Kimmo Vennonen

Submission

for the

GRADUATE DIPLOMA IN MUSIC

(COMPUTER MUSIC)

Australian Centre for the Arts and Technology

Canberra Institute of the Arts

Australian National University

B 19111174

Kimmo Vennonen

Submission for the

GRADUATE DIPLOMA IN MUSIC (COMPUTER MUSIC)

Australian Centre for the Arts and Technology
Canberra Institute of the Arts
Australian National University

This written submission consists of the following five sections, produced in accordance with requirements for the Graduate Diploma in Music (Computer Music).

1. **Literature Review:** Spatial Sound
2. **Hardware Project Description:** An Ambisonic Channel Card
3. **Composition:** *Reality Intrudes* Programme Note
4. **Analysis:** *Reality Intrudes*
5. **Composition:** *Mirage* Programme Note

In addition, a floppy disk containing the database compiled for the Literature Review and a Digital Audio Tape containing both Compositions is included with this written submission. Both Compositions will also be presented at a concert on a date to be announced.

Foreword

The material presented here and on the accompanying disk and tape represents work over the years 1990-1994, equivalent to two years part time study. There is a consistent thread running through most of it, being the intention of building an efficient system to project surround sound in three dimensions and then to employ it compositionally.

The Hardware Project and Literature Review are the most obvious examples of this aim and they turned out to have a large degree of symbiosis. I started my work with preconceived ideas for the hardware, which I discovered were not completely successful when implemented. I then turned my attention to the process of finding out what others around the world had done and realised that there had been much work done in this field over the last several decades. I also realised the limits of my own knowledge and launched myself into an attempt to understand as much as possible about the issues and techniques involved.

This desire to know resulted in a much more protracted and intense period of work on the Literature Search than I first expected. I am thankful I was advised to use a database for the presentation rather than a formal essay; there is certainly the potential for a fairly thick book in all those articles and papers.

This process eventuated in a new direction for the hardware and also involved a great expenditure of time writing software. I knew I was on the right track however and after a while the results of my research started to materialise. The system I put together is detailed in the Hardware Project Description and is very much an "evaluation testbed" for the ideas as well as future software. Of course, there is a lot of room for refinement and if I had started in 1994 the system would have been considerably smaller and faster. I was very lucky that the hardware I had designed previously could be adapted to a new approach that is a quantum leap ahead of my previous ideas.

At the same time as this activity was going on, I was working on *Reality Intrudes,* presented as part of this submission. This involved extensive use of the ACAT audio studio, working in an area that really pushed to the limit my own and the studio's resources. Sampling uses a lot of computer memory and often is very slow, due to the enormous number of calculations performed by the equipment I used. As a result, the piece took shape very slowly, usually very late at night. I am no longer so naive about the supposed advantages of an all digital way of working! My intention was to compose in such a way that the piece would work well in surround as well as stereo. The latter version cannot practically be presented with this submission, because it needs the system I have built to be reproduced as intended. However, it is not a completely unique system and others around the world are using the same standards. I hope to present it and my other surround piece, *Mirage* , at a concert at ACAT later this year.

It was an interesting process writing about *Reality Intrudes* for the Analysis component of this submission. I came to see it in a new light and I actually decided it was a better piece than when I had been working on its realisation. Analysing it was certainly a good way to close the lid on it; I knew that I would never again try to re-arrange it! However, there is always the temptation of yet another remix. . .

Mirage is the fulfilment of ideas I have long thought about but never really acted on. For many years I have been aware of the vast possibilities in working with "complex feedback". I even tried it on stage with Jim Denley and Mind/Body/Split, but it was never predictable and reliable enough to be used much. This time, the approach was to set up the necessary equipment on a table and make digital recordings of the feedback, then try to put it together into a piece that highlighted the timbral possibilities. The recording went well, but as the programme note for the piece describes, the arrangement was not easy. It is a very different piece to *Reality intrudes,* but some of the creative issues were much the same. Both pieces work with "slabs" of sound rather than discrete notes on a page, requiring a different sensitivity in the listening and sound shaping process. In the end, my initial optimism about "complex feedback was confirmed and I hope to incorporate it into future work.

Acknowledgments

The work contained in this submission did not happen in isolation there are some key people and organisations that made it all possible to get to this point.

Firstly, I am grateful to the Institute of the Arts for accepting me into a graduate course without the usual undergraduate work. It took some trust for the Institute to take me in and believe I could deliver what I promised. Likewise, without their understanding on the two occasions I needed to pause this work, I could not have finished. Many thanks also to Auriel and her colleagues at the School of Music Library: those wish lists of obscure papers and articles that you located really made a huge difference.

