the design of virtual spaces using digital VR technology. She argues that the most advanced virtual worlds are “acoustigraphic environments”<sup>5</sup> that couple 3D sound with 3D imagery. The multi-sensory interaction that would result from this coupling adds “additional layers of perception and meaning” (Bricken, 1991).

Eckel argues that imagery and sound “mutually enhance each other by playing complementary roles”. This mutual enhancement of visual and auditory cues contributes to the immersion of the user in a virtual environment, which makes both types of cues equally important. Eckel argues for an audiovisual approach to the design of virtual spaces:

“Therefore, the process generating the visual and auditory stimuli has to take into account the visual and the auditory scene as well as their relationship at a time.”

(Eckel, 1)

Eckel also draws attention to the fact that – in contrast to visual scene simulation – auditory scene simulation is always “event driven”. Sound is always the result of something happening in space and time. This requires that the visual and auditory cues be represented coherently when the event that causes the sound is visualized. They have to be simulated as heard at the same location and at the same time, reflecting the proper sonic characteristics of this specific situation (Eckel, 2).

Bernstein also notes – in relation to computer games, which can be considered as a popular version of immersive virtual space – that sonic cues can help convey complex visual phenomena or even events that aren’t represented visually (Bernstein, 1997: 2).

---

<sup>5</sup> The CyberStage Spatial Sound Server – a CAVE-like audiovisual projection system – that is described by Eckel can be considered as an example of an acoustigraphic environment (Eckel).

## 4.2 | *Realism and Virtual Space*

“Embedded in computer imagery are cultural and historical conventions which affect both aesthetic/artistic and technical scientific formations. In addition, these conventions reflect larger models of cultural reality. Both art and technology are affected by these models of reality.”

(Jones, 1989: 31)

In her article Jones suggests that the simulations and representations of reality made with the aid of computers may be based upon “scientific realism”. These simulations and representations utilize digitizing techniques and simulation algorithms to imitate our reality. This type of realism is related to what she calls “objective realism”. This type of realism considers reality to be structured as consisting of entities with certain properties and relations to other entities. Objective realists believe this structure can be modeled using theoretical models. This is of course exactly what computer generated representations do; imitate reality using algorithmic simulation techniques based on math (Jones, 1989: 34-35).

When we relate the type of computer generated representations that simulate reality to the “realisms” presented by Lister, verisimilitude seems most appropriate to characterize this practice. These simulations try to imitate reality using mathematical models to (re)produce imagery and sound as detailed and ‘natural’ as possible, with a multitude of corresponding audiovisual cues.

However, because digital representations can create alternative realities that are non-existent or fictional, verisimilitude is not the only realism to be reckoned with when it comes to computer generated virtual worlds. Representations can also attempt at photo-realism, reflecting cultural conventions of photography and film (take for example the insertion of CGI elements to filmic scenes). This still involves a multitude of corresponding audiovisual cues, with the difference that an extra element of simulation – that of cultural conventions of photography techniques – is introduced.

Digital representations can also be described as hyperrealism (especially in case of computer games). Because of the non-existent nature of virtual space – that exists purely as information – they are perfectly suited not only for the simulation of reality, but also for the simulation of various types of fiction. Digital simulation can therefore also represent things or situations that are impossible in reality. Exaggerating certain visual or auditory cues can for example do this.

With this type of representations it is not so much a question of how 'real' they are, but to what extent they are capable of creating a convincing experience. Further research in how these multi-sensory representations use visual and auditory cues that immerse the user in a compelling experience, which allows him to suspend his disbelief is therefore required.

#### 4.3 | *A Multitude of Audiovisual Cues*

In sophisticated – computer generated – virtual spaces visual and auditory cues form a digital version of the audiovisual contract as described by Chion. Visual and auditory cues can be considered being equally important to the representation as a whole and can mutually enhance each other. Because sound is always happening in time and space, it is required that events that are represented visually are provided with matching auditory cues, simulated to be heard at the same location and at the same time. Altogether these multi-sensory representations of virtual space can create complex and meaningful experiences.

Because computer generated representations imitate reality using algorithmic simulation techniques based on math they offer scientific objective realism. When the representations are simulating reality they mainly offer verisimilitude (or at least attempt to do so). However, because a virtual space is not confined to reality photorealism and hyperrealism are also possible. It seems more important for a – computer generated – virtual space to be convincing and offer a compelling experience, instead of being 'real' in this respect.

Computer generated imagery should not only be analyzed visually, but also 'audiovisually'. The designer of virtual worlds cannot only use visual cues to denote this world with its events; he also has the possibility of using auditory cues. These auditory cues can coexist and interrelate with visual cues, thus expanding the multitude of possible cues to denote a convincing world. By analyzing both of these cues – not only in their own respect but also their correlation – a deeper understanding of the creation of convincing virtual worlds can be achieved.

## **5 | Conclusion**

This paper offered an overview of various techniques used for the production of audiovisual – computer generated – virtual spaces, based on literary research. With this overview this paper tried to bring to attention the mostly neglected role of sound for the creation of a realistic virtual space within our audiovisual culture.

Today's sophisticated (3D) CGI are essentially made possible by the combination of some basic techniques for simulation our reality. Wire frame projections provide models with an illusion of depth. By using hidden line removal these models become solid. By adding surface detail with textures these solid models are further materialized. To further add detail to these images lighting and shading effects can be used. These techniques can be combined (and enhanced with further processing) to form a rather complex image by using a multitude of corresponding visual cues – generated with algorithms that simulate our visual reality – to create a sophisticated visual representation of (virtual) space.

