Centre for Geo-Information

Thesis Report GRS-2004-3

DISCOVERING THE SOUNDSCAPE IN A VIRTUAL 3D ENVIRONMENT

FEBRUARY 2004

VASILEIOS ROUMELIOTIS

WAGENINGEN UNIVERSITY

WAGENINGEN UR

Dedicated to Nikos, Fofo

Katerina and Hans

Preface

This book is my thesis work for obtaining the Master of Science degree on Geoinformation Science at Wageningen University, Alterra Department, and the Netherlands.
The basic and most important part of this work is based on the interpretation of the conceptual mode; of the thesis project, through a demo application by using behavioural software.

I would like to thank all these people who trusted and helped me during my work and gave some of their time to share their experiences and knowledge with me.
Special thanks to Peter Graf, for his significant contribution into the applicative part of the thesis project and his immediate on line support.
The contribution of my supervisor Ron van Lammeren to my thesis has been very important, not only for stimulating and sharing with me valuable scientific material and experience, but also for being a founded, positive and creative person. His initiative to order and provide me with Virtools developing software was meaningful for my work but it is also important for the students that will work in on a similar project in the future.

Finally, many thanks to my family for their entire support, not only during my thesis work but through the whole Msc course in Wageningen University.

INDEX

CHAPTER 5

Conceptual model

CHAPTER 6

Virtual soundscape implementation

CHAPTER 7

Remarks and conclusions

CHAPTER 1

Introduction

1.1 Context and background

It is true that an image counts just as a thousand words. And this has almost become a rule in the virtual reality world. The rapid enhancement of images and graphics details, are also accompanied by digital processed sound effects, see the developments in the virtual games environment.

Designers of virtual environments seek to create a feeling of presence, or immersion for users. Primarily, this goal has been pursued by creating worlds with convincing 3D interactive graphics. This sense of presence is achieved by chiefly engaging the human sense of sight. Unlike sight, the sense of hearing is often neglected in the implementation of a virtual world. Regardless of considerable evidence on its immersive potential, audio is often banished as the poor stepchild of virtual reality. This trend is in part due to technical resource limitations of computer systems. Designers were forced to sacrifice audio quality for graphics performance, immersive audio in graphically intensive virtual 3D environments.

In the domain of geoinformation and navigation in virtual 3D space however, the evolution of visual representation, is not followed by a respective involution of sound. The majority of geoinformative applications, although facilitating rich image resolutions and quick user interaction, present inefficiency when more sophisticated soundscapes are required.

The presence of sound in this kind of virtual application is perhaps one of the most depreciated parameters, mainly on account of the weak sequence between image and sound. Indeed, looking at several navigations through virtual landscapes on the websites, the lack of sound support is evident. Discriminating some applications with generous attempts to enrich the user’s navigation with some audio files, the virtual soundscape is generally characterised by a poor diversity. No matter if the users zooms in or out in the space and regardless of the transition through different thematic landscapes, the sound intensity and other attributes of sound remain identical. There is no dynamic sequence of sounds to supervene the progression of images.

Regarding the number of potential applications of sound in virtual reality-like orientational sound tips, navigational sound assistance or acoustic information when passing a certain thresholds-   it could help increase the sense of presence in virtual environments by relaying information about the environment and the objects within it. Such environmental awareness could be very beneficial in increasing the user's topological awareness in virtual environments. (C. Tonnesen, J. Steinmetz/1993)

1.2 Problem definition

A soundscape is characterised dynamic when comprise a number of static and variable (periodical or movable) sounds, adaptable to the static or moving visual field of the viewer, rendering more realistic sound. A dynamic soundscape must be created that can respond appropriately to any change caused by the user or the environment itself. (Jarrell Pair, URL1.3)

The creation of a dynamic soundscape and the interconnectivity with the graphics of the 3D virtual environment is the basic concepts involved in this research. Besides,

any ‘reality’ in which one could not operate equally autonomously within a meaningful acoustic space, would be even less virtual. (Richard Thorn/1996)
Spatial sound is at the point now of a having sufficient technological tools to solve most of the problems, but lacking a consensus on how to employ these tools for optimum results. (Kimmo Vennonen/1996)
Despite the suppleness and capabilities of the technological tools, the lack of more effective and integrated sound coverage accompanying the visualisation of the geometric and thematic geodatasets is evident: limited initiatives for more realistic rendered soundscapes in geoinformation visualisation, contrary to the videogame environment where sound records rapid developments.

1.3 Research objective

The main challenge of this research is the extension of geo-information-that come visualised through different levels of optical resolution- by sound data, leading to a representation of geo-information via sound-scapes.

Research questions

To be able to reach this objective, a number of conceptual, formal (structural) and implemental character questions need to be answered.

Q1. Is the 3D visual space composed from the scratch or it is imported as an end product?
Q2. Will the visualscape be further processed in order to render a more realistic impression of the real world?
The third research question is related to the geoinformative character of sound:
Q3. How is accomplished the extension of (visual) geoinformation by sounds, considering the abstract hypostasis of sound respect to image?

Proceeding with questions of more technical character, the association between sound and image is the dominating issue, whereunder a number of latent sub questions need an answer:

Q4. What kind of structure has the soundscape?
Q5. In which natural rules do the captured sounds obey?
Q6. How do the natural properties and general characteristics of sound come technically translated into the soundscape, in order to render a realistic impression of sound?
Q7. How the sound does follow the different resolution optical levels of the visualscape (image)?
Q8. Which structural approach is followed for the allocation of sounds within the virtual soundscape: raster or object oriented approach?

Along with the primary objective, a secondary one is the presentation of results by creating the feeling of immersion in the 3D virtual environment for the viewer. The questions rising in this point are:
Q9. What kind of visualisation tool is going to be used?
Q10. What type of user’s interface will be applied (interactive/ non interactive).
Proceeding with the implemental part of the project, the issue that dominates is
Q11. What type of software is going to be used for the creation of the soundscape and its association to the image?

A final point to be unravelled has to do with the same components that constitute the soundscape:

Q12. Are there going to be used 'store-bought' sounds from a sound library or there will be registered in nature?

Q13. Do sounds comprise a part of Geodata and do they behave as objects or features of the 3D space?

Q14. Is it possible to georeference the sounds and if it is feasible how this is achieved?

Q15. In which file format will the sounds be imported in the application?

1.4 Methodology

Continuing with the methodological part of the project, the first priority is to make understandable the way through which all the predefined objective questions will be answered.

(Q1) Focusing on the starting point of the project, a clarification of which parts will be developed from the scratch and which ones are 'store-bought', will determinate the time consumed for development of the corresponding soundscape. Considering the limitation of time from one side, and the focus on a more sophisticated soundscape from the other, the use of an 'end product visualscape' is the most appropriate solution.

(Q2) In addition, a further processing of the components that constitute the visualscape, by adding extra objects and animations, will improve the 3D and synchronously enrich the soundscape with more 'dynamic' sounds.

Although this project focuses on the dynamic presence of sound, the visualscape doesn't fall short of those characteristics that constitute an integrated virtual environment: the selection is based on such criteria that will accommodate a facile representation of geoinformation through sounds.

(Q3)Beginning with a brief description of the visualscape helps to familiarize with the geographic components of the 3D environment and to realize the variety of sounds to be used. Besides, the visual Level of Detail used for the visualization of the 3d world, will determine the Level of Detail that characterizes the corresponding soundscape (according always to the required acoustic accuracy).

The visualization and exploration of the 3D environment is accomplished through a VRML viewer that allows the exploration of the 3D space through different resolution levels of view (zooms).

(Q3) The extension of geoinformation by sounds is achieved by using sounds of physical or artificial nature. Sounds of physical nature, like the sound from a river or even the noise from a highway, are a kind of geoinformation, though with more abstract sense: the combination of sound and image serve the better perceptibility of the 3D space.

(Q4, Q5) For a better understanding of the conjunction between image and sound, it's wise to make a reference to the natural rules that the components of the soundscape obey, as well as the way this is interpreted (during the formation of the soundscape) through a media software.

The 3D virtual soundscape is a mixture of various discrete sounds (like a vehicle's horn) and more generic sounds (like the sough of the leaves). In particular circumstances however, some of the generic sounds could work as individual sounds and reversely. It is also possible to use the same type of sound to express either individual or generic character at the same time. This has mostly to do with the zooming detailed of the particular object /area under question. (The general sound of woodnote for example, when navigating over a forest area would abut in a particular woodnote after zooming in a particular bird on a branch. In this example, the various woodnotes generated by different species and considering the continuous movement

of birds induct an important parameter, basic for a successful sound realistic rendering: the randomisation of sound).
Working on the randomisation parameter, involves separating the sounds in 'static' and 'dynamic' (or 'movable') as well as a further discrimination between 'continuous' and 'recurrent' (or 'discrete') sounds.
A set of key auditory variables, like location and duration of a sound, loudness, order (sequence) of sounds and the pitch (frequency), introduced by Krygier (1994), will help to obtain an hierarchy of the sound and pose them properly to the soundscape.The separation between static and dynamic sounds is strongly related to the nature of characterised objects (immotile and mobile objects respectively).
The second discrimination has to do with permanent sounds that we meet in a soundscape (e.g. sound of a river) and discrete or periodical sounds (like a car's engine or the leading voice for assisting navigation).
In addition, sound diversity is another important parameter that will determine the variety of sounds to be comprised within the virtual space.
(Q6, Q7) The maximum level of detail (achieved during the zooming procedure), in combination with the geometrical detail of the objects, will define the range of sound diversity and intensity. In other words, there must be a logical consistency between the visual and the acoustic Level of Detail (LOD). The latter however is much more abstract and difficult to define, since it involves 'background' sounds from areas and objects that can't even be seen from the user's selected field of view. It is the difference between the 'visual background' (referring to the permanent visible backdrop of the picture) and the 'acoustic background, which is something that hovers in the space or even behind the observer.
Generally, the aim is to define a dynamic relation between the geographic entities and sounds. One of the greatest constraints to achieve this aim is the main difference between the two basic human's senses: while vision gathers a large amount of information in only one direction at a time, auditory perception is omni directional and as human species we are good at ascribing casual meanings to sound((Ballas 1993).
Talking in more technical terms, all the parameters describing a soundscape are related to the Level of Detail of the current (zoomed) image-view. Consequently the number of LOD used for the multiresolution images 'follows' the number of LOD of the soundscape. The different LOD structured layers used for the formation of the soundscape is not always identical to the number of the multiresolution picture format. According to the demanding simulative requirements, it can be of lesser, equal or higher number of LOD compared to the image LOD.
To simplify the associative interaction between the acoustic and visual structural layers of the project, there is an identical (1:1) one to one relation (or correspondence) between every visual and acoustic LOD.
In particular, assisted by the VRML viewer for visualisation of the virtual space, there will be selected three different optical zooms, of different resolutions which will be further represented through sounds of analogue resolution. The transition between the different acoustic levels of details follows the devolvement during the optical zooming (in/out).
(Q8) The virtual soundscape is separated in a 3 layer grid structure. All three grids are of different dimensions and are consisted of a different number of cells depending on the resolution of the current view. The 'overlaying' grid structure will accommodate the allocation of the sound features on the corresponding cell of the 'squared' geographical area.
(Q13, Q14) The right positioning of sounds in the 3D space is also related to the georeferency of geographical entities. Although sound stands for an abstract constitution, it can be accepted as a geographical entity (geodata), owing mainly to the importance of the geoinformation that provides. Particularly, sound is considered as a characteristic of several geographical entities or physical phenomena (sound

sources). Georeferency of sound is a rather sophisticated issue, depending on the 'nature' of the sound source but also from the technique that will be followed to georeference a certain sound.
(Q9) After completing the composition of the soundscape and associating it to the geographic entities, an additional objective is to discover this relation between sound and image by creating a feeling of immersion in the 3D environment. The assistance of a VRML viewer (as already mentioned above) is the most appropriate visualisation tool: a walk through the 3D virtual space, will give the opportunity to experiment with the various components of the soundscape, to explore their coherent change and the way they are related to the respective visualscape.

1.5 Feasibility

Trying to achieve optimal acoustic results within a virtual 3D visualscape, it's necessary to overcome any technical constraint, by working on 'neutral software', compatible with most of the pre-exported data formats.
(Q11) Virtools behavioural software, which is also very popular in the videogame design industry, is the most suitable solution for the implementing part of the thesis project. It can be used either as a VRML browser, or VRML player, allowing the composer to intervene dynamically into the 3D visual space with without requiring any special knowledge in computer language programming.
(Q2) Starting from the formation of the visualscape, the aforementioned 'Virtools developing package', disposes various object libraries -which can be even enriched by 3D objects from several websites- that can be used in order to save time for the construction of the virtual 3D environment.
(Q8) The three –different resolution- visual levels are achieved by using three sequential zooms from the same perspective view of the VRML (Virtools) viewer (URL 4.13). The same view is constantly following a human - 'avatar' during its walk through in the 3D space. Apart from the three resolution levels (that allow the easier discrimination of acoustic levels of detail respect to the current view), there is also a free navigation mode, liberated from the avatar's movements.
(Q12, Q15) All the sounds to be used in the application will be imported from sound libraries found in Virtools resources and in the web. This is a very convenient solution; that misses (from one side) the opportunity to experience and register sounds in nature, but from the other side saves a lot of time and effort. Virtools supports most of the 'commercial' sound formats among which 'wav' and 'mp3' (that saves a lot of space respect to the other sound formats).
(Q6) The predefined parameters, like sound diversity, sound intensity (loudness) and 'randomisation' of sound comprises the key for the appropriate allocation of sounds into the three raster sound layers. Sounds of generic nature (e.g. sough of leaves) are assigned to the lowest detail layer while more specific sounds are comprised in the 3rd layer of the maximum sound resolution. Some of the general character sounds contained in the first layer, are also comprised in the second and third layer, with extenuated loudness however to allow the hearing of the discrete nature sounds. Examining each layer separately, it is observed that the number of sounds linked in every raster block (or cell) is based on the variety of image features bounded in a single block.
(Q10) Moving on the virtual environment is one of the most critical points, considering the level of interactivity with the user and the information reception during the navigation. Allowing a free navigation in the virtual environment, the user has the opportunity to experience with all kind of sounds in every point of the 3D space.

The first four of the following chapters, are dedicated to the introduction of sound concepts and description of the methodological approach, baring the fourth chapter which his mainly referred to the technological evolution of sound. The 'feasibility' of the thesis project, which was sententiously reported in the paragraphs above, is analytically described in the sixth chapter.

CHAPTER 2

Introduction to sounds

Introduction:

In this chapter, an introduction to sound is attempted through given definitions and general properties of sound. Some of the sound concepts and characteristics of acoustic spaces met in this introduction will be used in the following chapters. A part of this chapter is dedicated to the human way of perceiving sounds and to the sound natural properties that is possible to simulate through a demo application. In addition, a further discussion is made regarding the natural rules on which the sounds obey as already noted in one of the research questions.

2.1 sound concepts

Starting with more basic concepts, like sound and its relevant properties, will accommodate the definition of more complex ‘acoustic’ compositions such as soundscape (in the following chapter.) Sound is a compression waveform that moves through air or other materials, having the standard characteristics of any waveform. From the physical point of view, sound consists of the transmission of vibrations generated by an oscillating body (sound source) through an "elastic" means, and the reception on the part of a sensor or receiver. Those vibrations of the sound source are translated into variations in the pressure and propagation speed of the air molecules, disturbances which, if particular prerequisites are present, can be perceived by the human ear. (Ron Kurtus/2002)
From humans’ perspective, it is the acoustic sensation that living beings equipped with the auditory sense experience when the hearing organ is struck by a sound stimulus.
Some of the basic natural properties that characterize sound are frequency (and period), amplitude, intensity, velocity, wavelength and phase. (URL 2.2)
Frequency is the number of complete waves or oscillations or cycles of a periodic quantity occurring in unit time (usually 1 second). Amplitude is the amount of pressure change generated by the motion of a vibrating object. The amplitude of a sound wave is the same thing as its loudness. Since sound is a compression wave, its loudness or amplitude would correspond to how much the wave is compressed.
Another important term, the intensity, refers to power and represents the amount of energy (expended in a second) measured over a particular area (m^2 or cm^2). Amplitude and intensity are both physical measurements (Intensity = the square of amplitude). Loudness however, is the perceptual counterpart: the greater the amplitude and intensity, the louder the perceived sound.
Velocity shows how fast a sound wave moves depending on density, temperature and elastic properties of the medium through which it is moving. It shouldn’t be confused however with speed of sound, since velocity usually includes direction of travel. Wavelength is another property of sound that represents the distance from one crest to another of a wave (distance between maximum compressions). Since the velocity of sound is approximately the same for all wavelengths, frequency is often used to better describe the effects of the different wavelengths. Finally the phase shows the relative timing of areas of high and low pressure. (URL 2.2)

2.2 experiencing sounds

- The effect of materials on sound

Proceeding with the natural properties and general characteristics of sound, there is a series of properties presented when other materials and objects affect sound. In particular a sound wave can be reflected, absorbed, transmitted, refracted or diffracted when it comes into contact with an object.
Refraction is caused when a waveform is transmitted or passes from one material to another, where the velocity is different in each material. On the contrary, diffraction is the slight bending of a waveform when it passes an edge.
Some materials absorb so much of the sound that it may be transmitted through the material for only a very short distance. Some sounds can even affect other sounds, like cancelling them out, as is seen in noise cancellation devices.

Geometric characteristics such as the size (or distance), the shape, the direction (orientation) and the topology of the sound source, respect to the listener, determine the way that sound is perceived. Especially the detection of direction is determined by comparing the sound heard by each ear. Detecting the distance however, is more difficult and often relates to loudness and quality of the sound heard. The mass of the object and the vibration velocity, are two more variables that determine the **quality** – therefore the perceptibility- of the sound produced. The quality of a sound is mainly defined by the pitch (high/low), the intensity (weak/strong), and the timbre (soft, clear, hollow, etc.). The latter due to different "harmonic" component of the sound itself, allows humans to distinguish sensations that are equal in pitch and intensity; that is for example, to distinguish one musical instrument from another or one voice from another. The pitch or tone of a sound wave is determined by its frequency, which is the wavelength divided by the speed of sound. The sound you hear consists of different frequencies or wavelengths, which determine their pitch. The amplitude of a sound wave determines its loudness. Humans and animals can only hear within a limited range of pitch or frequencies, depending on the species. This limitation affects their perception of the world around you, since there are sounds that you or animals can't hear that others can. (Ron Kurtus/2002)

When two or more sound waves are involved, then a constructive or destructive interference may appear: the former results in an increase of the wave’s amplitude while the second in a decrease in the amplitude of the resulting wave.
Generally, sound can be even separated in two categories: the periodic sound characterized by a waveform which has the same shape from period to period and aperiodic sound characterized by irregular and non repetitive forms.
Obstruction and occlusion represent two accessional characteristics of sound related to the geometry of the space: Obstruction appears when the sound is reflected off surfaces or other objects, and is not heard directly. Occlusion on the other side appears when the sound is muffled depending on the composition of the surface the sound goes through as well as the listener’s proximity to it.

-**Human perceptibility and experiencing mode**

Beside the natural properties of sound, there are also some particular characteristics, related to human’s perceptibility of sounds, and whose appearance depends on the way that humans experience the soundscape. Some of the parameters that could influence the human perception of sound could be their moving state (velocity), the surrounding space or the kind of sound source. The Doppler Effect is a representative example of the first case that refers to the change of the sound pitch according to the relative velocity between the sound source and the receiver. Harmonics is another characteristic, related to the listener’s surrounding environment

and is proportional to resonance. (URL 2.2) The position and quality of the source that the sound is originating from, can sensibly affect the listener's perceptibility in a high degree (e.g. natural sound versus sound heard through a stereo or 3D surround system.
Echo is a further characteristic-of human's perception of sound-that occurs when the incident wave vibrates your eardrum, but the eardrum has time to stop vibrating before it is set into vibration again by the reflected wave. (Ron Kurtus/2002)

Audible sound waves are measured in decibels, with the softest sounds that people can hear being around 10 decibels. Sounds become painful to the human ear around 120 decibels, but the human's endocrine and nervous system starts responding to noises at levels as low as 70 decibels.
The decibel scale (dB) is special designed to measure sounds that takes into account the amplitudes and intensities of sounds in relation to how we perceive sounds. It is notable that humans can hear around 1 trillion intensities. (URL 2.2)

'Human beings are usually very good at building mental models of their environments. In familiar surroundings, like at work or at home, we often have a fairly accurate picture of the activities in our immediate vicinity. We infer the doings of our colleagues or family members, even though we cannot observe them directly.
Sound plays an important role in the creation of such models, especially in an emotional way since it carries a number of subtle and familiar cues: they can give different feelings (cheerful, moaning, melancholic, anxious etc.) depending on things or events that are mentally connected with. Less obtrusive and remote sounds also play an important role. People are usually not aware of the hum of the ventilation system or the distant din of conversation, until sounds suddenly change in intensity or their name is mentioned. (URL 2.8)

Apart from the way that humans experience and perceive the sounds in the real world, it is also wise to examine the way that humans perceive the 'digitally created' multimedia models that represent 'virtually' the real world. Obviously, the structure and the methodological approach, followed for the formation of these digital models will mainly influence the perceptibility of the user, as it is further analyzed in the fifth chapter. Nevertheless, since the digitally created models are presented through digital -or analogue- electro-acoustic devices, the successful rendering of the real world sounds depends also on the available technological means and techniques for digital reproduction of sound, as it is later reported in the fourth chapter.