The Australian Centre for the Arts and Technology (ACAT) was the venue of most of my work, and without its facilities and staff I would not have had a hope. In particular, I must acknowledge the continuing support, interest and input from David Worrall. He trusted that I would come up with a system that worked and without his instrumental involvement with the *Polymedia Inc* dome, I would have been unable to evaluate my ideas. David, your compositional advice was very succinct when it was needed and your friendly manner was much appreciated. Stuart Ramsden, thank you for your interest over the years in my work and many helpful discussions. Thanks also to Tim Kreger, Niven Stines and Julie Fraser for help in your respective areas.

Capital Arts Patrons Organisation (CAPO) provided the funding in 1990 to purchase the ART effects processor used in the system I built. Although it took a bit longer than I expected to make the system work, your investment in this area has borne creative fruit.

Community Radio 2XX Inc is really where I started in this field of electronics and music many years ago, providing a crucial environment to sharpen my skills and just as importantly, to do it my way. As my current employer, I am very grateful for the flexibility the station could offer,

giving me the time I needed to get this work finished. Initial work on *Reality Intrudes* was also done in Studio C in 1991.

Finally and most importantly, it is hard to acknowledge enough the support of Virginia Killen. Your injections of realism caught me when I was losing touch with "the real world", and there is nothing more real than the two children you bore us. You were more patient than I could have hoped for, as my late nights and long days in the last few months put off all our plans. And to Ingrid and Karina, although you can't yet verbalise it, I know you'll be glad to see some more of me.

Kimmo Vennonen

LITERATURE REVIEW:

SPATIAL SOUND

PREFACE

Following is a description of work undertaken to fulfil part of the requirements for the Graduate Diploma in Music (Computer Music). It was elected to present the literature review as a database using Claris FileMaker Pro for Macintosh, included on a disk with this submission.

I have broadly kept to the original research proposal with the main change being to the title. The literature search revealed many more items than expected. I decided not to limit the search to a number of items or period of years, as many items informed other aspects of this submission and I wanted to attain a comprehensive understanding of the topic. Every one of the three hundred and forty nine articles and papers found was photocopied, read, summarised and entered into the database. In addition to basic details, the entries were categorised for easy access by other users. Many items belong to more than one category. Every entry was also assigned a level of difficulty parameter to assist readers of various backgrounds.

Due to the quantity of material found, it would be impossible to adequately discuss the many (and often contentious) issues in merely a few thousand words. Much of the material is technical and assumes prior knowledge of physics, electronics, acoustics, mathematics, computer music software or sound engineering. An informal overview is included that will be useful for many readers, but in no way is it intended to be a substitute for the database. In the separate introduction some comments on using the database and a statement of the author's own biases and interests are provided, along with a description of the search process.

Kimmo Vennonen

LITERATURE REVIEW:

SPATIAL SOUND

CONTENTS

PREFACE. i

RESEARCH PROPOSAL iii

1. INTRODUCTION 1

2. OVERVIEW. 3

RESEARCH PROPOSAL

Sound Manipulation in Real and Virtual Spaces

In brief, I propose to explore currently existing literature on sound manipulation and control in the spatial realm. These techniques have only become practical in the late twentieth century, adding hitherto unimagined aspects to the composition and performance of music. The topic may be loosely divided into the following areas:

1. Everyday three dimensional space

2. Stereo space as experienced in most recorded music

3. Virtual space, that is, the illusion of space created by electronic devices and reproduced by two or more loudspeakers.

I aim to reach a more profound understanding of the use of space in music, as both a tool and aid to the reproduction and recording of sound and in live performance, and as an important compositional parameter in its own right. New technological developments allowing the composer to place moving point sources in stereo or three dimensional space will be reviewed. Devices and techniques that create the illusion of depth, size, etc. by controlling events or manipulating sound fields in time and space domains, usually under computer control, will be explored. These developments necessarily rely on psychoacoustic phenomena and I will thus examine this area as well.

At the end of this research I intend to present an up to date overview of the literature I have found, summarising the articles for interested readers.

Kimmo Vennonen

LITERATURE REVIEW:

SPATIAL SOUND

1. INTRODUCTION

This database is a compilation of most of the key papers and articles on spatial sound from before the beginning of this century to early 1994. There is a small quantity of associated material, for instance instruction manuals for spatial processing equipment, that are the applied results of some of the papers. As well, there are articles from amateur or consumer oriented magazines that are aimed at the less advanced reader. These are invaluable first reading, as they clearly outline a topic without complex mathematics or much assumed prior knowledge.

The literature search process began with Computer Music Journal, which publishes occasional papers related to spatial sound. These then spawned more references to other journals and magazines, a process that repeated itself over many "generations" of references. The Journal of the Audio Engineering Society and the Journal of the Acoustical Society of America were two mainstays of the literature search. Both are respected journals covering wide areas and both have been long concerned with various and different aspects of spatial sound. At the same time, online database services and other sources like Engineering Index yielded more obscure papers not well related to an existing body of work. Another approach that was quite successful was a manual search, through every available issue of several key magazines that do not publish regular indexes or are not indexed elsewhere.