Digital technology makes it possible to (re)produce the pressure we perceive as sound digitally. Today's technology is very much capable of utilizing high quality audio in digital representations. Apart from the digital recording of real sound sources, sound can also be synthesized using digital technology allowing for a broad array of possible sounds. Sound can also be sampled in real-time allowing for complex interactive or adaptive sound experiences. Digital technology also facilitates (real-time) editing and manipulation of the sonic properties of a sound. This makes possible the altering of the sonic properties of a sonic cue to correspond a certain auditory state needed for the representation as a whole.

Digital sound comes in many different forms that serve different purposes in digital representations. Sound effects serve to denote actions and events within the space of representation and are often linked to visual objects in a virtual space, thus creating a correspondence between the visual and auditory cues. Background sound characterizes a specific scene, adding detail and atmosphere to it. Spatial sound encompasses auditory cues that characterize space and the way sound is affected by it. Contemporary sound technology is increasingly capable of simulating spatial sound, especially with the use of reverb and echo effects.

In sophisticated – computer generated – virtual spaces visual and auditory cues form a digital version of the audiovisual contract as described by Chion. Visual and auditory cues can be considered being equally important to the representation as a whole and can mutually enhance each other. Because sound is always happening in time and space, it is required that events that are represented visually are provided with matching auditory cues, simulated to be heard at the same location and at the same time.

Altogether these multi-sensory representations of virtual space can create complex and meaningful experiences.

Because computer generated representations imitate reality using algorithmic simulation techniques based on math they offer scientific objective realism. When the representations are simulating reality they mainly offer verisimilitude (or at least attempt to do so). However, because a virtual space is not confined to reality photorealism and hyperrealism are also possible. It seems more important for a – computer generated – virtual space to be convincing and offer a compelling experience, instead of being 'real' in this respect.

Computer generated imagery should not only be analyzed visually, but also 'audiovisually'. The designer of virtual worlds cannot only use visual cues to denote this world with its events; he also has the possibility of using auditory cues. These auditory cues can coexist and interrelate with visual cues, thus expanding the multitude of possible cues to denote a convincing world. By analyzing both of these cues – not only in their own respect but also their correlation – a deeper understanding of the creation of convincing virtual worlds can be achieved.

## **6 | Bibliography**

- Belinkie, Matthew. "Video Game Music: Not Just Kid Stuff". 1999.  
<<http://www.vgmusic.com/vgpaper.shtml>>.
- Bernstein, Daniel. "Creating an Interactive Audio Environment." [Gamasutra.com](http://www.gamasutra.com)  
November 14 1997.
- Brandon, Alexander. "Instrumentation: The Game Composer's New Challenge." [Gamasutra.com](http://www.gamasutra.com) April 30 2004.
- Bricken, Meredith. "Virtual Worlds: No Interface to Design". 1991.  
<<http://www.hitl.washington.edu/publications/papers/interface.html>>.
- Bridgett, Rob. "Hollywood Sound: Part Two." [Gamasutra.com](http://www.gamasutra.com) September 30 2005.
- Bull, Michael, and Les Back, eds. The Auditory Culture Reader. 2003. Oxford: Berg, 2004.  
Paperback Edition.
- Chion, Michel. Audio-Vision: Sound on Screen. Trans. Claudia Gorbman. New York:  
Columbia University Press, 1994.
- Eckel, Gerhard. "Applications of the Cyberstage Spatial Sound Server." AES 16th  
International Conference on Spatial Sound Reproduction
- Heger, Walter. "Vector Hidden Line Removal and Fractional Quantitative Invisibility".  
<<http://wheger.tripod.com/vhl/vhl.htm>>.
- Jones, Beverly. "Computer Imagery: Imitation and Representation of Realities." Leonardo: Journal of the International Society for the Arts, Sciences and  
Technology. Computer Art in Context Supplemental Issue (1989): 31-38.
- Kay, Alan, and Adele Goldberg. "Personal Dynamic Media." The New Media Reader.  
1977. Eds. Noah Wardrip-Fruin and Nick Montfort. Boston: The MIT Press, 2003.  
391-404.
- Lister, Martin, ed. New Media: A Critical Introduction. London: Routledge, 2003.

Manovich, Lev. "An Archeology of a Computer Screen". 1995.  
<[http://www.manovich.net/TEXT/digital\\_nature.html](http://www.manovich.net/TEXT/digital_nature.html)>.

Mitchell, William J. The Reconfigured Eye: Visual Truth in the Post-Photographic Era. 1992. Cambridge: The MIT Press, 2001. Fourth printing.

Mitchell, W.J.T. "There Are No Visual Media." Journal of Visual Culture 4.2 (2005): 257-66.

Prince, Bobby. "Tricks and Techniques for Sound Effect Design." Gamasutra.com June 1 1997.

Stockburger, Axel. "The Game Environment from an Auditive Perspective." Level Up: Digital Games Research Conference. Eds. Marinka Copier and Joost Raessens. Utrecht: Utrecht University, 2003.

Wenzel, Elizabeth M., Joel D. Miller, and Jonathan S. Abel. "A Software-Based System for Interactive Spatial Sound Synthesis." ICAD 2000. Atlanta, GA, 2000.