In the previous paragraphs a number of physical characteristics of sound have been presented, regarding the sound as a physical attribute (waveform) but also as a way that human's perceive the reality (sense of hearing).
When trying however to simulate the natural characteristics of sound in a virtual 'sound application', it would be impossible to involve all the aforementioned attributes.
There must be a wise selection of natural properties and general characteristics based on the simulative capabilities of the software but also depending on the humans' sensitivity and perceptibility of sounds through the 'sonic' application.

2.3 simulating sounds

Therefore, using a number of properties of sound and making a series of discriminations in the components of the acoustic space it is possible to facilitate the design of the virtual soundscape.
Natural properties like intensity (loudness) and frequency can be easily simulated by media software. In addition, the separation of sounds in permanent-discrete and

static- dynamic will also help the categorization of the sounds in the real world and facilitate their association to the graphic 3D environment.
Sound diversity, sequence of sound (order), randomization of sound and periodical sounds, are four **variables** imported for the easier assessment of the registered sounds in the virtual soundscape.
Sound diversity describes the number of different sounds (variety) within the area of research.
Randomization refers to the frequency that every recorded sound appears in the certain point (raster block) of the area. Periodical sounds are the sounds that appear in a certain period of time during a day: this facilitates the construction of an (appearing) 'loop' in the application template. Finally the sequence of sounds is a parameter that shows the sequence of appearance of the sounds in a certain area (facilitating the prioritizing of the recorded sounds).

2.4 conclusions

In the previous paragraphs, a general reference was made on the natural properties of sound, and the influence of different factors on sound perceptibility. In addition, a quick anaphora was made on the human's prehension of sounds depending on the way they experience them.
Consequently, a prime acquaintance with sound was attempted so far that constitutes an exordium for further introduction to more complex acoustic compositions and concepts, as they are described in the following chapters.

CHAPTER 3

Geoinformation and soundscapes

Introduction:

Having embedded some important properties of sound, it is time to introduce more complex sound concepts and acoustic compositions such as a soundscape. Apart from presenting soundscape concepts and the way humans perceive and interact with soundscapes, aim of this chapter is to prove the geoinformative character of sound, introducing soundscapes as an accessional component of geodata. Along with the geoinformative character of the soundscapes, some relevant issues are illuminated, related to georeferency of sounds and their consideration as ‘geodata’.

3.1 Soundscape: another representation of geo-information

When visualising geoinformation, it is observed a minimal presence of sound in 3D maps and an even complete absence of it, regarding 2D maps, graphs, tables and other ways of visualisation. In geoinformation, just as in day-to-day life, the importance of images is higher than sounds. If someone had to choose between eyesight and hearing, it is apparent the preference of the former human’s sense. Sound therefore should be an integrated part of a multi-medial representation of the real world in such a way that even visual disabled people could create a mental map (of the real world) based only through geoinformation via sounds.

One of the single most important resources of the natural world is its voice or **natural soundscape**. But what is actually a soundscape?
“Soundscape is an environment of sound (sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society. It thus depends on the relationship between the individual and any such environment. The term may refer to actual environments, or to abstract constructions such as musical compositions and tape montages, particularly when considered as an artificial environment.” (Matthew Gordon, URL 3.5)

‘Soundscape refers to any acoustic environment whether natural, urban, or rural. In undisturbed natural environments, creatures vocalize in relationship to one another. This system has evolved in a manner so that each voice can be heard distinctly and each creature can thrive as much through its iteration as any other aspect of its being. The introduction of noise into natural soundscape enhances the sense of loss because noise diminishes human experience of the wild. Creature behaviour is altered as a direct result of increased stress. Keeping in mind that human and non-human species respond differently to types, loudness, or combinations of mechanical noises, it becomes perceivable that many of these sounds introduce affliction in both worlds even though the victims may not seem conscious of the effect.’ (Bernie Krause/2001)
Consequently, the term ‘soundscape’ could have both positive and negative sense, considering that soundscape components can have a various effect in the temperament and behaviour of the living beings.

A soundscape can be both perceived as a model of real world sounds but also as a subset of the real world’s landscape. As a subset, it contains all the sound features and attributes that characterize all geographical entities, objects, phenomena and conditions of the visualscape. As a model, it constitutes a description of the spatial status, natural properties and interactivity rules for every sound component of the soundscape. The term ‘soundscape’ however, it doesn’t refer only to the real world but it is extended to include the representations of the real world as well, through the ‘virtual soundscapes’.

Soundscapes in natural environments are predominately composed of natural sounds produced by animals, birds, insects and other biotic sounds. Also present are abiotic sounds such as thunderstorms, rain, rushing rivers, and wind. In urban environments, the natural sounds are limited and the soundscape is dominated by man-made sounds, either biotic (voices) or abiotic (artificial sounds produced by machines or generally human activities. (Hessische Rundfunk/1997)

3.2 Creating soundscapes

There is a series of relative terms that someone can meet very often when involved with soundscape terminology, such as **soundscape design: ‘**a new interdisciplinary attempting to discover principles and to develop techniques by which the social, psychological and aesthetic quality of the acoustic environment or soundscape may be improved. The techniques of soundscape design include the elimination or restriction of certain sounds (e.g. noise) the evaluation of new sounds before they are introduced indiscriminately into the environment, as well as the preservation of certain sounds (sound marks) and above all the imaginative combination and balancing of sounds to create attractive acoustic environments.’ (Matthew Gordon, URL 3.5)
In addition, the study of the systematic relationships between man and sonic environments is called soundscape ecology, whereas the creation, improvement or modelling of any such environment is a matter of soundscape design. ‘Soundscape ecology is the study of the effects of the acoustic environment, or soundscape on the physical responses or behavioural characteristics of those living within it.’ (Matthew Gordon, URL 3.5)

Another basic concept related to soundscapes, is the **soundscape composition**: a form of electro acoustic music, characterized by the presence of recognizable environmental sounds and contexts, aiming to invoke the listener's associations, memories, and imagination related to the soundscape.’ (2) 'Framing' environmental sound by taking it out of context, where often it is ignored, and directing the listener's attention to it in a publication or public presentation, meant that the compositional technique involved was minimal, involving only selection and transparent editing. Others work use transformations of environmental sounds through analogue and digital studio techniques, with an inevitable increase in the level of abstraction. However, the intent is always to reveal a deeper level of signification inherent within the sound and to invoke the listener's semantic associations without obliterating the sound's reconcilability. (2)
Electro acoustic technology has profoundly changed not only the soundscape, but also the individual practice of listening and the social behaviour it leads to: in a visually dominant culture saturated with technologically based sound, much of our daily aural experience includes sound that is reproduced through loudspeakers. It can come from hidden sources or obvious ones, hugely powerful devices or personal attachments that speak only to our own ears. Electroacoustic sound is in the environment, but not of that environment, we cannot react to these imaginary sources as we could to actual people in the environment. Special care, however, needs to be exercised with media use of reproduced sounds, as the lack of the full context in the sound alone, and the listener's potential lack of knowledge about that

context (that is, how to interpret the sounds), must be compensated by appropriate commentary and guidance. (Hessische Rundfunk/1997)
‘Since a soundscape is shaped by both the conscious and subliminal perceptions of the listener, soundscape analysis is based on perceptual and cognitive attributes such as foreground, background, contour, rhythm, silence, density, space and volume, from which are derived such analytical concepts as soundmark, acoustic space and soundwalk. Soundwalk is a form of active participation in the soundscape. Though the variations are many, the essential purpose of the soundwalk is to encourage the participant to listen discriminatively, and moreover, to make critical judgments about the sounds heard and their contribution to the balance or imbalance of the sonic environment. (Matthew Gordon, URL 3.5)

3.3 Soundscape: another form of geo-information

Making a brief analysis on the aforementioned information about soundscapes and relevant concepts, it becomes obvious the geoinformative character that accompanies every soundscape composition: artificial or natural composed a soundscape constitutes a basic part of the geodata that characterises a certain area.
When referring to ‘geodata’, it is usually intelligible that sound is an irrelevant component of this concept.
The term geodata includes all the ‘geo’ and temporal referenced data about real world phenomena, structured via feature classes in a thematic (or semantic) domain and related geometric domain. Whether sounds constitute part of Geodata or not is obviously a very critical question to answer. It is important therefore to explain what kind of data is considered ‘geodata’ and what is not. Since geodata or geographical data are all georeferenced features
Substantially when talking about georeferency, moving entities (e.g. vehicles) must be excluded from the ‘list’ of geographical features. Abstract entities moving within certain boundaries (e.g. a river or sea) however, despite their moving state are possible to georeference.
Talking in absolute terms, sounds cannot be considered geographical objects (entities) since they appear as sound attributes of entities phenomena or even conditions. ‘Georeferency’ however, cannot constitute an absolute criterion for entitling entities as “geodata” especially for features of such a generic nature, as sound.
If for some of the visible entities is not possible to be georeferenced or temporal referenced (e.g. a car or an airplane) for the sounds is even harder. This is mostly because sounds don’t belong only to visible entities but emanate also from viewless entities- that cannot be visualized in a virtual 3D world. Or they can be a sound product of various natural phenomena(either visible such as the rain or invisible like a thunder) where it makes no sense to discuss about georeferency, but still it gives a lot of information about the (weather) conditions in a certain area.
In other cases however, sounds are possible to be georeferenced using withal different techniques: a **first approach** is based on extensive field work and consists in measuring (approximately) the buffer zone around the sound source (influential area) and georeferencing the buffer zone. An easier **approach (2nd)** could be based on the (approximate) detection and georeferencing of the same source that produces the sound.
Implicitly, sounds indeed constitute a part of geodata, since they can provide with a lot of information about a geographical area, and be ‘in a way’ georeferenced.
An example of the first approach could represent a buffer zone along a river (only for the sound of flowing water) or a buffer zone surrounding a forest (regarding the sounds of singing birds). A representative example for the second case could be the

georeferency of a volcano's caldera, as the epicentre of the emanating volcano sound. There is also a **third approach** to accomplish sound georeferency, by recording the sound 'ad loc' using a GPS receiver in every point of sound recording.
However there are two other parameters of sound such as the vertical range (propagation and effect in the vertical direction) and the duration, which are not strongly related to georeferency but still provide information about the soundscape of a geographical area.
Apparently, it is this 'geoinformative character' of sound that engages the core of this project: the extra information that someone gets when immersing in a 3D graphic environment and the contribution of sound into the perception of the real world's phenomena and conditions.

So far, some further definitions and main characteristics of the soundscapes in the real world have been given. The distinction 'real world' is used for discrimination from the virtual soundscapes. The name 'virtual' is used to determine the role of the soundscape developed by this project: it is virtual since it doesn't correspond directly to the real world but constitutes a representation of it through an association to a 3D graphic environment.

3.5 From the 'real' to the 'virtual' soundscape

Regarding the **constructive rules** of the 'virtual soundscape' –meaning all the physical attributes, structural rules and criteria to be taken into account for the simulation of a 'real world's soundscape' by a virtual one- , it's important first a profound understanding of the natural laws that rule a soundscape in the real world, and their transmission to the 'virtual' one (which is linked to an artificial 3D graphical environment).

The great range of sounds that someone can meet in a soundscape, apart from the aforementioned parameters, varies also according to the chronic moment of the day: examining either a natural or urban environment, there is a succession of different acoustic phases a soundscape passes through, starting from the dawn till the end of the day.
Reviewing the components of the impendent visualscape, there is always a number of objects which are 'sound inactive', to wit without producing any sound or even participating dynamically to a production of sound. In the opposite occasion, there are some objects with a respectable participation in the soundscape design, without even been visible from the viewer (e.g. birds' carol hidden under the branches of trees, wind), making the visualscape even more concise.
Every instant in the real world is unique, and so do humans from the surrounding environment receive the visual and acoustic signals. The simulation of such acoustic signals is the first priority set on this project, with the corresponding graphic environment not falling short of quality and resolution. Obviously a high-resolution structured soundscape should correspond to a visualscape with analogue characteristics and content.
'Each sound or soundscape has its own meanings and expressions and is like a spoken word: it has something to say about all living beings' behaviours and their relationship to their surroundings, about listening and sound making habits. Whether urban or rural, the sounds of our home environments give us - often unconsciously - a strong "sense of place". Since audio technology and recording equipment can now be used in similarly portable ways as a camera, the soundscape can be recorded, reproduced, composed and processed by more people than ever before.' (Hildegard Westerkamp/2002)

3.6 Conclusions

Going through this chapter becomes evident that purpose of the project is to develop a virtual soundscape based on a series of sound attributes that will be similar to the corresponding soundscape of the real world. Such a ‘virtual’ soundscape however and the corresponding visualscape that is associated to, are both essential parts of the integral geodatasets, describing the area of interest.
Although this association of soundscapes to visualscapes appears as an innovation regarding the domain of geoinformation, generally is not a new application at all:
The rapid progress noted in the video game sound industry, is a representative example of virtual soundscape compositions that everybody can meet and experience in the every day life: the quick reference in the following chapter, illuminates the possibilities of enhanced perception of the real world’s soundscapes through virtual soundscape compositions. Synchronously it makes evident the ability of exploiting the existing the ‘soundscape simulating’ technology for the sound representation of 3D visualisation models in geoinformation.

CHAPTER 4

Sound experiences-videogames

Introduction:

The 'virtual soundscapes' introduced in the previous chapter can find implementation through different applications of the every day life. By experiencing them within the video game environment, it is a step forward to understand the implementation of such 'artificial' soundscapes from a more technical point of view. A quick reference in the evolution of sound synthesis and reproduction- in videogame sound industry -of the last 30 years, will be succeeded by a cue to the latest technological products and techniques for enhanced sound quality.
Aim of this chapter is not to describe the soundscape technostructure or any sound model design as they are applied in the various videogame products, but to present the technological advanced tools that contribute in the better perception of digitally created soundscapes by humans, as it is already introduced in the second chapter. A further goal is to show the opportunities for imitating or expanding the videogame technology for the sound representation of visual geoinformation.
The fourth chapter constitutes also intimation for the kind of implementation tool to be used for the implementing part of the thesis project, since it is strongly related to the videogame design industry.

4.1 Sound and videogames

The videogame industry is the closest domain where someone meets such 'virtual soundscapes'. Sound in games always seems to get lost in the rush to talk about the visuals in games. However, sounds are more important than most people realize: a trial in playing an interactive game with all sounds turned off is the best proof. Reliable research showed that decent sound can actually fool the brain into thinking the picture is better. (Maragos Nich/2003)
"Sound was slow to develop, and if one recognizes the poor graphics quality of those early games compared with those of today, it's logical to realize the sound was equally poor, if not worse. Since PCs were considered business machines, and audio was not seen to have many business applications, it was not a priority for computer developers. Despite this important influence of video games on Western culture, there is still surprisingly little academic research into the subject. Although there have been many studies on the effects of game violence, and increasingly on the culture of video games or arcades, academic institutions have been slow to react to the importance of games, and research into games audio is virtually nonexistent. Timeline for the development of games audio technology computer and games consoles work effectively the same, with the exception that the advantage of the console is that the composer may know exactly how a game will sound on anyone's system, while PCs are all configured differently, and therefore mean there are variations between units." (Karen E., URL 4.5)
Examining the evolution of sound in videogames, from the most obsolete till the latest models, will give a better understanding of the technology and methods used for the performance of sound in the videogames.
Distinguishing the videogames according to certain criteria, such as game type (pc games, handheld consoles, game consoles, arcade) or the **number of bits** (8-bit t0 128-bit), will facilitate the reference-based on different attributes. The number of bits typically indicates how much data a computer's main processor can manipulate

simultaneously but it can also be used to describe sound fidelity or graphics. In sound, an 8-bit sound is comparable to AM radio, and16-bit sound comparable to CD quality. (Karen E., URL 4.5)

4.2 The 8-bit generation

In the following paragraphs, only the most important steps in videogame sound evolution are presented, though the numerous technical details that can be involved.
Starting with the 8 bit systems, they typically had one sound channel (meaning only one note could be played at a time) [(1)]. Taito's Midway Space Invaders in 1978, was the first game to include background "music" along with six sound effects, including a soundboard with its own amplifying circuit. (URL 4.5)
The majority of 8-bit machines used the Programmable Sound Generators (**PSG** sound): silicon chips designed for applications that generate sound according to user's specifications. Some companies however (like Atari and Commodore), designed their own sound chips to improve sound quality. Early PSGs used the **analogue synthesis**, or subtractive synthesis (filtering a wave form to attenuate or subtract specific frequencies and then passing through an amplifier to control the amplitude of the final resulting sound). Apart from the PSG chips, Pulse Code and Pulse width Modulation are two other methods used for sound modification. With **PCM** (Pulse Code Modulation) essentially, an analogue sound is converted into digital sound and stored in binary (1s and 0s), which is then decoded and played back as it was originally recorded. The fidelity of the sound depends upon the **sample rate** or quantization: the number of samples of a sound that are taken per second to represent the event digitally. The method is still used, for instance for DVDs or CDs, where the sample rate is 44,100 times per second or 16-bit (which accurately reproduces the audio frequencies up to 20,500 hertz, covering the full range of human hearing). The ADPCM was later introduced to the reduced the amount of sample data required. **PWM** (Pulse Width Modulation) modulation works by changing pulse waves by outputting pulses at a constant volume, while the width and spacing of the pulse gives the effect of different frequencies and volumes, controlled by the modulating signal's amplitude.
Access Software developed a PWM technique to play relatively realistic digital audio samples, without the addition of a sound card failing however to catch on, especially when soundcards became popular. (URL 4.5)
Intellivision 1984, more advanced in sound by using a PSG sound chip, that engaged among the other audio channels one **white noise** - a sound that contains every frequency within the range of human hearing (generally from 20 Hz to 20 kHz) in equal amounts. The Intellivoice add-on chip (released in 1981) used a speech chip (Orator), for voice data. By the late 1980s, a program was converting **MIDI** files directly into Intellivision music code. Musical Instrument Digital Interface (MIDI) is a protocol to allowing compatibility in a standardized format followed by General MIDI and GS MIDI standards. With MIDI commands, rather than actual sounds, are transmitted, reducing file size.
Undoubtedly the most popular of all 8-bit machines was the Nintendo Entertainment System, or NES (1983) that used a PSG direct memory access (DMA) among the other channels, playing samples from memory.
Regarding **handheld video games** most of them had very limited sound capabilities. Game Boy Advance (2001) however, added two 8-bit **digital-to-analogue converters** (DACs) to the original (3+1 one channel stereo sound PSG) configuration of previous editions. (URL 4.5)
Arcade machines did not differ significantly from home consoles of the contemporary period, although they were often slightly further advanced in graphics and audio.

The most significant advance came when Atari included a Texas Instruments speech chip- a form of audio reproduction known as Linear Predictive Coding, which could mimic human speech or be used for sound effects.

4.3 The 16-bit and 32-bit generation

Passing into the 16-bit era, Megadrive released by SEGA in '89, was the first real 16-bit engine PSG chip to handle effects, as well as a Yamaha FM chip with six channels of digitized stereo sound. The last version of mega drive, the 16- bit **Super SNES** owned a 16-bit Sony digital signal processor (DSP) and 16-bit stereo digital to analogue converter (DAC) - The method of converting analogue sound to digital sound). (URL 4.5)
Regarding 32-bit machines, Sega released their 32-bit **Saturn** in 1994, with two audio processors, consisted of a 32-channel PCM sound generator and a 16-bit DAC (Digital Analogue Conversion.. The only drawback to Saturn's sound system was the limited amount of RAM accorded to sound. Because audio samples had to be downloaded raw (decompressed) into the audio memory buffer, this meant there was a limited amount of space for simultaneous sounds, and so the sample rate was often reduced to conserve memory.
Sony decided to press ahead with its own 32-bit system, the Playstation where samples did not have to be compressed with sound quality greatly improved on the SEGA Saturn.
After splitting off from Sony, Nintendo went straight to a **64-bit** release in '96, the **Nintendo 64** producing 16-bit stereo sound at a slightly higher sample rate than CD quality with some games supporting even surround sound. (URL 4.5)

4.4 Towards the latest generations: 64-128-bit machines

The **Sega Dreamcast** was (the first 128-bit console), used a special Gigabyte Disc ROMs: samples did not have to be decompressed, improving audio capabilities. True 3D audio was supported, in CD-quality 64-channel sound, with effects such as delay, reverb and surround sound. In 2000 **Play Station 2** was launched with the ability to play DVD movies and fully supported the multi-channel DVD sound standards AC3, DTS and **Dolby Digital** which will be later discussed.
Regarding the latest video game releases, Nintendo **game cube** launched in 2001 with ADPCM 64-channel, Dolby, DTS and AC3 has an audio capability comparable to that of the Playstation 2.
With the **XBOX (2001)**, Microsoft entered in the game console business: featuring its own audio processor, supporting Direct X 8.0 and 64 voices using 3D positional audio. (URL 4.5)

4.5 Technological tools and techniques

3D positional audio, (head-related transfer function), uses signal processing to locate a single sound in a **specific location** in three-dimensional space around the listener, and is common in video games, because it can be interactive. With 3D positional audio, sound objects in a virtual space can maintain their location or path of motion while the gamer moves about. (URL4.5) Three-dimensional **(3D) sound** gives the listener the perception that the **sounds are emanating from a 360º** (three-dimensional) space. Apart from HRTF there are two other important terms regarding 3D audio that should be differentiated, specialization and virtualization.