Inevitably some biases appeared in the search priorities, related to my activity in other areas of work for this submission. In particular, an effort was made to find every possible item on

Ambisonics and a secondary focus was other multichannel approaches beyond stereo and quadraphonics.

Once most of the articles and papers had been found, they were read a second time, entered into the database and summaries written. Occasionally some personal opinions found their way into these summaries. The main bias is my preference for kernel methods of spatial coding and the advocation of surround technologies that convincingly approach the accuracy of natural hearing, whilst offering all kinds of manipulation possibilities. If I thought that an article was lacking in some important respect, an appropriate remark was made in the summary.

The database has been constructed in a way that should allow the user to glean much relevant information about a reference in a glance. The summarising process is bound to be a compromise between conciseness and comprehensiveness. The length of a summary is usually indicative of the breadth and complexity of issues presented in the paper or article. Use of the multiple choice "LEVEL" and "CATEGORY" fields in the database is the best way to navigate an area of spatial sound, followed by reading of the "SUMMARY" field for each item of interest found. The online help features of FileMaker Pro will be invaluable to the first-time user.

The following overview is intended to discuss various aspects of the topic in a loosely historical context, in order to introduce various database categories. These are shown in bold face on their first occurrence. It is not practical to reference every database entry in this overview, nor would it be fair to reference selected material only. Thus the most the overview can hope to present is an easy to comprehend *Gestalt* of many items. Spatial techniques as applied to music production and composition are also mentioned. As with the summaries, there is a degree of personal opinion in this overview.

2. OVERVIEW

It is easily confirmed that we hear sound in three dimensions and the perception of the spatial aspects of sound has been essential to our survival. For instance, ascertaining the location of a charging elephant, or crossing a busy city street both rely on this ability, called localisation. The ear-body-brain combination correctly decodes a handful of simultaneous and possibly conflicting spatial cues, often to within a few degrees of precision and in a fraction of a second.

The perception of music has always been an experience rich in spatial detail, but very much taken for granted due to its everyday nature. It is now acknowledged that the acoustic spaces music was performed in had a profound influence on the particular style. For example it is hard to imagine Gregorian Chant in an outdoors setting, or Balinese *gamelan* music in a cathedral.

The invention of sound recording in the previous century was a cultural milestone, bringing certain music to places, times and people it had never touched. The most primitive apparatus was an instant curiosity piece, yet the superiority of the live musical performance was rarely questioned. What was lacking in recorded sound was a sensation of ambience, a separation of the instruments and a perception of the context of the performance. By the early twentieth century a simple theory of human sound localisation had been developed, that enabled others to propose ways to convey this missing spatial detail.

Thus stereo was born, but it has not always been applied conforming with any **stereo theory.** For example, many audio practitioners are unaware that loudspeaker stereo relies on the conversion of intensity differences to phase differences around the head. Many sound engineers have still not adapted to stereo, preferring "multiple-mono" recording strategies. To some extent this is because the equipment makers have not produced appropriate hardware.

Psychoacoustics is a field broader in scope than spatial sound, but there are many applications of it to this topic. Most practical spatial systems (including stereo) rely on

psychoacoustics, endeavouring to create a natural impression, or an illusion of enveloping sound. Over the preceding few decades to the present there has been a considerable refinement of the theoretical and empirical aspects of the psychoacoustics of spatial sound. However we are still in the situation where there is no agreement on the finer details of what is perceptible, measurable or relevant. A consequence has been a lack of a unified approach on how to implement spatial sound beyond stereo, and indeed a lack of agreement on what are desirable characteristics of a spatial reproduction system. This lack of consensus is not confined just to the hardware, but is deeply related to what is the best way of encoding the spatial experience into an inevitably limited information bandwidth.

This problem could be broadly termed **spatial coding**, with discrete, matrix and kernel solutions being proposed. The discrete and matrix approaches both use speaker feed signals in the transmission medium, with desired sound locations between the speakers being reproduced with a lessened spatial precision or "stereoism" (literally, "solidity"). Kernel methods are more sophisticated, encoding an infinite number of source directions into a given number of channels, but abandoning the concept of a discrete "point source" being reproduced at the limited number of source locations corresponding to speaker locations. Kernel methods do not specify speaker locations implicitly, leaving that choice to the listener who must use the appropriate decoding for his/her speaker configuration.

Quadraphonics was a discrete and/or matrix solution, but there was no real agreement on what was intended to be achieved or what theoretical basis it had beyond "double stereo". Furthermore, its integration into **broadcasting** was found not to be straightforward. As a consequence, it failed. **Ambisonics** is an alternative solution from the same era that offers a very viable compromise using kernel coding, but missed the opportunity to be market tested.

The argument over desirable characteristics has also carried over into the areas of **microphones** and **loudspeakers.** In some ways one can consider this a related debate to that which often occurs in the art of **stereo recording**. Is the intention to record an acoustic event

with great precision, or to generate a perhaps even more pleasant spread of sonic images with not much relation to reality? Should one use highly directional loudspeakers in the aid of better imaging, or use less directional radiators that (in stereo) use room reflections to create a more natural listening experience? These issues are far from solved although it appears that in the cinema there is a greater consensus, because the listening environment and coding methods are more highly standardised.