Specialization, also known as stereo expansion, uses signal processing to give a listener the effect that **the sound is coming from a wider space** than that of the real location of speakers. Specialization disperses perceived locations of sounds so that a listener cannot determine the location of the sound. **Virtualization**, gives the impression to the listener that **speakers are more than actually are**. Rather than locating a specific sound in a specific location (see below), virtualization locates a specific channel to a specific location.
"Of course spatialized sound has been used before - lots of games split the mono sounds over two channels and use 3D distance from camera to determine each channel's volume. All of these speaker solutions only offer 2 dimensional sound systems. With 4 speakers, you can hear a sound anywhere around you in the horizontal, but not the vertical. You need a speaker over the top of you, and one beneath you for that to work correctly. Both Creative and Aureal mention that they offer some algorithms for simulating vertical sounds on a horizontal setup, but there's **no substitute for physical positioning of a sound** source. Creative Labs uses EAX software -EAX stands for Environmental Audio: the audio preferences of your surroundings reflect the type of surroundings you have. When you move from one area to another, you can change the surrounding area's aural characteristics. Microsoft has already included EAX support inside of DirectSound. Aureal offers a similar solution to EAX called A3D supporting **mp3 decompression**. The Miles sound system is more a layer of API that sits on top of whatever kind of card you have already supporting A3D, EAX, and DirectSound. It does all the initialization it needs to, figures out what kind of sound card you have, and basically does everything for you, plus more." (Jake Simpson, URL 4.2)
Dolby audio technology is the pioneer in Multichannel surrounds sound that continues to bring its audio technology to the video game industry, with top video game consoles such as Nintendo GameCube, Microsoft Xbox and Sony PlayStation 2. [9] "Dolby Surround encodes four sound channels (Left, Centre, Right, and Surround) into the two tracks of any stereo medium, including console- and PC-based game systems. Companies such as NVIDIA and C-Media have optimized their hardware to include the real-time Dolby Digital encoder technology (as found in the Microsoft Xbox), providing the highest quality digital audio signal for the PC. (URL 4.10) Dolby has designed the Virtual Incubator Program (VIP) to promote audio excellence in games and has also developed the groundbreaking Interactive Content Encoder, which enables real-time Dolby Digital encoding in game consoles. (Wootton Bassett /2003, URL 4.9)

Recently OpenAL was developed and set as a standard cross-platform API to allow developers to add sound to any title. OpenAL is basically an **audio library** that contains functions for playing back sounds and music in a game environment, allowing a programmer to load sounds and control characteristics like position, velocity, direction and angles for cones that determine how the sound is travelling. All sounds are positioned relative to the listener, which represents the current place in the game universe where the user is. (Dan Ricart, URL 4.4)
Technological development in sound for videogames however, also includes the evolution in sound cards. Sound cards take raw sample data and perform operations on that data, mix it together and then send the result to the speakers. Some of these operations include echo, reflections, Doppler shifts, and pitch bending and so on. The mixer mixes samples processed by the card, as well as outside analogue inputs, like the sound from your CD-ROM drive or mikes that you plug in. (URL 4.2)
Ad Lib Multimedia (1986) was the first popular card based on the Yamaha FM chip, enabling game developers to use a wider range of instruments and sounds. Creative Technology (Creative Labs), released the sound blaster, becoming soon a standard for game sound: using the same FM chip, but added digital audio capabilities for sampling, and a game port. Sound Blaster quickly became the standard for game

sound. One of the problems with the Sound Blaster cards was the fact that it had to mix the various sound channels into one or two output channels, resulting in a loss of resolution in the sound.
Developed by Sierra, Roland and Ad Lib, the MT-32 was a MIDI soundcard that used a form called Linear Arithmetic synthesis: What the human ear recognizes most about any particular sound is the attack transient. LA based synthesizers used this idea to reduce the amount of space required by the sound by combining the attack transients of a sample with simple subtractive synthesis waveforms. (URL 4.5)

4.6 The role of soundcards

The first soundcard to support 3D sound was Diamond Monster Sound, in 1997. Parameters such as room size and acoustic properties could be programmed into the audio, which would initiate filters and effects to simulate the space.
When CD-ROMS came out, MIDI was pretty much abandoned, and with it the notion of interactive music. Playing sound/music directly from a CD-ROM drive has the best quality. As during playing a game the CD-ROM drive must also access other large data such as graphics at the same time, the method of direct play sound/music may result in reducing the game speed dramatically even for a 4X speed CD-ROM drive. To maintain a reasonable speed, many CD games use the old method - play sound/music from PC audio hardware." (Richard S., URL 4.6) CD-ROMS ensured more fee space for audio in video games but it was still limited considering the space occupied the game: thus, various compression technologies were developed, the most important of which was MPEG level 3, known more commonly as MP3. MP3 meant that much less data would be required to store audio, and game companies quickly began incorporating MPEG compression into their games music

Aside from the conventional soundcards, **3D sound cards** deal with what the card does with 3D sounds. The difference between a 3d sound and a 2D sound is that a 3D sound has a 3D origin, is spatialized for 4 speakers. On the A3D card, once a sound is submitted, it can work out all the reflections for the sound inside a room on a dedicated on card chip, along with all the filters required for the sound - like Doppler effects, orientation effects and so on. (Jake Simpson, URL 4.2) "As the capabilities of sound cards and sound APIs have increased over the years, **3D sound** has played an increasingly important role in games. Creating an immersing experience for the gamer, goes beyond nice looking graphics. Done correctly, ambient sound and music can take games to a whole new level. To achieve this effect, different sound programming APIs has been developed such as Microsoft's DirectSound and DirectSound3D, Aureal's A3D and Creative Labs EAX." (Dan Ricart, URL 4.4)
DirectX, a multimedia **application programming interface**, improved the speed that sound and graphics cards could communicate, became soon a standard allowing for higher-quality 3-D graphics, and better control of sound mixing and output. One part of DirectX, DirectMusic, overhauls the old MIDI protocols by offering the industry-ratified DLS (**Downloadable Sound Levels** 1) specifications support for hardware acceleration and MIDI. (URL 4.5)
Music and sounds for the video game industry are encharged to **sound designers** (composers): they create sound effects by modifying an existing sound from an **audio library** or by finding and **recording the sound themselves**. A good cooperation between sound designers and game designers or programmers results in a proper blend between 3D graphics, plot and sound. "An important part of the job is to find creative ways to make sound. Sound designers edit almost every sound they use: they might lower a pitch, add an echo, and loop the sound to make it longer, or mix it with other sounds. They balance realism with the entertainment value of exaggeration, routinely sweetening natural sounds for dramatic effect. Other

sounds are triggered by an event. Sound designers are given a list of sounds the game needs. When choosing noises, sound designers also have to be aware of the game-playing environment. When designing for arcades, for example, they make effects loud and simple. For a home system, sounds can be more complex. Sound designers also need to learn the basics of computer hardware and software to predict how their compositions will sound to the player." (Olivia Crosby, URL 4.12)

It's not only hardware however that played an important role in developing sound in videogames: there were also **software developments** such as iMuse system that used the hardware capabilities to expand on possibilities for games composers (e.g. change volume, tempo, or add or remove instruments or sound effects in response to a given action in a game).

Having a quick view on the names of the latest and most popular videogames in the market, it's not coincidental the choice of Virtools Behavioural software for the development of the demo application of the research: '**Virtools** gives game studios the technology they need to develop complex, high-quality 3D games in record time. 100% of the games produced with Virtools have followed the full development cycle up through public release, to generate revenue and acquire a committed audience of gamers. Characterized by a versatile behavioural engine and an industry-quality rendering engine, Virtools enable XBox game developers to prototype in weeks rather than months, and concentrate on the quality of both graphics and gameplay'. (URL 4.13)

4.7 Conclusions

In the chapter above, the reader gains an impression of the "technological tools" already used for the 'soundscape investment' of the videogames the last years-the same or contiguous tools/techniques that can be used for the reproduction of geosounds.

Aim of the fourth chapter was not to approach the technostructure or model design of the 'virtual' soundscapes but to examine the existing technology invoked already in geoinformation's 'neighbouring' domain: the video game industry. The structure and methodological plane of the virtual soundscape is developed in the following chapter.

CHAPTER 5

Conceptual model

Introduction:

Aim of the fifth chapter is the description of the conceptual model that will lead to the representation of a virtual 3D environment by sound. The sound representation of the visualscape will result in a **virtual soundscape** in the form of a digital sound interface. This digital interface constitutes part of a complete **multimedia** model that will be presented into the next paragraphs of this chapter.
One of the basic research questions to answer through this chapter is the structure of the soundscape and the mode that is related to the visualscape. Particularly it is described how sound follows the different resolution levels of the image.
Regarding the proper positioning of sound on the 3D space, there will be introduced three approaches depending on the contribution of the virtual 3D space and the sound experiences from the real world.
A further point for discussion is the raster or object methodological approach to be used for the discrimination and attachment of sounds to geographical entities and 3D objects.

In the first paragraphs of this chapter, are introduced some conceptual frameworks (models) aiming to define the relations between some basic components (or systems). Between these components are also described different processes in order to acquire process, represent and perceive geographical information. They are also related to the way human see, hear, act and perceive the world (real, graphically or digitally represented).
Later, some of these components and relations will be selected and specified in order to represent the **flow of processing data** that will lead to the composition of the 'virtual soundscape'.

5.1 Obtaining the 'Visualization model'

A fundamental model whereat is based the synthesis of the multimedia model is the WDGM model- which stands for World Digital Graphic Mental model- composed by R.Van der Schans, (Schans V.d.R. ,1990). In this model are analyzed sixteen possible relationships between the components (or systems) that constitute the WDGM model: the real world, the digital models, the graphical representations and finally the mental representations by humans. All the relations between the aforementioned components are analytically described in figure 5.1.

Figure 5.1 WDGM MODEL BY R. V.D.SCHANS 1990

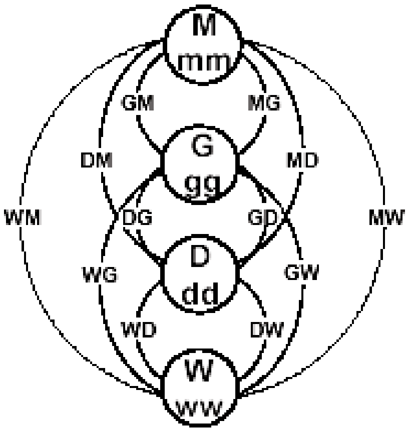
THE **WDGM** - MODEL.
a conceptual framework for description and analysis
of geographic information handling processes

In geographic information processing
four BASIC SYSTEMS are of primordial importance: W, D, G and M

Within and between these systems
sixteen BASIC PROCESSES can be distinguished: WW to MM

All basic systems and basic processes
can be represented in a simple matrix or graph form:

	W	D	G	M	
W	WW	WD	WG	WM	World
D	DW	DD	DG	DM	Digital models
G	GW	GD	GG	GM	Graphic representations
M	MW	MD	MG	MM	Mental representations



It's possible to build combinations of basic processes
e.g. DGMD is a (carto)graphic man - machine conversation.

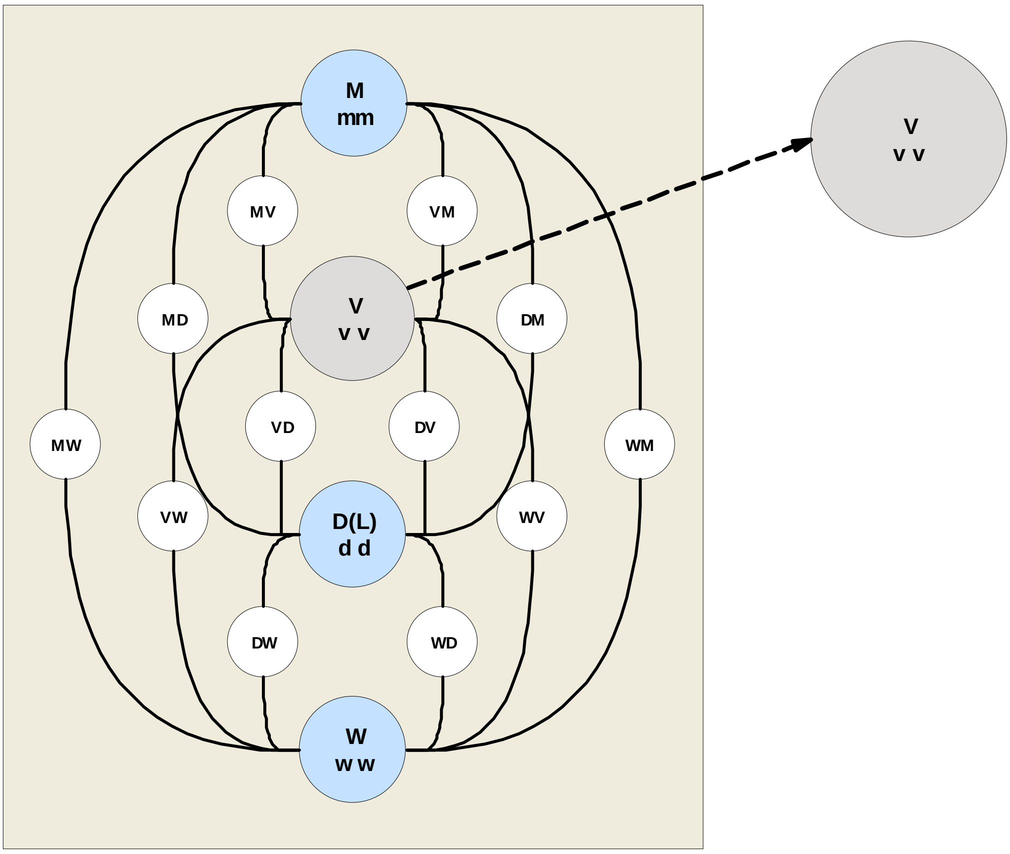
	W The World (observable)	D Digital Models (in machines, invisible)	G Graphic Representations (visible)	M Mental Representations (in humans)
W	**WW** **processes of change in the world** **(e.g. Erosion)**	WD digital data acquisition about the world by artificial sensors (e.g. remote sensing)	WG photographic recording of the world	WM observation of the world cy human senses
D	DW digitally controlled intervention in the world by machines	DD processes in digital models controlled by algorithms (analysis, simulation, data communication)	DG graphic representation of digital models (visualization)	DM non-visual information transfer from digital models to mental representations
G	GW changes in the world directly caused by graphic representations (almost non-existent)	GD automated transformation of graphic images to digital models (scanning)	GG reprographic processes, exchange of graphics	GM looking at graphic images (e. G. map reading)
M	MW direct intervention in the world by man's hand	MD human intervention in digital models, by means of keyboards, cursors, mouses, etc.	MG production of drawings directly by hand	MM mental processes of man

Examining the entire WDGM model as an attempt for description and analysis of all geographic information handling processes, it becomes obvious the generic character of this conceptual framework. The most crucial part of the WDGM model that plays a determinative role on the development of the multimedia model is the

‘graphic representations system’ and consequently the respective relations with the other systems of the model.
Considering that the graphic representations are mainly translated through ‘cartographic models’ or ‘digital visualisation models’, makes possible the specification of the WDGM model by concentrating on the visualisation domain. In figure 5.2 is described a specialised version of V.d.Schan’s WDGM model by replacing the ‘graphic representations system’ by the ‘visualisation system’, thusly forming the WDVM model.

The WDVM model

Figure 5.2



Looking carefully on the model above, it is possible to distinguish a further specification of the primary model: the modification of the ‘digital models’ component by the ‘digital landscape models’ (D(L)). This is a rather important notation regarding the ‘digital soundscape models’ which will be later introduced. Generally all the relations between the components of the last model are similar to the first model, expressed however in a more explicit character after the specification of the ‘Digital model’ and the ‘Graphic representation’ systems. The most important relations of the modified model are explained in the next paragraphs:

- VD (L): From digital visualization model to landscape model. Use of the ‘top view’ of a VRML browser to watch the 3D visualization model as a digital 2D map. Discrimination of thematic and geometric data by use of a VRML browser.
- D(L)V: From digital landscape model to digital visualization model.(3D)Graphic representation of thematic and geometric data
- VM: From the digital visualization model to human’s mental representation. Looking at 3D visualscapes (mainly through pc screens).

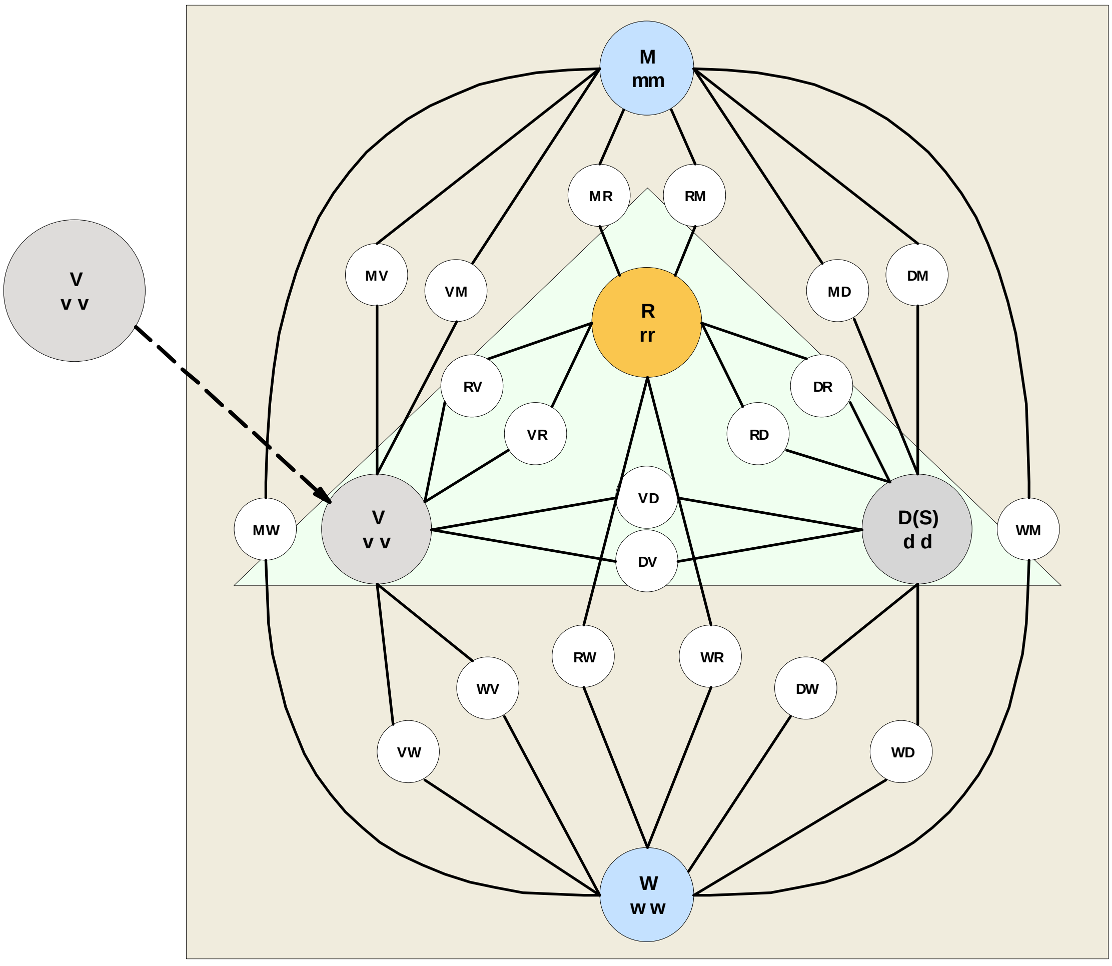
- MV: From human's mental representation to the digital visualization model. Digital construction of 3d worlds by use of 3D software (e.g. CAD)
- VW: From the digital visualization model to the real world. Dynamic changes on the 3D visualization model of the real world's features.
- WV: From the real world to the digital visualization model. Photographic recording of the world- use of photos in the 3D visualization model.

In the conceptual frameworks presented in the previous paragraphs, the digital visualization model is the most highlighted 'system' since it constitutes one of the basic components of the WRM conceptual framework, which is described in the next paragraph.

5.2 The WRM model

The WRM model, presented in figure 5.3, is a conceptual framework, similar to the previous models. Aim of the WRM model is to present the relations between the following systems: the real world (W), the visualization model (V), the digital soundscape model (D (L)), the multimedia model (R) and the mental representation by humans (M).Apparently most of the systems involved in this framework (W, V, M) are also present in the previous described models. The main difference from the previous models is the introduction of two new components: the digital soundscape model and the digital multimedia model.

Figure 5.3 **The WRM model**



The WRM conceptual framework is mostly focused on defining relations between sound and visual data, contrary to the previous frameworks where the sound component is absent.
In table 5.1 twenty five relations among the five components (or systems) that constitute the WRM model are described. The most important relations however, which will later determine the composition of the multimedia model, are comprised within the green triangle (figure 5.3). Particularly this triangle contains all the possible relations between the digital visualization model, the digital soundscape model and finally the digital multimedia model.
The WRM model has a similar structure to the previous conceptual models as well. The W and M systems remain identical to their positions, while the digital landscape model is replaced by the digital soundscape model (as long as the WRM model is concentrated on defining audio visual relations). In figures 5.2 and 5.3 the visualization model (V) is illustrated comprised in the WDVM conceptual framework, which is introduced in WRM model in order to be **acoustically represented by the digital soundscape model leading to the digital multimedia model**.
The digital visualization model constitutes a significant part of the WRM model. As it will be later discussed, absence of the digital visualization model influences negatively the definition of the **digital soundscape model** and the formation of the **digital multimedia model**.

Table 5.1

SYSTEMS	**W** **Real world**	**D(S)** **Digital soundscape model**	**V** **Visualization model**	**R** **Multimedia model**	**M** **Mental representation**
W	**WW** Processes in the real world (phenomena)	**WD** Digital sound data acquisition by recordings in real world	**WV** Photographic recording of real world to use in V.Model	**WR** Observation of virtual models through pc screens etc	**WM** observation of world by human senses
D(S) (or DSM)	**DW** Digitally control of world's sounds by audio devices	**DD** Processes within digital S.scape models	**DV** Ascribing sounds to Vis. model	**DR** Sound files& characteristics properly positioned in 3D space	**DM** Sound files, not properly perceived owing to luck of V.Model
V	**VW** Dynamic change of real world features within the V.Model	**VD** Detection of sound sources in Vis. model	**VV** Visualisation processes within the model	**VR** Definition of 3D space to link sounds	**VM** Looking at 3D modes using a VRML browser
R (or DMM)	**RW** Simulation of reality by graphic and sound repres. of dig. models	**RD** Proper position of sound in the virtual 3D space	**RV** Identical to the Visualisation model with sound switched off	**RR** Rendering processes of the Mult.model	**RM** Presentation of audio visual model using a VRML browser
M	**MW** Direct intervention in real world by humans	**MD** Human intervention in digital sound models by sound software	**MV** Digital construction of real world (by GIS,CAD etc. software)	**MR** Combination of V.Model and Sound Models using special software	**MM** Mental processes of humans

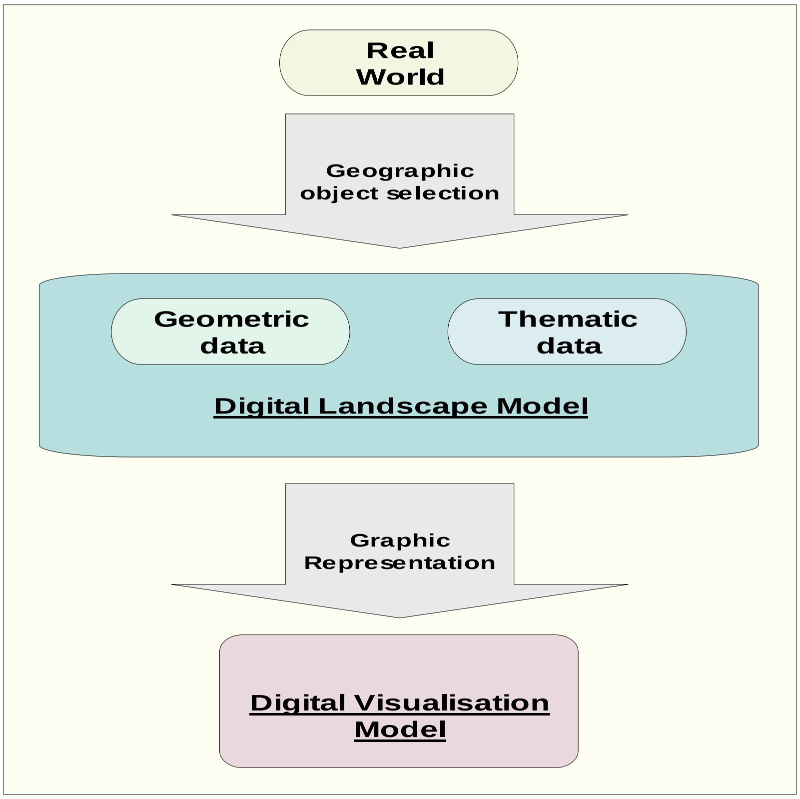
Before proceeding with the descriptions of the soundscape and multimedia digital models, it is important to distinguish some relations within the conceptual models which describe the flow of data processing in order to come up with the multimedia model.

5.3 Data flow processing

In the conceptual frameworks presented in the previous paragraphs, only the relationships between the various components that constitute each conceptual model are described. All the relations between two different components in each model, are 'two ways' expressing a reciprocal relation in two directions.
Describing the formation of a digital multimedia model however presupposes the definition of one way relationships in order to illustrate the flow of data processing between the various components (within each conceptual framework).
In the WDVM model, the distinction of two – of major importance - relations, the WD and DV is graphically illustrated in the next figure 5.4.

Formation of digital visualization model

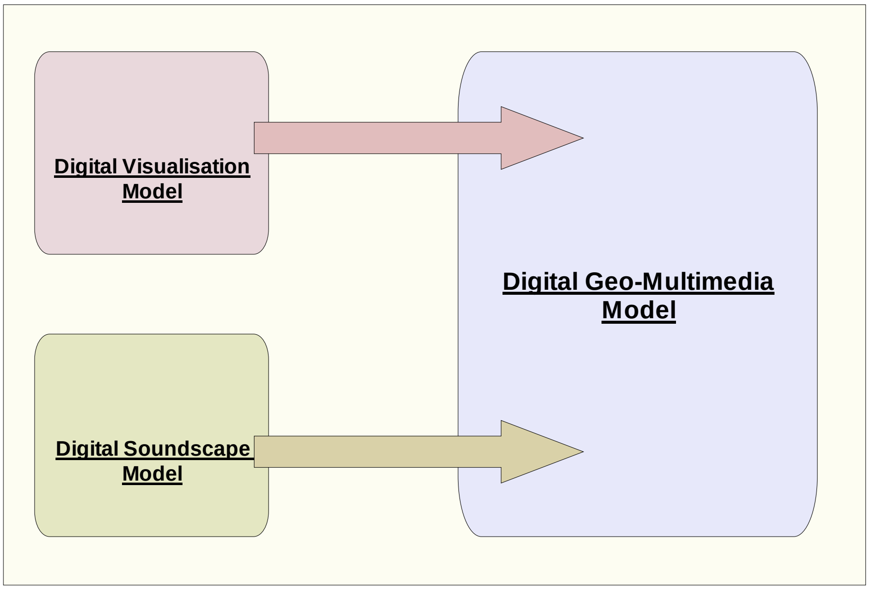
Figure 5.4



The schematic representation above resembles the 'nature flow of geospatial data' described by Menno-Jan Kraak and Ferjan Ormeling (Cartography, 2003), diverging however in the graphic representation of the digital landscape model. The model describes the data process from reality, via model construction and selection of a digital landscape model towards a digital visualization model (3D model and not 2D cartographic model).

Proceeding with the WRM model, the most important relations to distinguish are the VR and the DsR (one way relations) which are schematically represented in figure 5.5.

Formation of digital geo-multimedia model
Figure 5.5

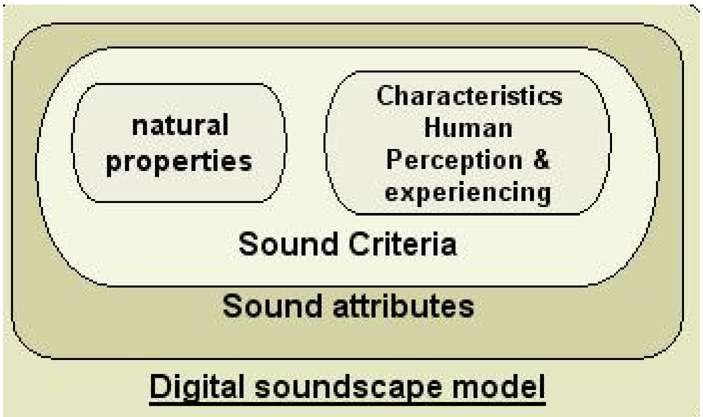


This model constitutes an extension of the previous visualization model. The digital visualization model (as an end product of previous processing), is acoustically represented through a digital soundscape model, leading finally to a digital geo-multimedia model.
Apart from the VR and DsR relations, in the WRM model are presented other significant relations such as the one between multimedia model and human mental representation, or between digital visualization and digital soundscape models. The first relation represents the human perceptibility of the real world through the multimedia model, while the second is later discussed within the description of the digital multimedia model.

5.4 The digital soundscape model (DSM)

Before proceeding with the analysis of the soundscape model, it is important to define sound as aspect of the real world as well as the relation between sound and the real world's entities and phenomena.
Sound is mainly recorded within the real world's soundscapes in order to contribute in the acoustic investment of the virtual soundscapes, without excluding however the use of 'imitating' sounds produced by electronic means. To be able to associate a sound within a visualisation model – or image in general- the nature of sound has to be defined first. References of sounds in literature as 'sound objects' are very common, though without examining the real hypostasis of sound itself or corresponding with the rest of the components or phenomena of the surrounding space.
Considering **sound** as an **attribute** that accompanies the various geographical entities, objects and natural phenomena of the real world, implies that sound is always depended on the presence of these 'factors'. In other words, sound cannot stand as an 'independent object'- since it is **not an object**- in the real world but appears as a depended parameter –as a physical feature.

In addition, sound as an **attribute** of the real world's entities or phenomena is characterised by a number of **natural properties** such as frequency, pitch, intensity (analytically described in the second chapter). A representative example that describes the nature of sound is the man's voice: as an attribute of human nature, presents some natural properties like frequency, pitch, loudness etc.
Apart from the natural properties, sound presents some **characteristics** (that cannot be categorised as natural attributes), depending on the way that human experience the sound and the current conditions (e.g. the Doppler Effect, harmonics etc).
When composing a virtual soundscape, all the **natural properties** and (experiencing) **characteristics** of sounds must be taken into account in order a proper rendering of sound to appear: in other words some **sound criteria** must be defined that is a **combination of sound properties and characteristics.** Part of sound criteria can also comprise some **variables** (like sequence of sound) which are not describing the sound as an attribute but show the relation with other sounds. These sound criteria are part of the digital soundscape model and are described in the following section.



The digital sound model consists of a sound – attribute interface which serves for the sound representation of the visualisation model. It is important to declare that any congruency between the digital soundscape model (DSM) and the virtual soundscape cannot exist since the sound attribute interface of the first – in the form of raw(not processed) sound data- is not positioned yet in the 3D virtual space. In other words, it is the digital visualisation model that serves as a conductor for the proper displacement of the sounds into the 3D virtual space. A further issue regarding the DSM has to do with the technique used for capturing those sounds.
When dealing with attributes of general nature, such as the aural ones, different approaches can be followed for capturing and processing the sound features comprised in the sound model. A first approach is 'visualisation model' depended. Since the distinction and selection of sounds is based on the way of experiencing the digital visualisation model (e.g. by a VRML viewer) from different views (perspective, top, side) and through varying detailed distances (zooms), the detection of sounds emanating from visible sources is achieved.

The main disadvantage of this approach is the loss of a respectable part from those sounds originating either from 'invisible sources' or natural phenomena that cannot be included or forecasted in a DVM. The 'invisible sources' refers to entities that cannot be visible from the maximum resolution level of the visualisation model. Obviously, a second approach related to experiencing the research area without the interference of the visualisation model works as more integrated sound-sample collections. The second approach is mainly based on the human's perception of sound, meaning that the distinction and collection of sounds –on sight in the area of interest- can be even made with the eyes closed! Apparently, the second approach is more laborious than the first: the exploration of the real world is no longer made in a quick and virtual way. The only view point is the explorer's eyes and the field work of recording 'invisible' sounds could be unlimited!

However, the last approach is the only way to experience the temporal dimension of 'both visible and non visible' sounds, aiming to a more accurate aural simulation. For example observing the weather's changing conditions in an area and variant duration sound effects can provide useful information for the formation of a multimedia model, contrary to the 'deaf' or 'time independent' DVM. Nevertheless, the digital visualisation model is indispensable in both of the aforementioned approaches for the proper allocation of sounds to the virtual 3D world and assignment in different detailed levels (as it will be further analysed afterwards).

The sounds to be used in both approaches have to be simple and clear or in other terms 'lucid sounds' and not 'mixed' recordings of different sound sources at a time. Evidently, the registration of 'lucid' sounds in complex or dynamically changing soundscapes of the real world is an insistent work that requires further processing in sound studios or use of 'store bought' sounds from a sound library. A separation of sounds in 'point' sounds and 'background' sounds is ascribed to the nature of sound sources: in the case of background sounds (e.g. sound of rain) it is possible to be imported in the multimedia model without the contribution of the visualisation model, owing to the generic nature of the sounds.

A third methodological approach under the scope of locating and capturing sounds in natural soundscape is related to the geo-informative character of sounds. Similar to the second approach, the composer is 'ignoring' the use of the visualisation model. Instead the composer registers sounds at preferable points of the research area and contemporaneously geo-references the 'recording points' using a GPS device. The association of those points to the digital visualisation model presupposes a georeferenced virtual 3D space as well, which is a major criterion for choosing or rejecting this approach. The use of the last approach is not applicable in moving sound sources while the registration of sounds in the bounding-georeferenced- points results mainly in mixed sounds, difficult to individually handle and assort (in the different acoustic level of details of the multimedia model). Consequently the research is based on the first two approaches analysed in the previous paragraphs.

5.5 The sound criteria

Frequency and **pitch** are the most common of the natural sound properties. A modification in sound frequency when introducing a sound in a virtual soundscape, is sometimes essential for a proper rendering, respect to the rest existing sounds. In an analogue way, sound pitch has to be adjusted in order to contribute in realistic sound rendering.

Sound intensity is one more significant attribute of sound and just like the pitch and the frequency, has to be correctly adjusted in order to accomplish the relevant loudness for every sound. For example, a car's horn must be heard louder than its engine which is consecutively louder than the bird singing.

The **duration** of each sound is also an important attribute and the right adjustment within the multimedia model depends on the experience of the composer with the sounds in the natural soundscape.

In the following criteria the state and the behavioural rules of a certain sound respect to the other sounds and the observer is described.

Static and **kinetic** are the two moving states that describe the source where the sound is originating from. Dealing with kinetic sounds is one of the most critical cases

in the composition of a virtual soundscape since they cannot be georeferenced or detected in a specific location of the visualization model.
Separation of sounds in **permanent and temporary** are two other criteria strongly related to the temporal period the application is referred to. This is why a series of chronic criteria like the season, the time of the day, or the weather conditions must be predefined. An example that shows the relativity of a permanent sound is the flow of the river and the bird singing. The sound from the river flow is always constant while the carol of the singing birds in a forest could vary during the time of the day or the season of the year. In addition, temporary sounds, which are sounds that last only for a certain period (few seconds), like a car's horn or the sound of an airplane.
Just like the discrimination of spatial phenomena in **continuous and discrete**, the same concept can be used for sounds as well. Usually there is a correspondence between the optical and aural discretion. A big forest for example is considered a 'continuous' viewing element, while a car is a discrete one. The sound emanating from a forest is also a continuous element while the engine of a singe car is a discrete sound.
Periodical sounds are reproduced in a soundscape characterized by a certain period (e.g. the clock of the church ringing every hour). **Random** sounds however are more frequent in nature (like the vehicles that pass from a road in unequal time intervals or even the moos from cows laying in a meadow).
Diversity and variety of sound are two more criteria that have to be specified for the separation of sounds into the three LODs. For the simplification of the soundscape- as well as for limitation of dispensable time- some sounds have to be homogenized and comprised within a single category. A representative example rises again from the environment of singing birds, where varying melodies can be substituted by a representative sample of bird singing, recorded even in a different area!

Visible and non visible sources constitute a further parameter to take into account. Some of the entities in nature have a more dynamic presence in the acoustic space than in the visible one! A representative example is the sound emanating from the bank of a channel or a river. Sounds reproduced by ducks, frogs or other amphibians have a significant contribution in the soundscape of the area, and an overlooking presence in the visualscape.
Making a distinction between **important** and **meaningless** sounds can save much time and effort during the design of the soundscape. Sounds that have a minor contribution in the soundscape or have a minimum possibility to be heard can be easily omitted. This discrimination represents also a criterion for the assignment of sounds to the proper level of sound detail (LOSD).
Some sounds that belong to entities, characterized by a specific **animation**- meaning move within a certain frame and following a predefined route- are a special case of sounds to mention: a representative example is the sound of the flowing river, which is hard to categorize it either as static or kinetic sound attribute.
An important **variable** used for the processing of sound attributes is the **sequence** of the participating sounds. Some sounds have to wait the completion of other sounds in order to give optimal acoustic results. The thunderclaps for example, are posed just before and during the rain and never afterwards. A wise selection of sound sequence would render to the application an even more realistic character.

A further discrimination of sounds in **natural** and **artificial** is an essential criterion for the easier detection and separation of the sound sources either in the real world or through the visualisation model.
The **coexistence** of a sound together with other sounds in a soundscape is a particular attribute of sound related to other attributes such as the duration and sequence. A 'strong' wind or the sounds of a storm for example cannot coexist with

the carols of the singing of birds because in intensive weather conditions the birds are silent!

Finally the **sound density -** which describes the number of sound sources within a predefined surface unit of an area - is one more criterion with a determinative role in the methodological approach for the audio-visual linking.

It is important to refer that the aforementioned criteria cannot constitute absolute criteria, especially when they are used to define such an abstract concept as sound is – without optical borders - much more abstract than images.

Beside the reported **criteria** (properties & characteristics), sounds can be even easier to trace and handle when separated into the following general categories (based on the nature of the sound source):

- Sounds belonging to natural phenomena, like rain, wind and thunders.
- Sounds belonging to natural (geographic) elements, in static state (e.g. the sough of leaves in a forest)
- Sounds belonging to natural (geographic) elements ,in dynamic state, moving within or without boundaries (e.g. the river flow)
- Sounds belonging to live organisms like humans (voice, footsteps) and animals.
- Sounds belonging to artificial elements or human activities ,in static state (e.g. the campanile of the church)
- Sounds belonging to artificial elements or human activities, in dynamic state (like the sound of the car or the airplane)

The sound categorization above can be used in any of the three reported approaches for discrimination and collection of sounds.

Another component of the digital sound model contains all the sound characteristics from the human way of experiencing a soundscape. Characteristics like the Doppler or the Roll off effects, which are depended on the human's velocity and distance from the various sound sources, add to the multimedia model an even more realistic mode of experience.

5.6 The digital (geo)-multimedia model (DMM)

As already described, the digital multimedia model constitutes the final product resulting from the sound representation of the digital visualisation model through the digital soundscape model. Substantially, the digital visualisation model is acoustically represented by a digital soundscape 'interface', consisting of sound attributes that characterize the geographical entities and objects (that belong to the visualization model). This digital soundscape 'interface' is based both on experiences of the real world but also on the exploration of the area through the visualisation model.

The prefix 'Geo' in front of the 'multimedia' term is given to reflect the geoinformative character of the model. As already mentioned in previous chapter, the geoinformative character of sounds cannot be absolutely depended on parameters such as georeferency, though the generic nature of sound compared to the image. Besides, the way some sounds are registered in a real world's soundscape or later attached to a 3D visualisation model is not always accurately positioned respect to the visible sources. Georeferency is not only a weak point of the sound features since it cannot even be determined for many of the visible sources. In addition there is always a group of sounds of even more generic nature (like the sounds of weather phenomena) that characterise a whole geographical area and can't be associated to a point sound source.

Consequently, sounds that are not officially considered as 'geodata' can provide substantial geographical information.

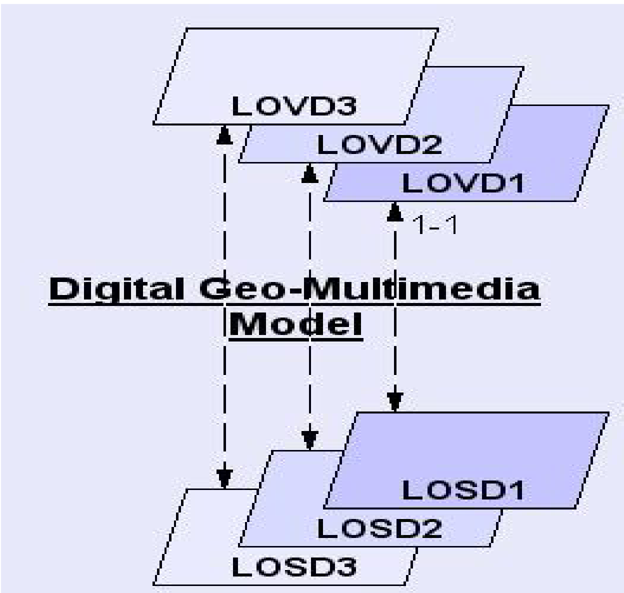
In the next paragraph, the geo-multimedia model is presented, which originates from the two earlier described models: the digital visualization model and the digital soundscape model.

5.7 Levels of detail

The allocation of sounds into the digital visualization model involves the combination of **two** different **approaches**. The first one deals with the arrangement of sounds in the 3D virtual space, based on the sound criteria that were described within the digital soundscape model. The second approach is related to the assortment of sounds in three visual levels of different resolution.
Starting from the **second approach**, apparently the sounds can be first categorised in three sound resolution levels and consecutively linked to the corresponding visual resolution levels (figure 5.6).
Each level of different sound resolution (or **Level of Sound Detail**-**LOSD**) comprises sounds of different 'sound range' and 'nature'. For example the sound of an airplane can be heard from all Levels of Sound Detail while the sound emanating from an animal (e.g. sheep) is perceived only by the first or second LOSD. From this example it becomes obvious that the Levels of Sound of detail are highly depended on the definition of the Levels of Visual Detail.
In an similar way to the LOSD, each **Level of Visual Detail** (**LOVD**) comprises geographical entities or objects of different magnitude and of different optical detail. Depending always from the predefined resolution -or distance from the current view (zoom) in the visualisation model - the discrimination of the various sound sources belonging to each visual level is achieved. In this case, there is an identical relation (one to one) between the LOVD and the LOSD meaning that every level of sound detail corresponds to only on level of visual detail.
.

Structure of geo-multimedia model

FIGURE 5.6



The number of the different detail levels however is not absolute and could be increased or reduced depending on the research area and the requirements of the multimedia product. In addition, the correspondence between the visual (LOVD) and the LOSD could vary according to the importance of using more visual or acoustic number of levels. In that way for example, two levels of sound detail could correspond to one level of visual detail, meaning that a change between the visual levels wouldn't be followed by a corresponding change of sound. For simplified reasons the number of Levels of Sound Detail (or LOSD) is identical in both of the audiovisual components of the multimedia model.