The question of what is really practical or appropriate has not been settled either, although some will say that the buying public knows best. Most homes cannot accommodate sixteen speaker sound systems, nor is it acceptable yet to hand out three dimensional **virtual reality** headphones to rock concert patrons. Different technologies evolve to suit different contexts.

In the last decade of the twentieth century there has been a growing interest in spatial sound arising from the development of new communications and media technologies. Virtual Reality and High Definition Television (**HDTV**) both require spatial sound reproduction in excess of what ordinary stereo can offer. Many rental videotapes and television shows are now Dolby encoded for **cinema** style surround sound. New equipment and software is beginning to proliferate, claiming to deliver spatial sound.

There is no commercially available system that successfully conveys the natural spatial hearing experience. Apart from Ambisonics, a close approach is made by **binaural** stereo and **transaural** stereo methods, both capable of delivering a quite natural sounding three dimensional effect when used with well recorded source material. However, the limitation is that one must wear headphones with binaural stereo , or in the case of transaural stereo, there is only one good listening position for hearing the full spatial effect through the loudspeakers.

In the laboratory much more is possible, at a much greater cost. If digital audio technology continues to become more affordable then some of these systems will become commercially feasible, but that is no guarantee of their universality or intrinsic merits. Indeed, the rise and fall

of quadraphonics could easily be emulated in the next decade by some new candidate. As before, the two key concerns will probably be compatibility with existing stereo and standardisation of software format, including **perceptual coding** to reduce the digital data rate. **Multispeaker stereo**, using a small number of frontal speakers is proposed as a worthwhile improvement over stereo, while retaining backwards compatibility.

It remains to be seen whether the recording industry is willing to set aside the stereo paradigm and update all the equipment to a new format, when most consumers would still be using stereo (and even mono) for many more years. Once HDTV with multichannel sound becomes commonplace, there may be an irresistible pressure for the audio-only industry to settle on the same or an even better surround format. In the next few years, these transmission and distribution decisions must also be made for digital audio broadcasting (DAB). The use of simple matrixing and speaker-feed signals seems particularly short-sighted to this author, yet that is exactly what is being proposed for HDTV by some European and Japanese researchers.

Apart from what occurs in the mass marketplace, there will always be distinct applications for **other multichannel systems** intended for psychoacoustic and localisation research, for **auralization** (used to simulate building acoustics etc) and even for contemporary music concerts. Current systems are all very specific to their context with not many common features, apart from the use of many amplifiers and loudspeakers. Likewise there appears not to be any accepted spatial coding standard, apart from ad hoc discrete methods. Often the spatial effect is achieved by a "brute force" approach at high expense. However, these multichannel systems can achieve striking effects when combined with the appropriate **spatial software** or **spatial processors.**

Equipment based on digital signal processing (DSP) has become commonplace in both the domestic and professional arenas. Processors for recording studios or on-stage applications employ complex but easily useable algorithms, creating a variety of effects with many consequences in stereo space.

Much effort has gone into the digital synthesis of **reverberation.** Nowadays it is possible to buy for a very reasonable price a device that produces stereo reverberation indistinguishable from a stereo recording of the real event. This technology is commonplace in recording and production studios. Lately a variant has been seen in the domestic sphere in the form of ambience simulators, often requiring the addition of an extra pair of loudspeakers. The devices enable one to play recorded music via a choice of synthesised acoustics, for instance a jazz club or a concert hall. This technology cannot fool the experienced listener, as individual instruments cannot be separated from the stereo mix and given individual spatial characteristics, as would occur in reality. It is not often recognised that reverberation is just one aspect of a perceived **depth/distance** effect, an area in need of more psychoacoustic research.

The biggest current research effort in spatial sound is being directed toward **head related functions**, based on the shadowing effects of the head and on **pinna cues**. Using these functions it is now possible to generate binaural signals that convey three dimensional sound over stereo headphones, as a synthetic parallel to conventional binaural recordings. The main problem is that to achieve acceptable accuracy, the equipment must be calibrated to the individual's spatial location versus spectral response, caused by the pinna. If a generic head related transfer function could be discovered that suited all individual variations in ear shape and size, it could be the equivalent of the Rosetta Stone for virtual binaural audio.

These techniques are being proposed for applications where a large amount of aural information must be quickly processed, for example in an aircraft cockpit where the pilot is making an approach to a busy airport. At the moment the head related techniques are far from reliable and cause many "front/back confusions" at the best of times. In spite of this, an immediately practical application may be in teleconferencing, where it is difficult to resolve the voices of several simultaneous participants by conventional means.