Far from the separation of sounds in different levels of detail, the sequence between sound and image is also depended on the predefined **'sound range'**, which is substantially the influence of every sound on the surrounding space (intensity).

The definition of the different levels of visual detail within the digital visualisation model are defined through a VRML browser

Regarding the **first approach**, mentioned in the beginning of the paragraph, the assortment of sounds on the 3D visualisation model (by the use of sound criteria) can be assisted using a grid (raster) layer.

Following a raster oriented approach, is a hard way of working since a series of parameters have to be justified in order to avoid any obscurities on the model. The size of the area of interest for example is of major importance since it is straight associated to the size of raster cell to apply for each level.

A cell is an object, which represents a (rectangular) spatial extent on the ground and encodes an attribute or attributes for that extent (Ervin S.M., 1993). In the multimedia model, a cell encloses a specific geographic part with certain sound attributes. The extent of the geographical area along with the variety of sounds - included in each cell – depends on the cell's shape and size as well as the geo-diversity of the area. Different classes of cells (belonging to the first second or third LOD), have different displays depending on the visualization views of each level and the available tools for exploration of the virtual space.

Examining the research area by using the top and perspective views of the visualization model, three different focusing (zooming) views are selected, which represent the three resolution levels of the multimedia model (as mentioned before).

Obviously, the number of raster cells as well as the size of grid applied on the top view of the visualization model varies according to the resolution level (or zooming distance): While the dimensions of the grid increase when overlays a higher resolution visual level, the dimension of each cell decreases so as to encapsulate a manageable number of higher detail entities and sounds.

Contrary to Stephen M. Ervin's model (Ervin S.M., 1993), each raster cell is not characterized by certain behaviour but serves as an auxiliary tool for the geometrical apportionment of each level. In other words, the composer has to recognize which sound sources are included into every cell of each level, without bothering in attaching the sounds to entities (see example in the appendix).

After attaching the sounds to the relative raster cells, all the sounds are adjusted based on their physical properties and criteria described in the digital sound model. A further adjustment of those sound attributes related to the human way of perception, leads to a realistic mode of experiencing the multimedia model.

The digital geo-multimedia model is an extension of the term 'virtual soundscape' because the visualization model is already integrated and incorporated with the sound.

5.8 Conclusions

In this chapter, the two basic components of the conceptual model of the thesis project were presented and analyzed: the digital soundscape model and the digital geo-multimedia model. The latter originates from the sound representation of the visualization model, stepping on a series of sound attributes and criteria for the proper allocation of sound in the virtual space.

The demo application described in the following chapter, moves on the aforementioned methodological lines through the implementation of the different resolution levels between the digital soundscape model and the digital visualization model.

Chapter 6

Virtual soundscape implementation

Introduction:

The sixth chapter constitutes the implementation part of the research project, dealing with questions of more technical character. The methodological approach introduced in the previous chapter, (see pictures 5.1 and 5.2) is implemented through a demo application using behavioural software as it is described in the following paragraphs.
Through chapter six, questions of more technical character will be answered, such as how the physical characteristics of sound are technically translated into the soundscape- in order to render a realistic impression of sound -or on which logic the software is based upon for the implementation of the soundscape. In addition, the source and the sound format of the sound files to be used in the demo application will be defined, as well as the level of interactivity in the user's interface.
A further basic technical issue discussed in this chapter is the way that soundscape is linked and follows the sequence of the images during the exploration of the 3D world.
Finally, topics such as the format of sound files, the source of derivation and the selection criteria for the application, will be clarified through this chapter.

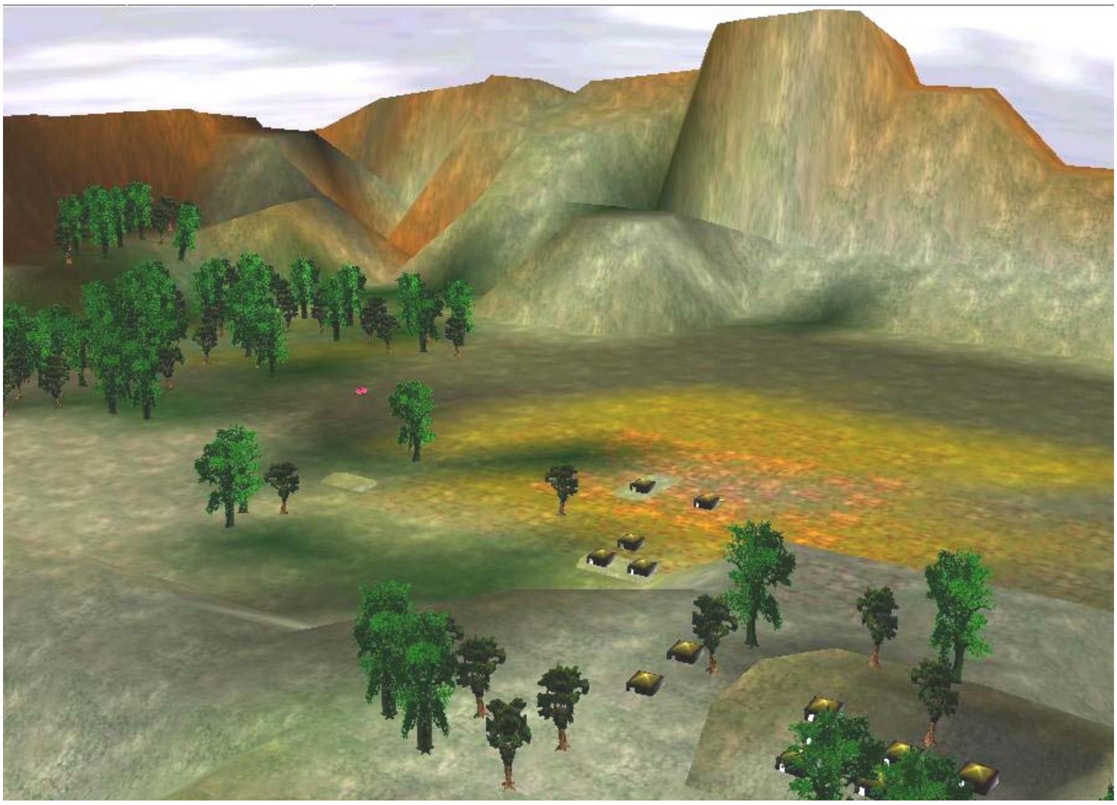
6.1 Describing the 3D virtual environment

Before proceeding with the description of the graphical and auditory components of the application, a brief description will be made on the software used for the applicative part of the project.
Behind the conventional techniques for 3D visualization –through conventional VRML viewers lie Cortona or Cosmo Player - and based on the concept of behavioural interactivity for real time 3D, Virtools 'dynamically immerses' users in an 'entirely interactive environment', with each 3D object acting and interacting according to its own specific behaviours. The term 'dynamical immersion' refers to the efficiency of the user to move, act and explore the 3D space in variant modes. In addition, the term 'entire interactive environment' expresses the ability of the virtual space's components to interact 'inter se' but also according to the user's reactions, though predefined behaviours.
Apart from the already existing object resources and behaviours -the so called "behaviour building blocks" or BBs- Virtools composition files allows you to download new objects, media files and interactive behaviours or even create your own behaviours blocks. (URL 4.13)
As an object oriented software, media files are treated as behavioural objects: supporting a great gamma of behaviour BBs that could be attached on them, but also the possibility of being imported using streaming technology, for improving performance and reducing the file size.
Regarding the demo application, previous to the soundscape composition precedes an analytical description of the 3D 'virtual environment' that presents the area of interest. The term 'virtual environment' refers to the 3D space as the end product of digital processing and various graphic representations –through visualisation software like arc view, Lightware, Maya , 3Ds Max - before been inserted to Virtools software. The visualisation of the 'virtual environment 'through the Virtools software –

either inserted from the Virtools resources or from an external database - constitutes the digital visualisation model that will be later represented through sounds.
The 3D environment that will show some of the soundscape concepts is based on existing Virtools demo resources. The 3D scene represents an uneven valley and a small forest wholly surrounded by rocks and mountains. Some group of houses situated over mounds are also part of the visual space. This 3D scene however, is not complete - at least not at the desired level in order to be associated with the virtual soundscape. Therefore a number of extra 3D objects and behaviours are added in order to be associated to even more variegated sounds.

A perspective view of the area

Picture 6.1



A character (human figure) is one of the first components added to the visualscape: it creates in every user of the application, the feeling of immersion in the 3D world, enabling the exploration of the visual and acoustic space through the eyes and movements of the character. Substantially this character represents the 'potential listener' that experiences the soundscape.
All characters in Virtools are actually 'blind and deaf. The use of different cameras gives the opportunity to see and hear the 3D space from different points and field of views. In this application there are imported 4 cameras: the first camera is looking the world through the character's eyes, two cameras are following the user from a certain distance and finally a free camera for quicker navigation in the virtual world. The advantage of the free camera respect to the others is that it allows a liberal navigation in all altitudes and places of the virtual world, but even out of the limits of the perceptive area. From one hand, the ability of this unlimited way of movement, undoubtedly gives a much quicker overview of the area. However, it presents a 'virtual' way of visual/soundscape exploration, which a potential user would never be able to follow in reality, not even assisted by flying equipment (e.g. airplane, helicopter or balloon)! From the other hand, the rest three cameras, continuously

focused on the character-walker from a predefined distance, serve a three Level-Of-Detail optical and acoustic zoom: a following mode that comes closer to reality, suspending the navigation however under the characters walking or running speed.
Making some interference in the visualisation model, by adding some dynamic objects - has as a consequence, the enrichment of the corresponding soundscape with sounds originating from the added objects. Especially when the moving state of the added objects is 'dynamic' - for example a moving vehicle or a river- then the proper rendering of their acoustic effect becomes a real challenge.

6.2 Sound features on three levels of visualization

Going through with the 3d virtual scene, one of the next tasks is to arrange the assortment of the sounds involved in the application according to the three optical levels of detail. Based on the methodology developed on the previous chapter, each visual level will be attached to the corresponding acoustic level. The basic tool to achieve such a kind of assortment is nothing else than the camera.
Using a number of behaviours available for behavioural objects (such as a camera), it is possible to keep each camera (that is constantly watching the character), in three different distances. The maximum detailed camera (LOD 1) is at zero distance, becoming "the eyes" of the character. The other two cameras, representing the second and third LOD, are placed in a distance of thirty and a hundred meters respectively from the character, maintaining the same direction and position settings –moving always on the same imaginable oblique line. During the character's movement, the last two cameras adapt a constant field of view- immediately or after a small delay-depending on the percentage of their following speed.
The selection and association of the sounds to the virtual environment is accomplished first in the 2D dimension and then in the 3D space. Using the Virtools display management functions, it is possible to obtain an extended zoom (zoom all) of the virtual world's top or perspective view. Assisted by the standard available cameras of Virtools (top camera and perspective camera) as well as the two cameras of the second and third LOD, a number of successive zooms in the 3D environment : the aim is to select the "potential" geographical objects and 3D entities that produce sound, either visible or non visible.
Following literately the methodological approach in point of the assortment of sounds, a raster surface is required to be applied on the different-range camera shots to facilitate the classification of sounds to each raster cell. Part of the Virtools 3D layout management tools, is the 'dynamic' grid. It is mentioned as dynamic owing to the ability of adjusting the shape and number of the raster cells, but even because it can be applied horizontally on the landscape's surface, independently from the selected camera's view. Using different camera's zoom, and applying an horizontal grid- with the same cell size but with the number depending on the resolution of the camera shot), it is achieved a classification of sounds respect to the visible objects belonging to each grid.
It becomes evident so far, that using Virtools is feasible the construction of a 2D sound map by following the methodology to the letter. In practise however the assortment of sound in the 2D- horizontal space is accomplished in a more empirical way. Especially when the examined area is more 'natural than urban' and the number of distinct sounds is limited, the use of a grid is not necessary for the positioning of sounds on the 2D environment: entities that produce sound become distinguishable and easy placed on the surface by zooming in and out in the 3D world.
As already mentioned before, in Virtools sound is considered a behavioural object. Talking with technical terms however, this is just a way of classifying the media components in order to assign specified behaviours –attributes on each sound.

The sounds in Virtools can be separated in two categories: point sounds and background sounds. Unlike the background sounds, point sounds have to be attached to an object in order to exhibit the source of origin. In reality, a sound cannot stand alone in a Virtools composition (as a usual object), without the support of a 'player' behaviour building block (BB). For a proper function of a Virtools composition, sounds have to be attached in (invisible) 3D frames. Consecutively these 3D frames can be attached to an object (sound source) or anywhere in the 3D space.
Using an object oriented software but following a raster oriented methodology could lead to hazes and confusion: the raster approach, as described in the fifth chapter, is not a mode of visualising the geographical features and 3D entities but just a matter of implementing the association between the visualisation model and the sound model.
For easier understanding of the methodological approaches, following in the next paragraphs, it is wise to make a quick reference over the 'operability' of three different 'classes' used in the Virtools environment: the objects, the frames and the grid.
Principally, all the above mentioned components are considered behavioural entities; consequently a number of behaviours (properties) can be ascribed on them. In addition, it is possible to be presented in either visible or invisible mode and change shape and size. The main difference between 3D frames and 3D objects is the way of appearance: the 3D frames are not part of the real world objects, therefore are used in invisible mode, as auxiliary means, holding a number of attributes (e.g. in the present application with sounds). Usually these frames are attached –placed within or close -to visible 3D objects, in order to ascribe the attributes they carry.
The 2D frames are also invisible frames that are placed on the user's interface, having as main task to visualise 2D objects or texts on predefined positions of the screen. The grid is a 2D surface, serving also as an auxiliary tool –thus not part of the real world- for the 'geographical parcelling' of the area in manageable squares (cells). Usually it is used in invisible mode, after allocating the objects and 3D frames in the proper cells.
When following a raster oriented approach- after the assortment of sounds in the horizontal grid - 3D frames are imported in the application and positioned over the corresponding grid cells (one frame per sound). The elevation of the 3D frames from the ground depends on the altitude of the sound source. A vertical projection of each frame on the grid surface would show the corresponding grid cell on the 2D sound map.
This seems an easy technique for sounds that belong to immobile objects, but for animated objects, it is wiser to avoid the raster oriented approach. Using a horizontal grid, the moving object could activate the over passing cells of the horizontal grid, which would succession ally activate the 3D sound frames belonging to each cell. However using this technique would never render properly the sound attributes as it would do using the object oriented approach.
Using the last technique, the sounds are also linked to 3D frames, and the frames sequentially attached to the 'sound source' objects, without interference of a grid.
It is important to note however that using Virtools, the sound files can also be directed linked to the objects without any intermediation of 3D frames. Using 3D frames however is easier to handle a sound. The 3D frames as mentioned above, are mainly used in invisible mode, something that facilitates the relative positioning of a sound respect to the visible object when needed (e.g. positioning a sound in a different elevation from an object so as to be easier perceived by a camera.)
Implicitly, afar from the methodological approach, working on the applicative part of the project, the selection of each technique depends on the number, the complexity and the type of sounds involved in the soundscape composition.

6.3 Sound criteria and categorization

Talking in technical terms, the applicative part of the project approaches the object oriented way, though the limited number of sounds and the presence of movable sounds, set aside the use of a grid.
Due to the limited sound samples in the Virtools resources, the majority of sounds files used in the application derive from sound libraries in websites. Virtools software supports most of the existing sound formats, including wav and mp3 format which are mainly used in the demo composition.
The selection of the sounds that will form the soundscape is not only depending on the components forming the visualscape but also from a series of criteria that influence a soundscape in a unique way. The chronic period regarding the season of the year, has an important influence on the soundscape 'appearance': the weather conditions prevailing on each period in combination with the 'vivid' entities that nominate a landscape each chronic season can greatly vary the effect of the acoustic space.
The chronic time of the day, is another important parameter that has to be defined since the variety, frequency and intensity of acoustic signals during day and night can also greatly diversify. The geographic characteristics of the area and the influence from human activities (urban- extra urban- natural environment) play also a decisive role in the composition of the soundscape.
The demo application refers to a rainy and windy day of spring. As already mentioned in the beginning of the chapter, the visualscape is outlined by a natural environment (valley), keeping however some artificial elements (houses on the banks) but also involving some human activities (car and airplane passage).
In the table 6.1 are presented all the sounds involved in the demo application, together with all attributes and characteristics as they are rendered from Virtools software.
In addition, in the same table fourteen sounds (of different nature) are analyzed and categorised based on the their natural properties and characteristics- mentioned in the previous chapter- but also depending on the way that are used in the application.
Starting from the first rows of the table, an assortment of sounds is attempted to the three levels of detail, depending always on how far they are perceived by each camera. Once again, it is wise to remind that the different levels of audio-visual detail are assessed by the three cameras pointing on the character. The user of the demo application has the possibility to switch between the cameras and observe (hear and listen) the virtual environment from three different points of view. The sounds marked with the symbol (-/x) indicate that their reception from the camera (LOD3) depends on the position of the character on the 3D space.
In the following rows of the table, are described some 'behavioural' characteristics of sounds and substantially the way they invoke in the application. All these behavioural characteristics are analytically described in the previous chapter, but here are briefly reported from a more technical side.
The way these sounds are behaving is not unique and can vary according to the composer's preferences. For example more than one sounds involved in the application could appear as periodic (recurrently in equal chronic intervals) but this would have a different –not so realistic- effect on the virtual soundscape production.
Some of the sounds in the table, such as the engine of the car or the noise from the footsteps, appear as continuous sounds: although they are not reproduced continuously throughout the application, they indicate a constant sound since their sound source is activated. It is also remarkable that the sounds of the immotile birds (standing on the trees) belong both to continuous and discrete sounds. This is because there is more than one carol (only for the immotile birds) characterised by a different duration.

Table of sound properties and assortment of sounds

Table 6.1

	animals					humans		nature				machines	
Sound samples													
Sounds	1-frog	2-Static birds	3-Flying birds	4-Cow	5-Duck	6-Footsteps	7-Voices	8-River flow	9-Wind	10-Rain	11-Thunders	13-Car	14-Airplane
Assortment to LODs													
1st level of detail	x	x	x	x	x	x	x	x	x	x	x	x	x
2nd level of detail		x	x	x	x			x	x	x	x	x	x
3rd level of detail		-/x	-/x						x	x	x		x
Properties-characteristics													
Random	x	x		x			x		x	x	x	x	
Periodic					x								
Continuous		x	x			x		x	x	x		x	x
Discrete	x	x		x	x		x				x		
Static	x	x		x			x		x		x		
Kinetic-Dynamic			x		x	x	x					x	x
Animation texture								x		x			
Sequence-order	x	x	x							x	x		
Duration (sec)	<3	<5	D	<5	<3	D	<5	D	<40	<50	<5	<15	<8
Coexisting with				All		All	All	All		11	10	All	All
Not coexisting with	11	9-11	9-11						1,2,3	1,2,3	1,2,3		
Natural	x	x	x	x	x	x	x	x	x	x	x		
Artificial												x	x
Visible source				x	x	x	x	x		x		x	x
Software settings													
Linked to 3D frame	x	x	x	x	x		x	x			x	x	x
Point sound	x	x	x	x	x	x	x	x			x	x	x
Background sound									x	x	x		
Doppler effect		x	x									x	x
Fader		x	x						x	x			
Intensity	x	x	x	x	x	x	x	x	x	x	x	x	x
Pitch	x	x	x	x	x	x	x	x	x	x	x	x	x
Frequency	x	x	x	x	x	x	x	x	x	x	x	x	x
Sound script/graph	x	x	x	x	x	x	x	x	x	x	x	x	x
Position on curve		x						x				x	x
Equalizer bar	x	x		x		x		x	x	x	x	x	

Two more noticeable points are the sounds of the flowing river and rain- rather hard to classify as static or kinetic- thus they are assorted in the category with animated texture.
The sequence is another variable, that shows which sounds have to follow a specific order for their reproduction. A number of thunders for example, could be heard before and during the rain, but not afterwards. The duration shows the maximum time each sound could last. The 'D' notation indicates that the duration of sound depends on the reaction or position of the walker in the 3D space.
Some of the sounds in the real world are almost never reproduced contemporaneously e.g. birds stop singing when it is thundering or raining. This characteristic of sound is expressed through the 'coexisting' rows in the table of attributes.
Determining which sounds belong to visible sources (objects) is more complex than it appears to be, always from a technical point of view. Working with visible sources, the 3D frame with the attached sound file is directly linked to the (visible) object that produces the sound. In the opposite case, when the sounds don't belong to visible sources, first it has to be clarified if it concerns a point sound or a background sound. In the last case, a 3D frame is needless, since the Virtools software will equally distribute the sound to the space, giving the impression of coming from nowhere (or is better to say everywhere). When it is about a point sound however, it is usually difficult to detect the 'epicentre' of the sound source, consequently a higher number of 3D frames should be used close to the invisible source so as to render better the sound effect. A representative example of both cases is the sound of thunder clapping: it can be used synchronously as a background and as an attached sound to several distributed 3D frames, giving the feeling of 'distance thunder clapping '.

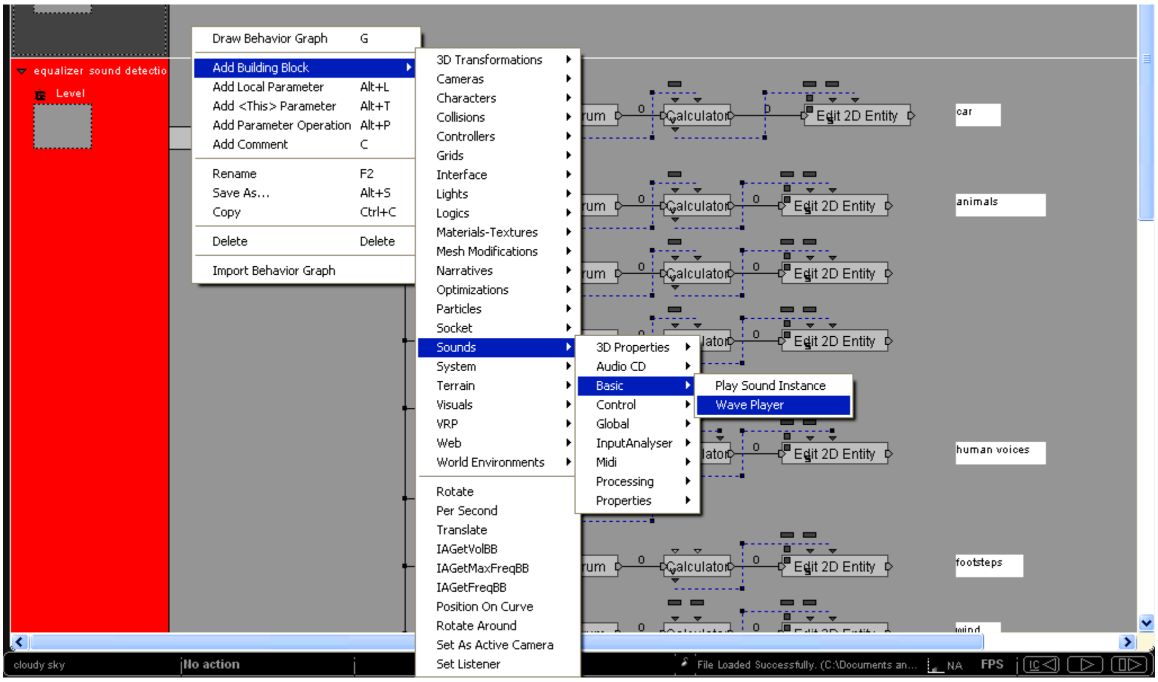
6.4 Virtools scripts: audio and visual improvement

In the final rows of the table 6.1, are listed those sound properties that are adjusted and rendered by Virtools software. Apart from the 3D frames and point-background sounds discussed already above, there is a number of 'natural properties' of sounds attempted to be simulated by Virtools software.
Intensity, pitch and frequency are some of those natural properties than can be adjusted from the sound properties menu of each imported sound. In addition, importing some behavioural building blocks (BBs) some more sophisticated sound characteristics are attained, such as the Doppler effect (for objects with a considerable speed) and the fader effect (e.g. for scalable decrease of the rain sound and successively increase the sound of bird singing).

Apart from the standard sound settings, it is possible to import several behaviour building blocks from the Virtools 'behavioural library' and combine them in a unique way, creating in such a way numerous comporting scripts (pic 6.2).
The majority of the 'behaviour scripts' are dedicated to the proper positioning and reproduction of sound. However there are also a number of scripts concerning the visual improvement of the 3D virtual space. Considering all the behavioural scripts are contained in the appendix, there will follow only an epigrammatic description of the most important ones.
The constant movement of the clouds and the flowing texture of the river are two representative examples of scripts, aiming to create more realistic visual effects. Some of the behaviour scripts however, can both contribute into the visual and acoustic enhancement of the 3D space: creating a number of invisible curves and using a series of behavioural BBs can enforce any object or 3D (sound) frame to move on the curve with the desirable frequency and velocity. A characteristic example found within the demo, is the movement of the car (or the airplane) on a pre

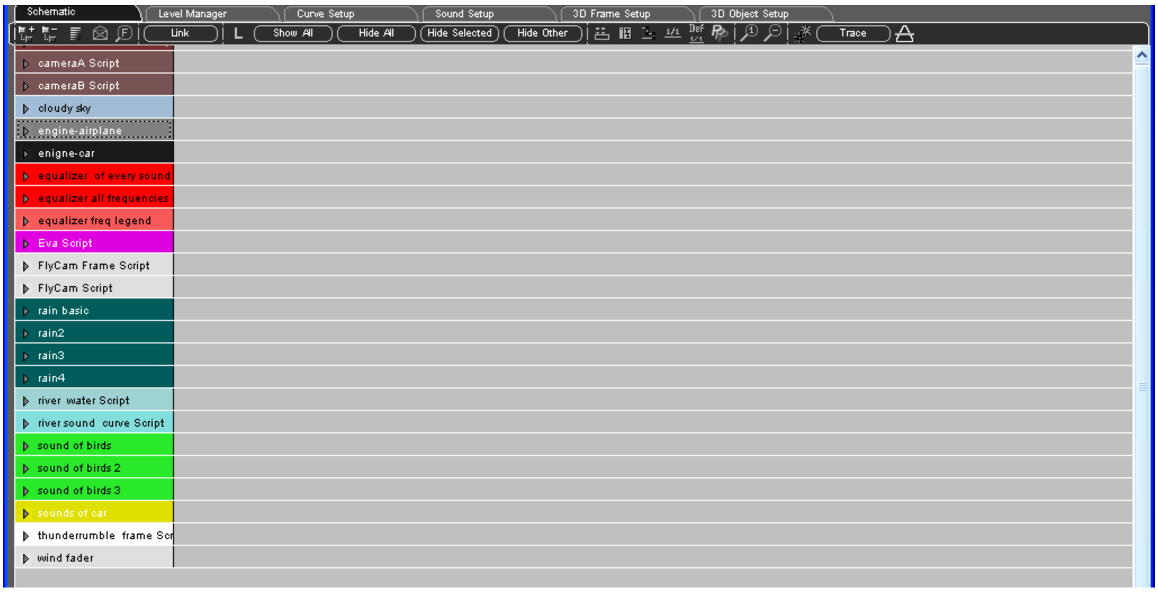
incised curve, having attached a 3D frame which is further linked to a car (or airplane) engine sound file. Curves can be also used for non visible sound sources or when is difficult to detect the epicentre of the sound and want to give a certain direction to the sound. For the flying birds in the forest, a number of 3D frames with different sound files attached to them render a more realistic carol effect. In the flowing water case, a number of 'water sound files' , attached each to a 3D frame, are sequentially following a 2D curve along the river's surface, giving a more dynamic perception of the water's movement.

Picture 6.2



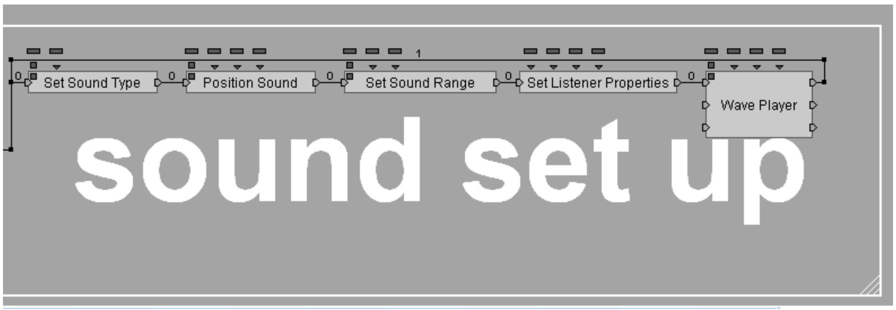
All the sound files imported in the demo composition- as well as all cameras and most important objects and functions- own their personal script that includes single behaviour BBs or combination of BBs (behaviour graphs).

Picture 6.3



For the background sounds of the application, the use of a simple 'player behaviour block' is enough to reproduce it properly. For the 'point sounds' however, the use of appropriate Behaviour blocks and Behaviour Graphs is essential for the right 'pick up' from each of the four cameras-listeners. A simple example of a behaviour block array, -responsible for the adjustment of each sound- is presented in the following picture (5.3).
More analytical descriptions and samples from Virtools behavioural scripts are contained in the appendix.

Picture 6.4



The simulation of flashes and rain is a spanking part of the composition; not only owing to the visual effect but also for the chronic coherence and randomization they are presented. The use of particles attached to series of invisible 3D frames, creates a prominent rain drop effect but this is just the beginning: the sequential or random order of rain when combined with other audio-visual components (flash light, thunder sound, bird's carol, rain sound and others) is accomplished through the use of logic and narrative behaviour blocks. These BBs can activate or deactivate scripts or send/receive messages to/from other scripts in predefined or arbitrary time intervals resulting in even more realistic audio-visual renderings.

6.5 User's interface- interactivity and navigation

Working on the user's interface, two graphic equalizers are created in order to provide extra information for the reproduced sound. Task of the first equalizer, is to measure the medium frequency of all sounds involved in the application, in the following frequency intervals: 0-60-170-310-600-1000-3000-6000-12000-14000-16000 Hz. The second equalizer detects every type of sound that is active in the frequency interval 0-16 KHz. In the last equalizer, sounds that belong to similar 'cluster of sound sources' (e.g. animals or birds) are incorporated into the same graphic bar.
On the user's interface, there is continuously displayed information about the selective-active camera, as well as the distance of the free camera from the walker. Generally the graphical user's interface is maintained as simple and loose as possible: besides, main scope of this demo is to create a totally immersive feeling in the 3D world, without distracting the user attention with 2D/3D web links or pull down menus. Although Virtools is oriented in continuously more interactive environments, the impendent application is purged from highly interactive options: beyond the basic control of the character's movements and the selection of the active camera, the user is liberally co opted in the exploration of the virtual soundscape.
The Virtools software, apart from the Virtools Composer, disposes also its own VRML player which can be freely installed at any pc to facilitate the reproduction of any Virtools composition. Exporting any Virtools composition to the corresponding player

not only reduces the file size but also improves the refreshing frame rate (display refreshing quality) during the reproduction of the file.

6.6 Conclusions

Based on the methodological steps incised in the previous chapter, the implementation part of the project was presented through the composition of a soundscape-lying within the geo-multimedia model: a model that originates from geovisualisation of geodata. In this geovisualisation model a sound attribute layer is added via the use of behavioural software.

The idea of the three audio visual levels of detail is accomplished through the use of different positioned cameras. In addition, for the assortment of sounds –first to the 2D sound map consequently to the 3D space- there was no need of using the grid methodological approach owing to the simplicity of the 3D environment and the limited number of sounds.

Chapter 7

Remarks and conclusions

Introduction:

The final chapter constitutes a general report, over the previous chapter of the thesis project. Some of the concepts or research questions that were not sufficiently illuminated are fatherly discussed in the following paragraphs. This chapter can be mentally separated in three parts: in the first, there is a reference to the most significant points that were met in the first five chapters. Particular anaphora is made on the geoinformative character of sounds, the digital soundscape model and the methodological approaches for the composition of a virtual soundscape.The second part, is more technically oriented and mainly dedicated to the functionality of the multimedia model as described in the sixth chapter. Finally the third part constitutes a 'force account' of the thesis work, giving synchronously some recommendations for further development of the present report.

7.1 Comments of conceptual character

Beginning from the area of interest, substantially it does not exist in reality but as it is already mentioned in the sixth chapter, the visualisation model refers to an imaginary 3D environment inserted from the resources of Virtools. Implicitly the composer doesn't have the opportunity to experience the area on 'sight': based on his own (aural) experiences and assisted by the Visualisation model he manages to scan through the 'virtual' area of interest for the allocation of sounds. Evidently it's not possible to use of both the two techniques for the detection of sounds,-as they are described in the fifth chapter- since the virtual world doesn't correspond to a real one. This doesn't mean of course that the virtual soundscape falls apart of those characteristics of real world's soundscape. Besides, the proper simulation of a sound in a virtual mode is highly depending on the general experience of the composer from natural or urban soundscapes.
Generally the creation of a virtual soundscape – as an abstract composition - is more complicated than the visualscape that belongs to. This is the basic reason that justifies any deviation from the predefined methodological approaches (which determine the structural design of the soundscape). Looking on the implemental part of the thesis project, it becomes obvious that both of the approaches for the formation of the digital sound model and the digital multimedia model are not literally applied. As previously mentioned, the 'localization' of sounds and the modulation of the sound properties that constitute the digital sound model, are not 'dynamically' experienced in the real world, resulting in a divergence from the predefined methodological 'path'.
An analogue deviation is observed in the implementation of the digital geo-multimedia model. The raster based approach -defined in the fifth chapter- for the assortment of sounds on the 'squarely framed' geographical surface, is not applied, giving way to a rather object oriented approach. This is not owing to the inapplicability of the raster approach, but due to the simplicity of the visualization model: the limited number of sounds – corresponding to each cell of the impendent grid- combined with the level of optical accuracy in commission) , are favouring the use of a less sophisticated technique.

Another important issue also mentioned in the third chapter, represents sound as a part of the geographical data of an area. The question that has to be further illuminated is whether sound, as a feature of an entity can be comprised as a part of geodata. It is obvious that georeferency cannot constitute the only criterion for classification of sound as a part of geodata, considering the abstract character of sound. Consequently, criteria of more generic nature, such as the occurrence of a sound within a geographical area, or the importance of geoinformation that it provides, could comprise two of the specifiable factors.
Considering 'georeferency' as the 'ultimate' criterion for determination of geodata impels the study of sound effect from different perspectives. For example considering the sound around an airport, as product of movable sources (airplane) is not possible to georeference it. Working with buffer zones however, that show the potential influence of the airplanes' sound around the airport –after an extensive recording field work- can result in a georeferency of the ambient area. An analogue example refers to the georeferenced sound pollution (generated around the traffic zones). Substantially, the impendent approach for sound georeferencing, reminds the discrimination of entities in discrete and continuous.
From the aforementioned it becomes obvious that georeferency doesn't depend only on the character of a sound, but also from the methodological approach on the sound effect. Beside the georeferency of a sound, it is also important to study the 'temporal reference' of a sound.
When there is talk about georeferency of a geodata, implies translation of position in 2D coordinates, but also parameters like type of ellipse and projection. Sound however, as a more abstract concept respect to image, is even more difficult to represent graphically and follow the conventional rules applied for visualization of maps and images. Graphic representation of sound through 2D or 3D maps is beyond the scope of the thesis project, Nevertheless it is accepted that sound as geographical data, it is possible to be graphically represented, depending however on the nature and state of the sound source (as it is further described in the third chapter).
Duration of sound, in a real world environment, is a rather sophisticated characteristic, rough to simulate through a virtual soundscape. By experiencing the duration of each sound in real world and by making a series of assumptions (setting time limitations in the reproduction of different sounds), it is achieved the required 'time compression' when passing from the real to the virtual soundscape. The simulation of 'time' within a virtual soundscape also depends on the duration of the geo-multimedia model's reproduction. This is mostly concerns sounds accompanying natural phenomena (for example rain or wind): the longer the multimedia reproduction, the more time that is dedicated for 'continuous' sounds within the virtual application.
Regarding the fourth chapter of the thesis report, it doesn't follow 'faithfully' the conceptual coherence of the previous chapters about sound: it has a rather technical character, predisposing the reader for the dispensable technological tools, some of which will be used in the implemental part of the thesis. The reference on the development of sound in the videogame industry, doesn't study the structure of the 'virtual soundscapes' within the games but emphasizes the persistence and 'each time demand' for enhanced sound quality.

7.2 Comments of technical character

The number of sound attributes that are simulated within the demo application is not depending only on the capabilities of the software in use, but also from the geo-diversity of the selected area. The visualization model, the way it was primarily imported from the Virtools resources, represents a rather 'loose' natural environment, without however any occurrence of 'dynamic natural phenomena (e.g. rain, thunder, wind) or elements (e.g. river)', and with an absence of any human activity(e.g. humans, machines). An evenly 'dynamic' intervention in the natural landscape for the extension with those 'dynamic' features would equally enrich the corresponding soundscape. In other words, the visualization model is 'dynamically' improved in order to be (later) acoustically represented by a greater gamma of sound characteristics and physical properties.

Indeed the introduction of particular objects, whose moving speed is fairly higher respect to the characters moving velocity, allows the simulation of more sophisticated sound effects, which express the human impression of sound experiencing. The Doppler Effect is perhaps the most representative example, which gives higher detail in the sound simulation of the moving car or the airplane of the demo application: if the listener and the sound buffer are both moving, the system automatically calculates the relative velocity and adjusts the Doppler Effect accordingly.
Apart from the Doppler Effect, other characteristics of sound related to the human perceptibility and experience of sounds, are the harmonics and the roll off effect. The roll off effect is easily implemented within the application, rendering ideally the attenuation of sound according to the 'remotion' of the character from any sound source. Regarding the 'harmonic' effect of sound, although is not supported by the current 'behavioural blocks' residing in the Virtools resources, the composer has the possibility to create and introduce new 'behaviours' written in C++ programming language. However, even if it is possible to simulate characteristics like harmonics or echo, the geo-morphological characteristics of the surrounding area, don't permit the existence of those types of sound effects.
Simpler characteristics of sound such as sound intensity or pitch are handily adjusted through the sound settings and by the use of already existing behaviours. All the sounds inserted in the application, have a 'global 'effect' on the surrounding soundscape. Specifically, it is assumed that the sound originating from each sound source is propagated to all directions- acting on a spherical range around the source- without the use of (more complicated) 'sound cones'.
Apparently, other visualization would be acoustically represented by sound models that emphasize different characteristics of sound.
The technical capabilities of sound simulative (or behavioural) software are not the only criterion to determine the efficient reproduction of a geo-multimedia model. The use of high fidelity audio systems plays also determinative role for the qualitative rendering of multimedia product. Especially when the sound files of the multimedia model have been digitally processed- for separation of sound in different audio channels- then the use of digital audio processors and proper loudspeakers, would impressively improve the sound results. That was, besides, one of the main purposes of the fourth chapter, that is the design and presentation of virtual soundscapes by efficient use of the existing technology.
The simulation of sound physical properties is complemented by a fully interactive user's interface, where the viewer has the possibility to visit any desirable extent of the 3D space and receive a respective 'sound feedback'.

Making a quick evaluation on the software used in the applicative part of the thesis work, it appears that Virtools is one of the most suitable tools for simulating 'reality'. Apart from the composition of the virtual soundscape, Virtools developing software contributed in the enhancement of the visual 3D environment by adding 3D objects, improving textures and creating animations. In addition, as behavioural software, Virtools adds an extra component to the visualization model that cannot be rendered by a simple browser: the dynamically changing phenomena or conditions. The simulation of weather changes (and a further extension by the corresponding sounds) is a representative example from the demo application.

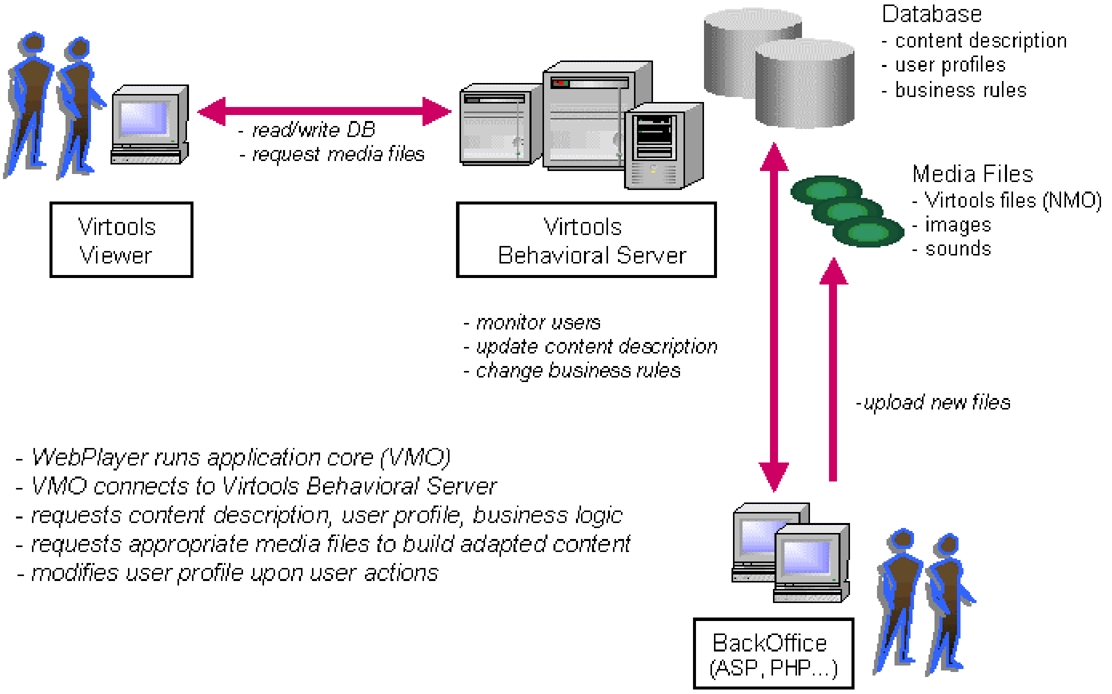
However, shifting between 2.1 and 2.5 Virtools Developing versions, it generated compatibility problems in the proper rendering of the soundscape, resulting in the remake of the application from scratch.
Besides Virtools software, there were also used Adobe Photoshop (for the formation of pictures and textures) and Sound forge (for registration of voices, processing of sound files and conversion of data in mp3 format).

7.3 Users and utility of the geo-multimedia model

The applicative part of the thesis project, in the form of the Virtools application, is designated for presentation to a limited number of spectators. As long as the application is not accessed through an internet site but is distributed via portable discs (cd), there is no question regarding the size occupied by the sound files within the demo file. Especially the use of mp3 sound format, supported by Virtools, minimizes even more the final size of the demo file.
Nevertheless, the remote access of the geo-multimedia model through a web page, would neither affect negatively the proper reproduction of the application, seeing the advanced streaming technology of Virtools. Thanks to Virtools Behavioural Server's unique streaming technique, end-users don't have to wait for complete download before starting to visualize and interact with applications.