The application of this constellation of technological approaches to music composition and production, as opposed to reproduction, is quite varied. In many respects the popular music industry leads the way in terms of using the latest processing technology, albeit only in the realm of stereo. Equipment like Roland Sound Space (RSS) using a hybrid of methods has achieved a limited acceptance, commensurate with its limitations in terms of listening positions.

For computer and contemporary music no one spatial approach has dominated over stereo, with the possible (and lingering) exception of quadraphonics. Ambisonics is used by some and offers considerable advantages when a true three dimensional effect is required, combined with computational elegance. Others have built unique multichannel systems specific to a given venue. In any case, most people accept the need to use loudspeakers for delivering the music to a live audience, ruling out binaural or head related methods.

In computer and contemporary music composition the use of space is very much an individual matter. To a large degree it depends on the technological and software resources available if one is working in a format beyond stereo. Among others, individuals like Chowning, Stockhausen and Wishart have been influential, providing a glimpse of what is possible. The greatest consensus is found in the world of cinema, where the spatial possibilities are limited by the coding standard and informal conventions delineate what is acceptable.

My personal conclusion is that 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. This has a resulted in a highly fragmented field, not aided at all by a general ignorance about spatial psychoacoustics and the competing marketing departments of key corporate players. What is required is the cultivation of an informed and general outlook on spatial sound, combined with an appreciation of past mistakes and achievements.

This database is one starting point.

Kimmo Vennonen

HARDWARE PROJECT DESCRIPTION:

An Ambisonic Channel Card

PREFACE

Following is a description of the work undertaken to fulfil part of the requirements for the Graduate Diploma in Music (Computer Music). As the ultimate result has been a physical electronic circuit very specific to a certain computer and a larger system of hardware and software, it may not actually be appropriate to submit that item. Hence this description is submitted in lieu of the actual hardware. Should it become necessary to inspect the circuit board, that may be arranged with this author. Alternatively or in addition, it may be possible to demonstrate the circuit board functioning in its intended context.

The original research proposal is included below as a reference point. It should be noted that since 1991 the direction of the Hardware Project changed somewhat, following concurrent work in the Literature Review component of this submission. However this did not alter the intentions described below and indeed necessitated the construction of a comprehensive system in order to evaluate the circuit board. Furthermore, that system has been utilised in the realization of both musical works also included in this submission. Software for both computers used was also written and can be demonstrated in context if necessary. This software work has outstripped the initial intention of being created solely for testing purposes, but I am clear that it is not directly part of this assessment, in the same way that the extra system hardware is not.

Kimmo Vennonen

HARDWARE PROJECT DESCRIPTION:

An Ambisonic Channel Card

CONTENTS

PREFACE. i

HARDWARE PROPOSAL iii

1. INTRODUCTION. 1

2. BACKGROUND. 2

3. AMBISONICS. 3

4. SYSTEM IMPLEMENTATION. 6

5. CHANNEL CARD DETAILS. 10

6. CRITICAL EVALUATION. 13

7. SUMMARY. 15

REFERENCES. 16

APPENDIX A. 18

HARDWARE PROPOSAL

Spatial Distributor Channel Card

I aim to undertake a hardware project directly related to the research proposal. My intention is to design, construct, test and debug the basic building block of a spatial distributor, eventually capable of distributing eight or more sound sources to a multi-channel sound system in three dimensions. This project will also necessitate the writing of a small amount of computer software in FORTH for testing purposes, and some technical documentation.

The initial channel card, as I am calling it, could then be duplicated by others and interfaced to computer software to assemble a complete distributor. Design priorities will include a practical, low cost implementation with easily obtainable components, built on printed circuit board using CAD software. In terms of technical standards, low noise and distortion, good frequency response over the human hearing range and fast movement of sounds at audio rates will be delivered by the card. I am confident that I have the necessary skills and experience to undertake this project within the permitted time and I have already done some preliminary work in my own time.

The final distributor will permit exploration of the spatial aspects of sound at a performance to price ratio otherwise unobtainable. Furthermore, the channel card may be modified if desired, software written, and used in a more general application as part of a computer controlled audio mixer.

Kimmo Vennonen

HARDWARE PROJECT DESCRIPTION:
An Ambisonic Channel Card

1. INTRODUCTION

Recent years have seen a considerable revival of interest in the possibilities afforded by surround sound. Since the days of quad, many different systems have been proposed and developed for widely varying applications, resulting in mutual incompatibilities and an absence of standards: there is no universally accepted solution to the problem.

It has been my project to develop a practical surround sound system based around a computer and specifically to design and build a plug-in board (called the "channel card") to handle the actual placement and movement of sound sources. A geodesic dome supporting sixteen loudspeakers is an ideal test bed for such hardware (Fig.1), [27]. As it happened, I decided to alter the course of development midway, the result being not only two channel cards, but a complete system ready for compositional use. I will describe the process leading up to that result in this report, with full details of the channel cards. Only general details will be given on other parts of the system, as they not part of this assessment and are still under development.

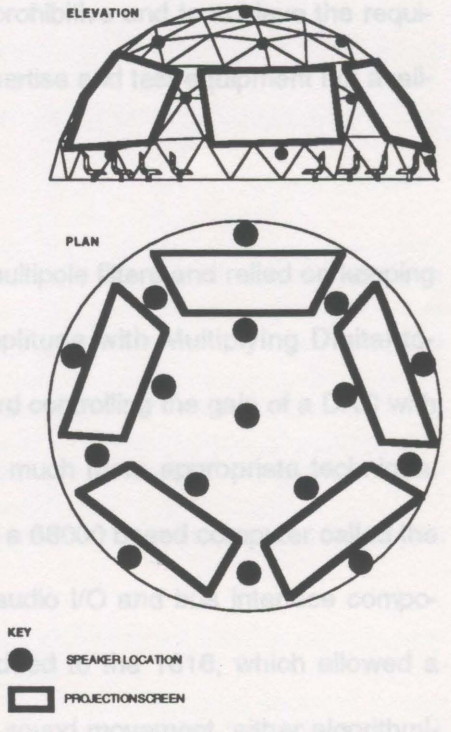


Fig.1 Dome plan and elevation views

2. BACKGROUND

I have made several attempts to build a truly dynamic surround sound system, and the current configuration represents the third generation in my thinking. Over the years, I have changed tack several times with regard to the spatial sound projection techniques, and I am confident that the current system is the most efficient and versatile approach available. The core of the system has been called the Spatial Distributor, a device with a small number of inputs and many outputs). Under appropriate software control, it would move sounds around by varying the volumes sent to the speakers, much like ordinary stereo panning.

Initial attempts involved building hardware for generating digital staircase signals for controlling 16 bit multipliers, that modulated the amplitude of 16 bit digital audio. This approach was found to be cumbersome, due to the need for high quality Analogue-to-Digital (A-D) and Digital-to-Analogue (D-A) stages with many associated filters. For instance a two input sixteen output device would have required 18 filters alone, not to mention the A-D/D-A stages and related control hardware. In 1990, the cost of such circuitry was prohibitive and to achieve the requisite quality in practice was found to have necessitated expertise and test equipment not available to me.

The next approach did away with A-D-A conversion and multipole filters and relied on keeping the audio in the analog domain, but controlling its amplitude with Multiplying Digital-to-Analogue Converters (MDAC's). This involves a digital word controlling the gain of a DAC with the audio fed into the reference voltage pin. This was a much more appropriate technique. Dedicated ramping and control hardware was replaced by a 68000 based computer called the Applix 1616, and plug-in cards containing the MDAC's, audio I/O and bus interface components were designed and built. A MIDI interface was added to the 1616, which allowed a Macintosh running user interface software to control the sound movement, either algorithmically or in real time with a mouse. All programming was in the Forth language.

Although this system worked, it was certainly not optimal. For a start, each MDAC channel card had one audio input and only six outputs, due to lack of board space. Taking advantage of the five-fold geometry of the dome, five outputs normally fed the lower ring of five speakers and the sixth was connected to the speaker at the zenith position. Panning for height effects gave quite indefinite localization with this patch. Images tended to "stick" to the space around the speakers and the 1.5dB steps in gain control afforded by the MDAC's were often audible when moving a sound in real time. A more conceptual worry also arose - the software was a descendant of one dimensional stereo/quad panpot techniques and in retrospect, it was found to be unsuitable for a surround environment. In the software, speakers were grouped together into five triplets, and panning only occurred within an active triplet. The paradigm was one of trying to define a phantom point image as in intensity stereo. Waves or moving masses of sound were not achievable, as was the situation where every speaker was to be equally excited. Even more worryingly, the hardware and software could not easily be adapted to different speaker configurations. It became clear that a new approach, a third stage of development, was required.

2. AMBISONICS

Over some months three hundred and forty nine papers and articles were found dealing with stereo and surround sound. (A database listing and annotating these works has been compiled by the author, forming another part of this submission.) Upon digesting this volume of information and research, it became clear that others had preceded me in terms of my general intentions, but had developed quite different systems. Several had even worked with venues like the dome [1] - [3].

There is now a growing body of theoretical understanding into how we hear three dimensionally in nature, and how techniques like stereo and quadraphonics work (and don't work !). References [4]-[14] are a representative and introductory sample.

On first reading, the Ambisonics proposals were extremely interesting. A 3D sound field could be encoded into four audio channels and then decoded with the appropriate circuits to suit any playback space. A microphone is available that registers and encodes 3D sound in the same format [19-21]. It became obvious that there is a considerable theoretical basis to this system, and it became clear that it could be adapted to the dome environment, as well as other speaker configurations. I saw that it would be a tremendous advantage to be able to encode a piece just once in a standard format, and know it will spatially work anywhere else given the right decoding parameters. References [8]-[11], [15]-[21] are recommended reading as good background material for Ambisonics.