Virtools Behavioural Server guarantees fast immersion in 3D environments enriched with behavioural interactivity, maintaining synchronously a sufficient 'refreshing rate' of the 'image frame' display. (URL 4.13) Any sound or media file, is 'dynamically' loaded from a remote database, which is can be updated at any time, as shown in the next scheme (pic 7.1).

Picture 7.1



Considering the commodity of accessing the application in a remote mode-using advanced technology- it becomes obvious that possible publications of the digital multimedia model though the internet, could constitute a remarkable way of evaluating the model.
The composition of the digital multimedia model, through the implementation of the demo application, addresses to a broad number of potential users. Users from all the fields of utilization of geoinformation (e.g. policy makers, tourists, municipality and prefecture, land planning users, environmentalists, and tourists) can take advance of the extra information that provides the multimedia model via sound.
The geomultimedia model, however, accosts also to a distinct part of users, for whom any kind of real or virtual soundscape, has a special significance for the perception of the surrounding space: visual disabled people can really estimate the contribution of a virtual soundscape since in this case acts as the only auxiliary tool for the perception of the simulated area.

7.4 General overview- recommendations

Making a quick overview over the thesis work, the general goal of the research project is apparently achieved. The extension of geoinformation by sounds and the acoustic representation of visualization model through a virtual soundscape are verified by looking on the demo application.

Examining the geoinformative character of sounds is an **innovating** procedure but with very **broad sense** as well. Consequently requires explicit definition of the encountered concepts and continuous specification of the relations within the conceptual models in order to avoid generalities and hazes.
Some of the concepts or principles that are met during the thesis project can be differently interpreted by every researcher. In the WDVM and WRM conceptual frameworks, presented in the fifth chapter, are defined respectively sixteen and twenty five relation ships, but only few of them are selected and analyzed. Obviously this depends on the components of the model the researcher aims to emphasize but also from the methodological approach that is followed. Particularly, there have been highlighted and further developed the relations between the visualization, the soundscape and the multimedia digital models. In addition, this thesis project could be extended in different directions by focusing on relations between the aforementioned models and the real world or the human perceptibility of each (of these) models.

Further alternatives for **extending the thesis work** can be either through the soundscape structure approach or the use of different levels of detail.
Regarding the methodological approach for the assignment of sounds on the 3D virtual space, the use of a raster (grid) –the way it is described in the fifth chapter- appears to be applicable for more ‘complex’ sound and visual scapes. The selection of a more complicated environment (e.g. urban environment) combined with an even higher resolution visual-and audio- level of detail, could be ideally assisted by a grid layer approach.
As a first attempt to represent different levels of detail within a multimedia model, the number of audio-visual levels of detail (and the way they are related between each other) was low and simple. Extending the model with a higher number of LODs and by using different ‘correspondences’ between the visual and audio levels (not 1 t0 1), it is possible to exaggerate and consequently to emphasize some of the aural characteristics of the real world’s environments.

Referring to the implemental part of the thesis project, the use of four cameras and a character (avatar) is just a mode between numerous ways of navigating in the 3D space.
Depending on the characteristics of the virtual space to be accentuated, every composer could use his own personal mode of exploring the virtual 3D space.
Discussing about the assumptions made in the project, a consideration of sound as an attribute of the real world's entities and phenomena, leads to respective methodological approaches and implemental techniques. This doesn't implies that possible consideration of sound as 'object' of the real world-by other researchers-would lead to hazes or invalid results, but the relations between sounds and the real world's entities would follow a different approach.
Finally, accepting sound as a part of geographical data (which presents some natural properties and characteristics) associates a number of different approaches for georeferencing sounds: undoubtedly a topic with very broad sense that could constitute a different topic for thesis research.

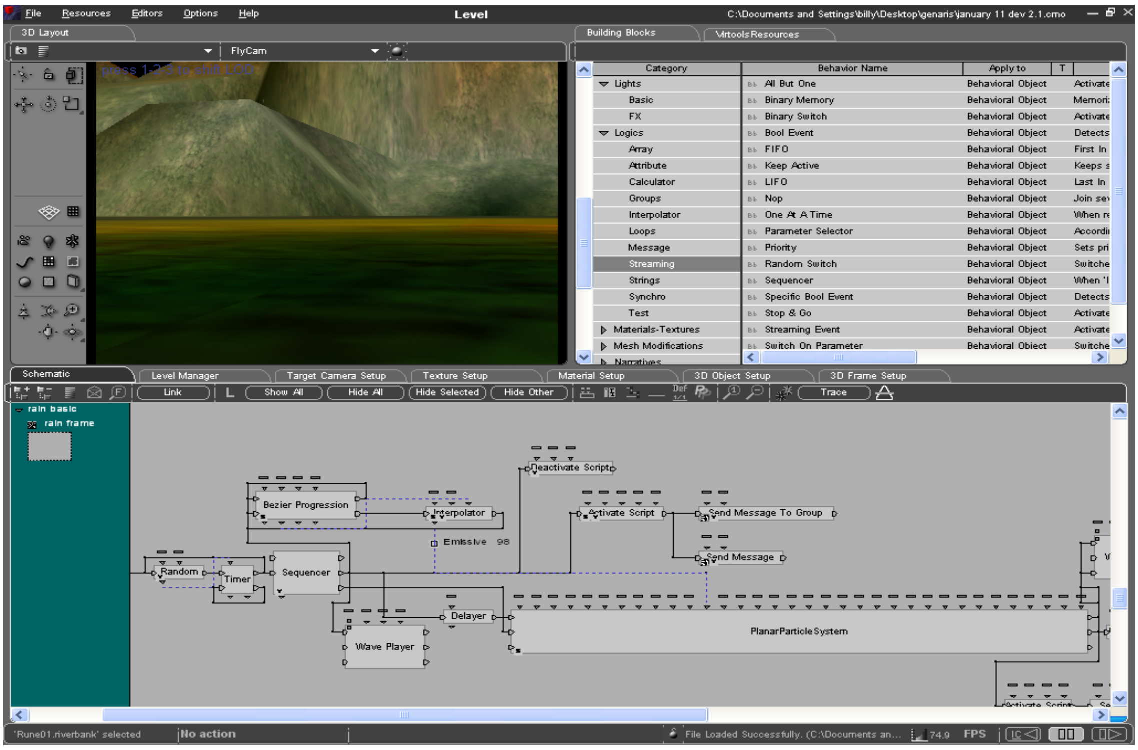
Some of the most significant or ambiguous points reported in the previous chapters, are elucidated in the paragraphs above. Regarding the implemental part of the thesis project, more technical examples are illustrated in the appendix of the report.

Appendix

Examples from the implemental part of the thesis project developed by Virtools software

Images and scripts taken from Virtools demo application

Picture A1



The Virtools Developing software interface

Picture A2

Top view of the area. The red perimetric parts are the mountains that surround the valley (of interest).

Picture A3

First Level of Detail (low)–camera A

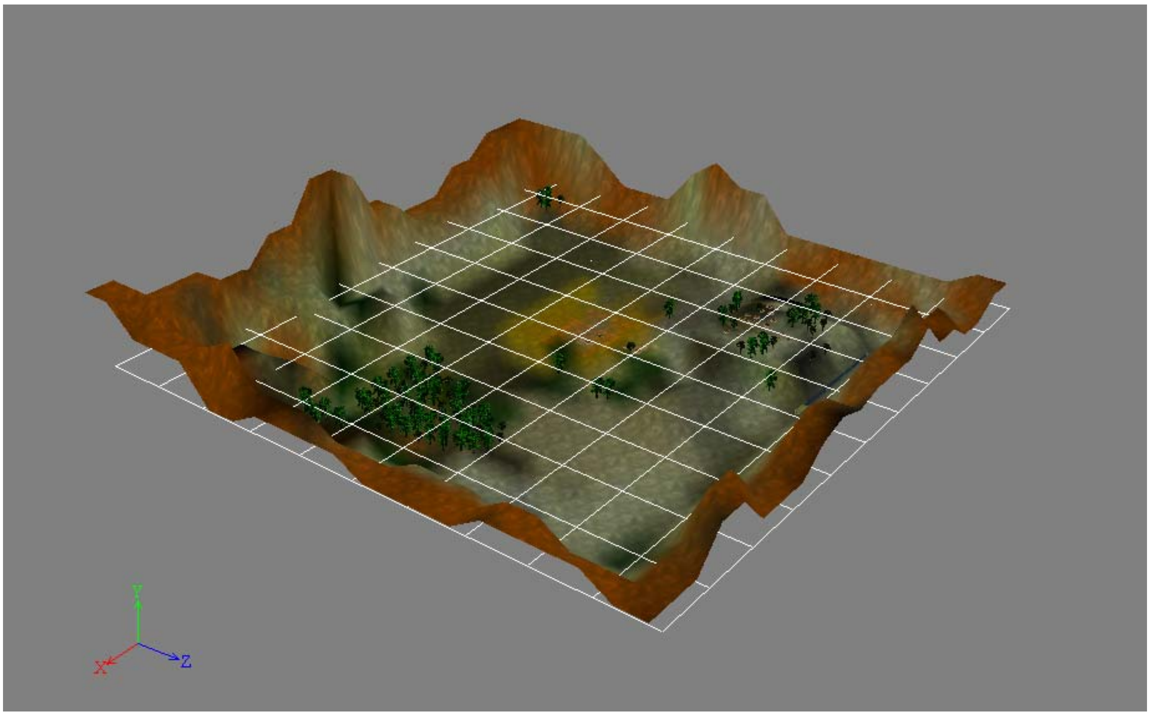
Picture A4

Second Level of Detail (medium)– camera B

Picture A5

Third Level of Detail (high)–camera C

Picture A6

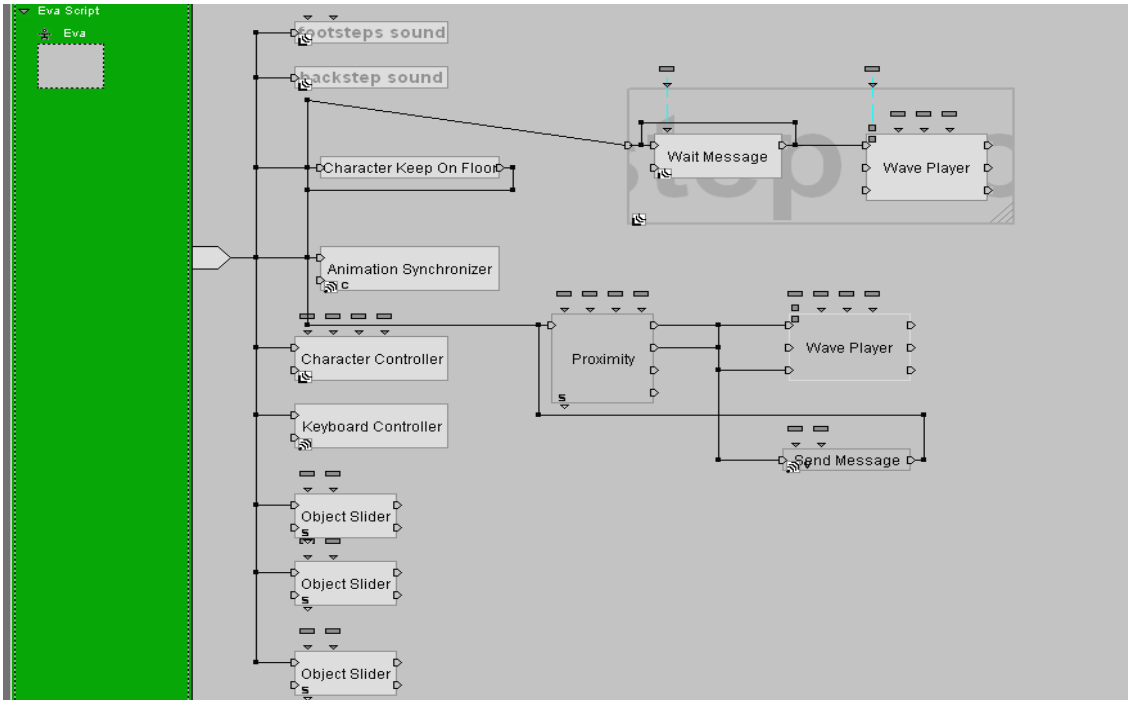


Application of horizontal grid on the area of interest. It is possible to adjust the dimensions of the grid, the number of cells, the cell size and the position of the grid respect to the area (height, orientation).

Picture A7

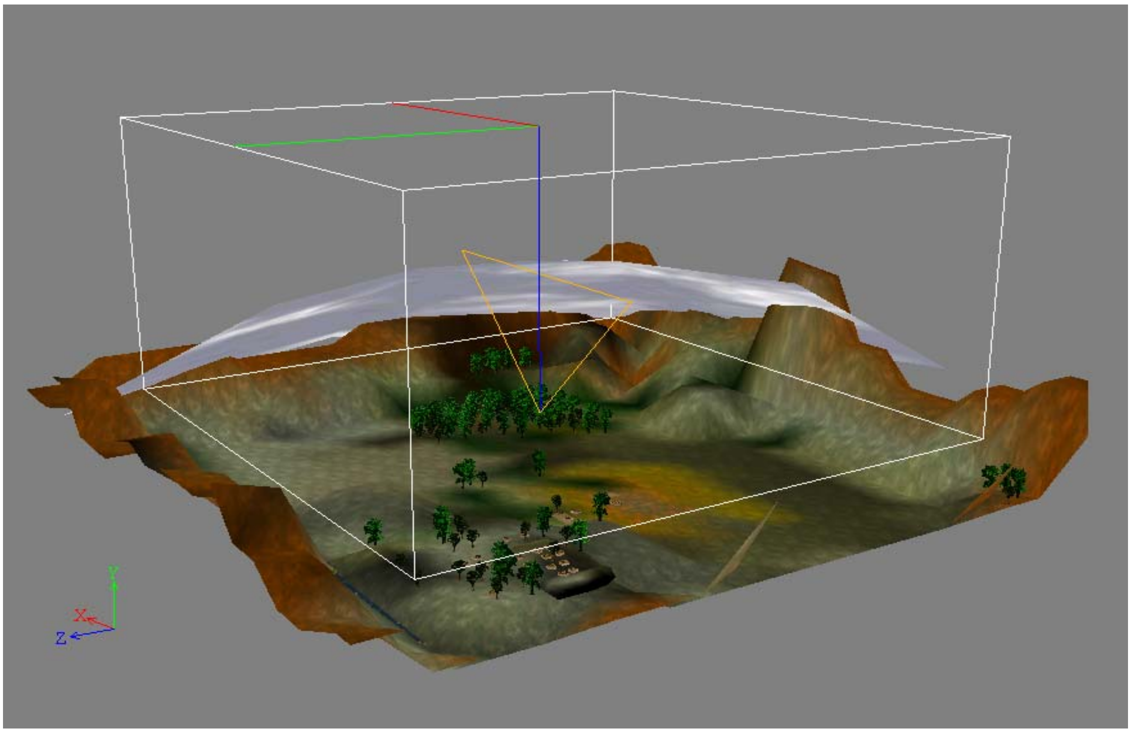
A character (avatar) where the cameras are focused on.

Picture A8



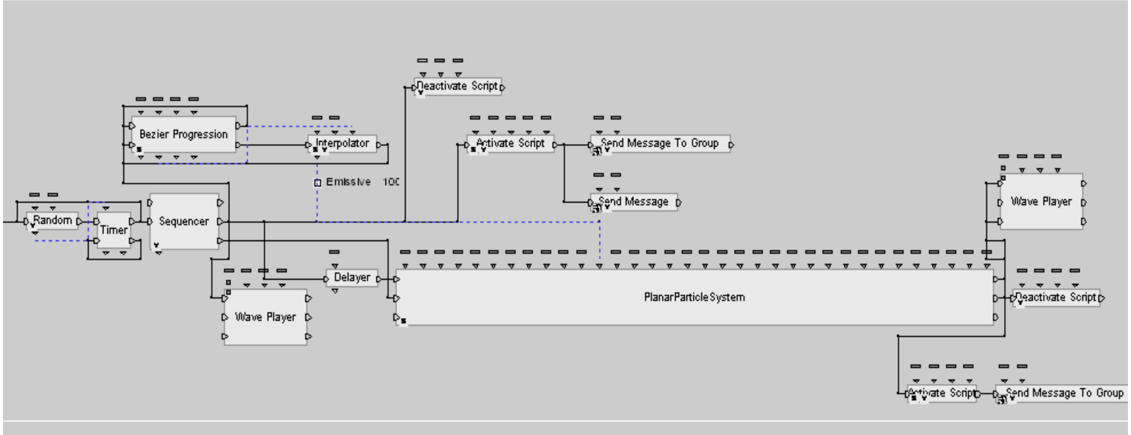
Script of character control and behavior. Apart from the keyboard controllers, it simulates the character's sound footsteps, and the human voice when approaching a second character in the 3D space.

Picture A9



One of the 3D (rain) frames. In this 3D frame is attached a particle system that causes that produces the rain drops. There is no need to attach a 'rain' sound file in this frame since rain is considered a 'background sound'.

Picture A10



Script of rain frame set up. Before the rain particles begin, a message is send to the singing birds to pause. After the rain stops, with a second message the birds start singing again. The two wave player Building Blocks involved in the script serve for the sounds of rain and thunders. A lighting effect could also be involved in this script. The timer and sequencer BBs serve for the random appearance of the rain effect.

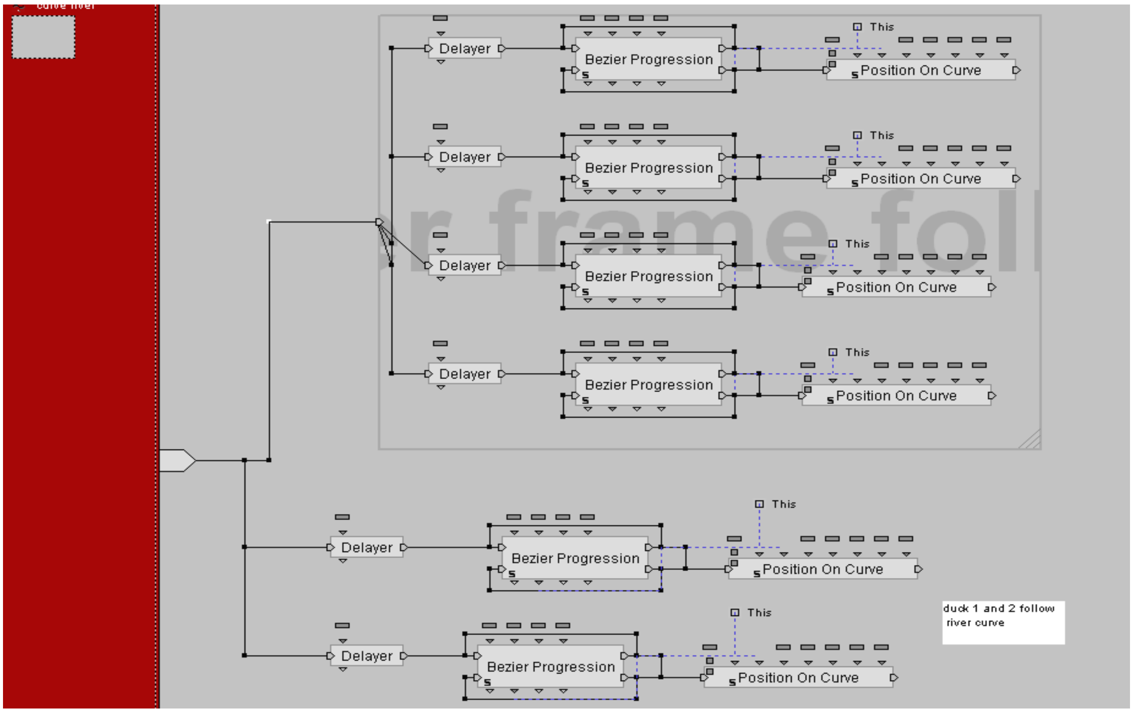
Picture A11

Two 3D frames following the forest curve. In each frame are attached different sound files of singing birds giving the sense of woodnote in the forest.

Picture A12

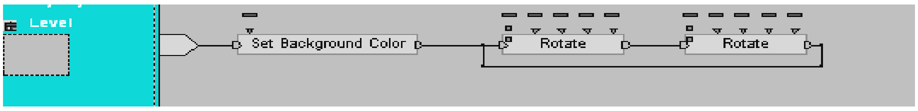
A river 3D frame used for sound simulation. The white cube frame containing a cross (3D frame) follows continuously the river curve (red line) in order to simulate the water flowing sound. A flowing water sound file is attached to each of the four river 3D frames. The same ‘philosophy’ is used for the 3D frames of the moving vehicle, the airplane and the singing birds in the forest.

Picture A13



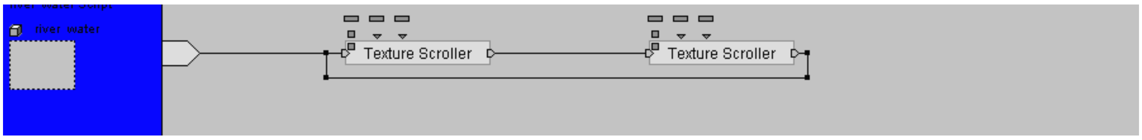
Script of river curve: followed by four 3D frames (attached to river sound files) and by two objects. Similar scripts are also used for objects or 3D frames that follow a curve.