The most general description possible would be to say that the system offers a psychoacoustically optimized way of encoding an infinite number of sound directions into a limited number of channels, and then decoding them to a given loudspeaker layout on playback. A theory of the human sound localization process has been refined [29] and Ambisonics takes this into account, dealing with low and high frequencies in different ways using phase and amplitude.

Current Ambisonics also deals with "zero and first-order directional components" of a sound field in the transduction process. (Higher order systems are yet to be practically realized and need nine or more encoded channels.) First order directivity patterns, like cardiod or figure-of-eight for instance, are characteristic of nearly all directional microphones and correspond to particle velocity caused by transverse waves in a medium. An omnidirectional microphone responds only to zero order components, caused by the pressure component of sound waves. Ambisonics could be described as as system that simulates one omnidirectional microphone and three figure of eight directional microphones, aligned along the X, Y and Z axes.

On playback, this has the consequence that an encoded sound source is decoded not with one or two speakers, but many. In very approximate terms, the volume of each speaker corresponds to its distance from the desired sound location. One interesting result of the encoding outlined above, combined with the different ways low and high frequency material is dealt with

in the decoding, is that some speakers will reproduce low frequencies out of phase relative to other speakers. One can see that this system is much more sophisticated than stereo or standard quad, which are using the one dimensional paradigm of pair-wise intensity panning to place sounds.

There are a handful of Ambisonic sub-formats, corresponding to various places in the signal path from microphone to loudspeakers. In this paper we are concerned with B-format, intended for professional applications. Pressure and velocity components are kept separate, allowing considerable production flexibility and maximising the spatial resolution possible with first order coding. The four values W,X,Y and Z given by the relations below are multiplied against the incoming audio, resulting in four B-format signals [18].

To simulate a source on the circumference of a circle centred on the listener:

for ∂ being the anticlockwise angle from front

and ß the elevation,

$W = 1/\sqrt{2}$ (omnidirectional pressure, circular directivity)

$X = \cos\partial \cos ß$ (left-right velocity, figure-of-eight directivity)

$Y = \sin\partial \cos ß$ (back-front velocity, figure-of-eight directivity)

$Z = \sin ß$ (up-down velocity, figure-of-eight directivity)

To simulate movement inside the circle:

$$W = 1 - (1-1/\sqrt{2})(X^2 + Y^2 + Z^2)$$

Further manipulation of the B-format signals to simulate distance, etc. and global effects like soundfield rotation, tumbling etc. can be achieved by multiplying them by appropriate coefficients.

For encoding Ambisonic signals into a two channel form more suitable to broadcasting or mass reproduction, there is a definition called C-format, which is a slight compromise involving loss of some spatial detail. A decoder at the end of the Ambisonic chain turns B (or C) format signals into D-format, which is adapted for the speaker layout used. For instance, in the 1970's this was often just the standard quad signals LF, LB, RF, and RB. In our case, it is also the sixteen signals sent to the speakers at various angles and elevations.

It was stated earlier that Ambisonics is psychoacoustically optimised - this means that sound is reproduced in a way that corresponds closely to how we actually localise sounds in nature: Frequencies above and below about 800Hz are decoded differently, corresponding to interaural time and interaural intensity cues respectively. Michael Gerzon provides the classic introduction to this topic [8]. This optimization, most effective for typical small to medium size rooms, occurs at the decoder by the use accurately matched all-pass shelf filters. The requirement is to smoothly alter the gain of W relative to X, Y and Z at frequencies centred on 400Hz, with about 90 degrees relative phase lag at that frequency compared to several kHz. The gain proportions are different for horizontal or three dimensional listening. A secondary optimization is loudspeaker distance compensation filtering, which is neccessary to reduce low frequency phasiness for loudspeaker distances under about five metres, due to wavefront curvature. Phasiness is a characteristic of many types of reproduction systems, and is manifest as a localization difficulty, often causing listening fatigue due to unnatural phase relations at the ears. For large distances, the filters are switched out of the signal path.

3. SYSTEM IMPLEMENTATION

The Ambisonic prototype system as implemented here consists of the following major components (Fig. 2). It was adapted from the previous system based around the 1616 computer with channel cards.