Picture A14



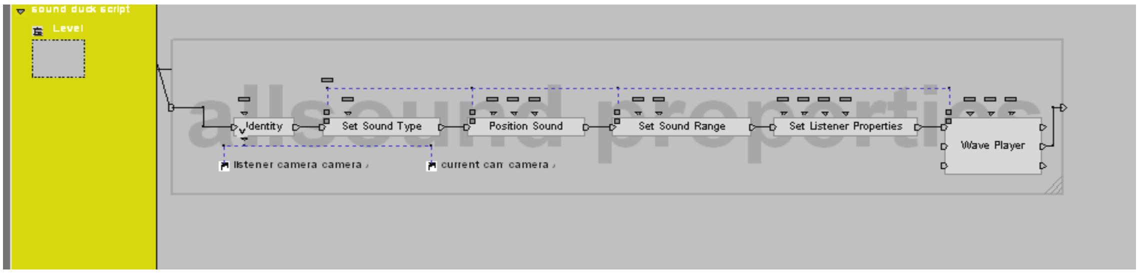
Script of environmental improvement: adds animation to clouds

Picture A15



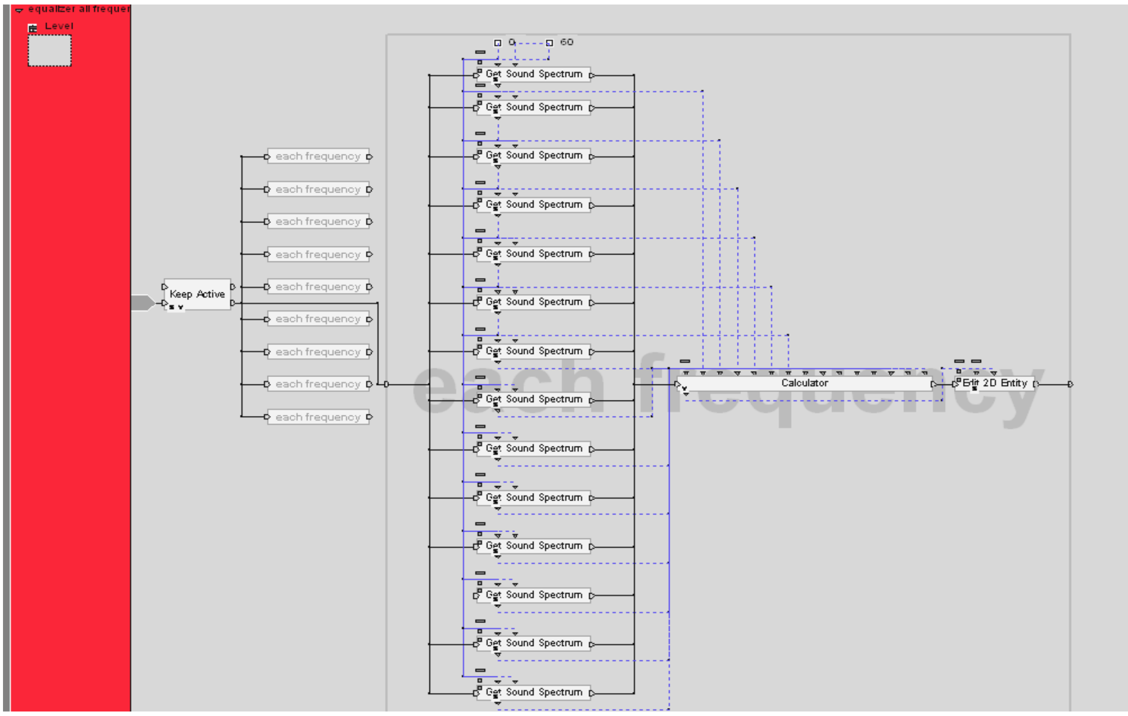
Script of environmental improvement: serves for simulation of the river's flow.

Picture A16



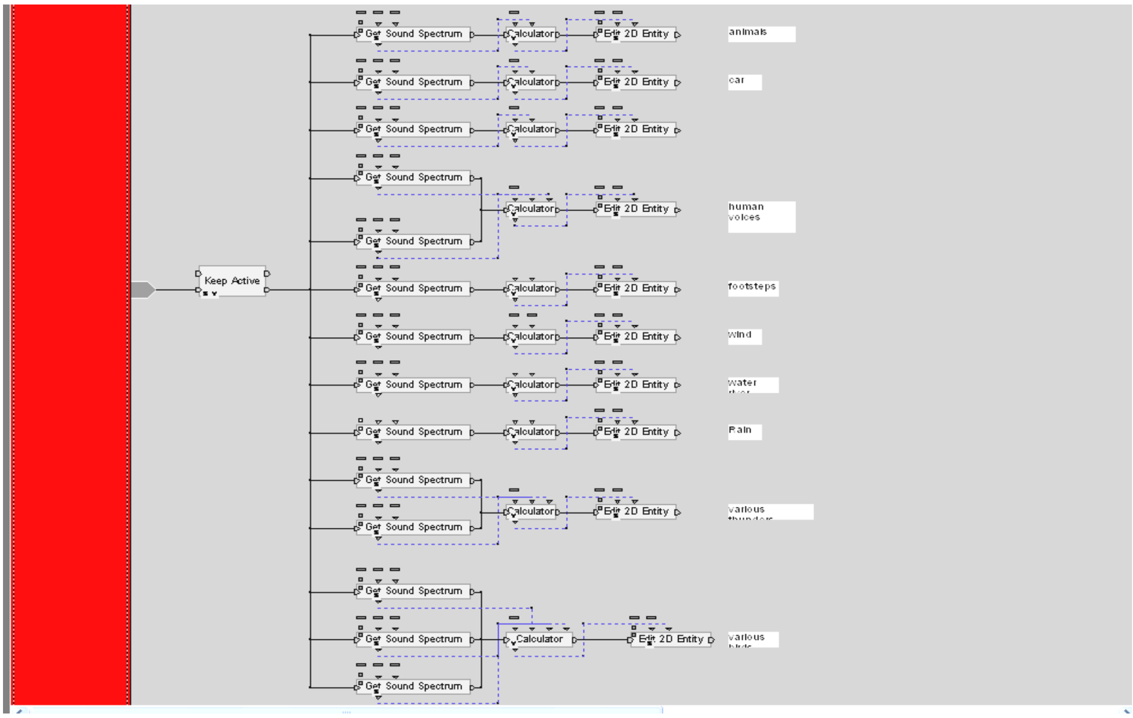
Script of general sound set up: used for adjusting the properties of point sounds, (either static or kinetic) such as the sounds of animals, human voices, machines etc. Adjustments regard the sound type(point-background), position sound, sound range, listener properties and wave player.

Picture A17



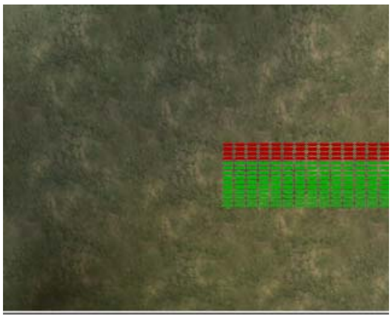
Script of frequency equalizer. Represents the medium frequency of all sounds within the intervals 0-60-170-310-600-1000-3000-6000-12000-14000-16000 Hz.

Picture A18



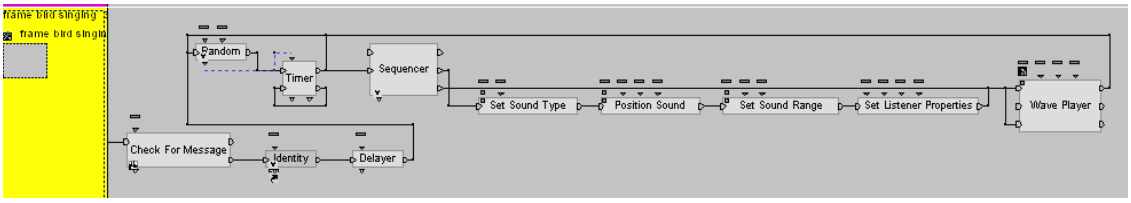
Script for sound detection. Similar to equalizer, gets the sound spectrum of every sound and translates it graphically on the users interface showing which sounds are currently reproduced.

Picture 19



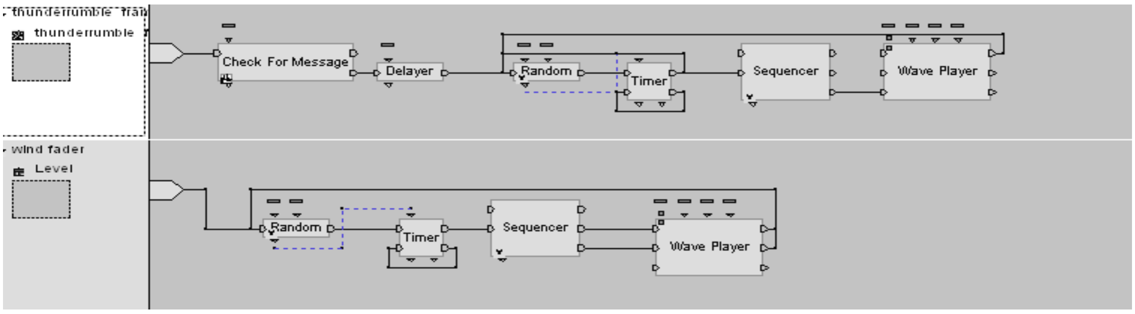
A 2D frame on the users interface visualizing a Graphic equalizer

Picture A20



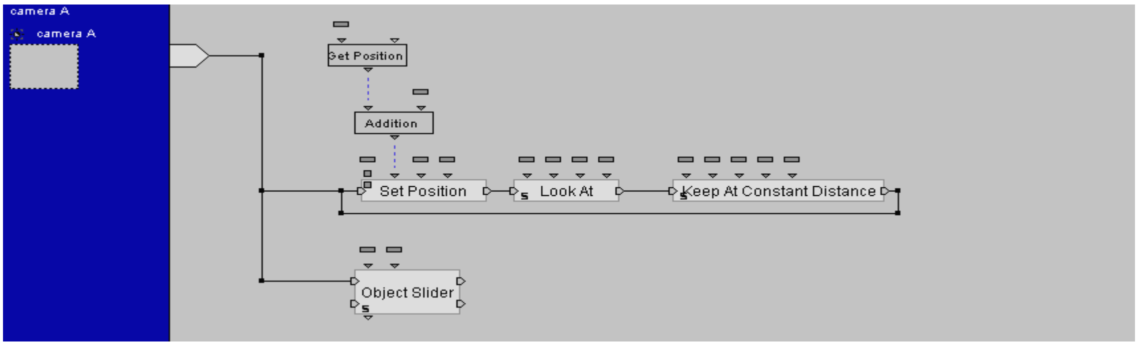
Script of singing birds: alternatively activated and deactivated depending on the weather conditions. The activation is based on a message sent by the rain script

Picture A 21



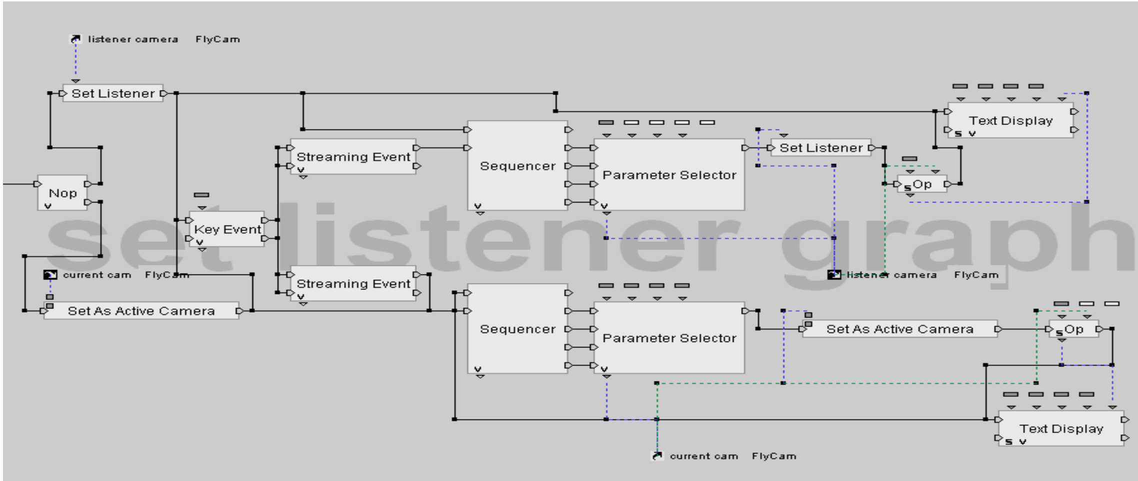
Scripts of background sounds: thunder and wind. The sound of thunder is depended on the rain script, thus is constantly checking for a ‘message’. The sound of the wind appears in random intervals.

Picture A22



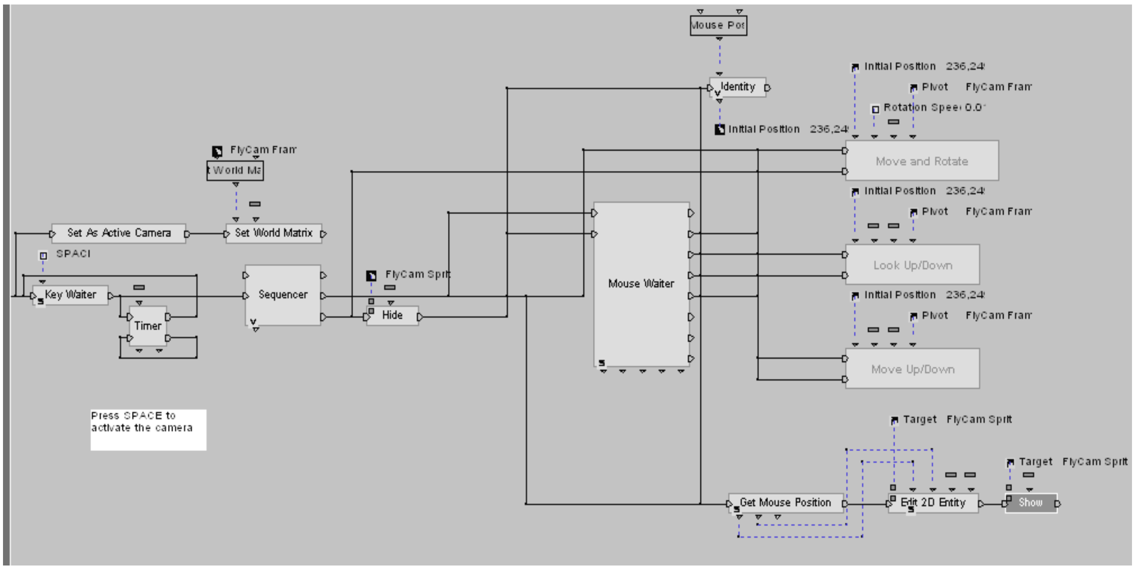
Script of camera A. The same script is also used for cameras B and C, setting however different focusing distances from the character.

Picture A23



Script of the camera audio listener switcher: using the same button of the keyboard the user can see and synchronously listen from each of the four cameras.

Picture A24



Script of the Fly camera (available in Virtools resources): allows the free navigation in the 3D space without any constraint to follow the character-avatar. The user can experience every point of the soundscape by listening through the fly camera.

References

Chapter 1

Bourdakis Vassilis 2003, ‘Modelling and City Planning’ (report)

Camara S. Antonio, Raper Jonathan /1999, ‘Spatial Multimedia and Virtual Reality’

Fisher Peter, David Unwin/2002, ‘Virtual Reality in Geography’

Halskov Madsen Kim /Feb-2003, ‘3D Sound in 3D space’, University of Aarhus, Denmark (URL 1.2)

Kobayashi Minoru**,** Schmandt Chris/Cambridge 1997, ‘Dynamic Soundscape: mapping time to space for audio browsing’

Pair Kim, ‘Virtual Environment Sound Design Primer’

Thorn Richard/ March 1996, ‘Virtual Reality: A Sound Proposition’: Paper for ‘Hearing is Believing 2’

Tonnesen Cindy and Steinmetz Joe/fall 1993, ‘The Encyclopedia of Virtual Environments’

Van Lammeren Ron/2003, Capita Selecta Geo-Visualisation

Chapter 2

Kilander Fredrik, Peter L¨onnqvist/2002, ‘A whisper in the woods-An ambient soundscape for peripheral awareness of remote processes’ (International Conference on Auditory Display Kyoto, Japan 2002)

Kurtus Ron/February 2002, ‘sound waves’

Leslie Van Allen, The Soundscape of New York City in the 1930's

Truax Barry /1978, "The Handbook for acoustic ecology", A.R.C Publications, Vancouver, Canada

Worrall David, ‘PHYSICS AND PSYCHOPHYSICS OF MUSIC’

Chapter 3

Hessische Rundfunk, Kassel/Germany, September 1997, ‘Listening the electroacoustic community’

Krause Bernie/January 2001, ‘Loss of Natural Soundscape’, Global Implications of Its Effect on Humans and Other Creatures

Truax Barry, Soundscape Composition

Truax Barry/1978, ‘The Handbook for acoustic ecology’ published by A.R.C Publications, Vancouver, Canada

Westerkamp Hildegard/2002, ‘Linking Soundscape Composition and Acoustic Ecology

Chapter 4

Maragos Nich/2003 ‘Videogames as narrative medium’

Richard S., ‘BEST AUDIO SYSTEM FOR PC GAMES’

Wireman Casey, ‘Sound Formats and Their Uses in Games’

Chapter 5

Ervin M. Stephen /March 1993, “Landscape Visualization with EMAPS”

Kraak Menno Jan & Ormeling Ferjan/2003, ‘CARTOGRAPHY: VISUALIZATION OF GERSPATIAL DATA’, 2nd edition

Schans Van der R., "The WDGM Model, a formal Systems View on GIS", International Journal of Geographical Information Systems’

WEBSITES –URLS

Chapter 1

http://www.hitl.washington.edu/scivw/EVE/I.B.1.3DSoundSynthesis.html (URL1)

http://www.pervasive.dk/projects/3Dsound/3Dsound_summary.htm (URL2)

http://www.cc.gatech.edu/gvu/people/jarrell.pair/primer.html (URL3)

http://interact.uoregon.edu/MediaLit/WFAE/readings/Thorn.html (URL4)

http://www.anu.edu.au/ITA/ACAT/Ambisonic/spatialoverview.html (URL5)

http://www.i3a.org/i_flashpix.html (URL6)

http://www.macromedia.com (URL7)

http://www.virtools.com/solutions/products/virtools_multiuser.asp (URL8)

http://www.esri.com/news/arcnews/summer03articles/introducing-arcglobe.html (URL9)

http://www.terradesk.nl/onderwerpen/go3d.htm (URL10)

http://www.skylinesoft.com/corporate/technology/ (URL11)

http:// www.planet9.com

http://www.67100.gr/

http://www.parallelgraphics.com/products/cortona/best/

http://www.navigram.com/

http://www.geo-informatie.nl/stuwwal/grs50302.htm

http://www.acm.org/sigchi/chi97/proceedings/paper/kob.htm#U22

Chapter 2

http://interact.uoregon.edu/MediaLit/wfae/readings/vanallen.html#3 (URL2.1)

http://www.d.umn.edu/~ameredit/spch%20and%20%20hrg%20science/Characteristics %20of%20Sound%20Wavesnotes.htm (URL2.2)

http://www.school-for-champions.com/science/sound.htm (URL2.3)

http://www.digistar.com.my/rcf/1_1.htm (URL2.4)

http://www.linguistics.ucsb.edu/faculty/gordon/106/acoustics.pdf (URL2.5)

http://www.coldstare.50megs.com/Pagetwo.html (URL2.6)

http://www.ndr.hu/sound/sound/sound3.htm (URL2.7)

http://www.dsv.su.se/~peterl/publications/ICAD.pdf (URL2.8)

Chapter 3

http://www.sfu.ca/~truax/ganzohr.html (URL3.1)

http://www.sfu.ca/~truax/scomp.html (URL3.2)

http://www.digistar.com.my/rcf/1_1.htm (URL3.3)

http://www.linguistics.ucsb.edu/faculty/gordon/106/acoustics.pdf (URL3.4)

http://www.coldstare.50megs.com/Pagetwo.html (URL3.5)

http://www.ndr.hu/sound/sound/sound3.htm (URL3.6)

http://www.wildsanctuary.com/BK_WAC.pdf (URL3.7)

http://www.sfu.ca/~westerka/writings/linking.html (URL3.8)

Chapter 4

http://www.tetsuboushi.com/games/thesis06.shtml (URL4.1)

http://www.gamedev.net/reference/articles/article1130.asp (URL4.2)

http://www.gamedev.net/reference/articles/article1902.asp (URL4.3)

http://www.gamedev.net/reference/articles/article1958.asp (URL4.4)

http://www.theblackbook.net/acad/student/collins/texts/vgaudio.pdf (URL4.5)

http://www.ibiblio.org/GameBytes/issue21/misc/sound.html (URL4.6)

http://www.dolby.com/press/aes/pr_pr_0310_AES2003.pdf (URL4.7)

http://www.dolby.com/press/pr_pr_0307_IBC2003Preview.pdf (URL4.8)

http://www.dolby.com/press/co_pr_0201_CES2002Gaming.pdf (URL4.9)

http://www.dolby.com/press/pc_pr_0303_Dolby_PC_ProgramGDC.pdf (URL4.10)

http://www.eds.com/thought/innovates/trends_games.pdf (URL4.11)

http://www.gignews.com/crosby2.htm (URL4.12)

www.virtools.com (URL4.13)

Chapter 5

http://www.gsd.harvard.edu/~servin/ieee/emaps.html (URL5.1)