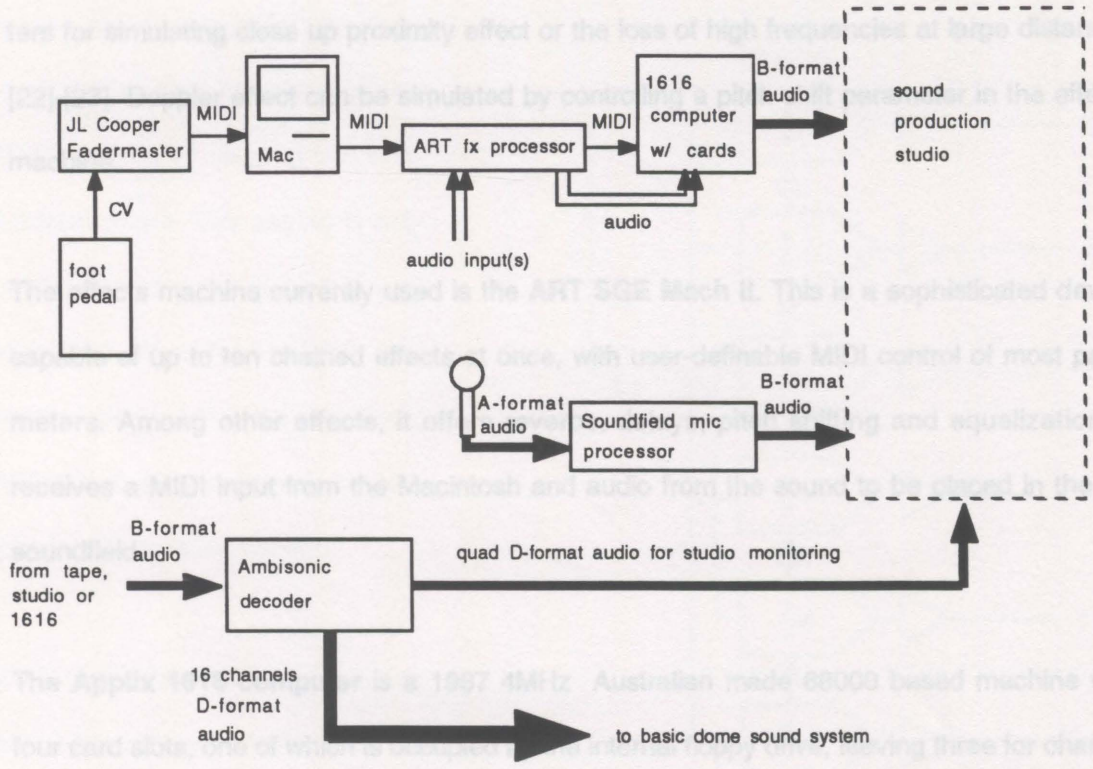


Fig. 2 The Ambisonic system

The **Macintosh** performs the role of calculating all B-format coefficients, either in real time using a mouse and foot pedal interface, or otherwise calculating tables of coefficients for down-loading to the 1616 for faster movements (samples mode). All spatial data is sent over MIDI, using the System Exclusive format. Program Changes, Continous Controllers and Note On/ Off commands are also used to control the 1616. Preliminary spatial software on the Macintosh has been developed by myself, with the use of some aspects of *Streamer*, a Forth application designed by David Worrall [28].

In both real time and samples modes, the horizontal path of the sound is drawn by a mouse, while elevation is controlled by a MIDI foot pedal. A **JL Cooper Fadermaster** converts the variable resistance of the pedal into MIDI Control Change values. The combination of mouse and pedal interfaces allows the operator to completely specify the location of the source in an intuitively obvious way. Distance is simulated by scaling the B-format coefficients according to mouse position, which also can be made to send Control Changes to an effects machine for simulation of reverberation at greater distances for instance, and changing equalizer parame-

ters for simulating close up proximity effect or the loss of high frequencies at large distances [22]-[23]. Doppler effect can be simulated by controlling a pitch shift parameter in the effects machine.

The effects machine currently used is the **ART SGE Mach II**. This is a sophisticated device capable of up to ten chained effects at once, with user-definable MIDI control of most parameters. Among other effects, it offers reverbs, delays, pitch shifting and equalization. It receives a MIDI input from the Macintosh and audio from the sound to be placed in the 3D soundfield.

The **Applix 1616 computer** is a 1987 4MHz Australian made 68000 based machine with four card slots, one of which is occupied by the internal floppy drive, leaving three for channel cards. A MIDI interface was added to make it compatible with the output from the Macintosh via the ART effects machine. The Forth software extracts incoming coefficients from the MIDI stream, either directly memory mapping them to the channel cards (real time-mode) or building arrays in memory for subsequent playback (spatial samples mode). In the latter mode, Controllers and Note On/Offs are interpreted to control sample playback in real time, for instance to alter the speed of a moving sound.

Effected audio is applied to **channel cards** first designed for the non-Ambisonic version of the system, which sit on the 68000 bus and are written to by the 1616. Each card (Fig.3) has one input and six outputs, and consists of six MDAC's, bus interface components, audio input stage and an inverting stage on an extension sub-board which modifies the card to output a B-format signal. The individual card outputs are added by a mixing stage circuit board in the 1616 to make a common B-format output. The reason for using six MDAC's is that each chip can only perform 2-quadrant multiplication, whereas B-format requires the gain control element to invert phase when required, as well as scale the input audio. Thus one chip is required for the non-inverting phase, and a second is fed by a phase-inverted signal. The coefficients are arranged by the Macintosh software such that only one chip of the pair is

passing the signal at any one time, so that one can simply add the MDAC outputs to obtain a true four-quadrant effect. This pairing is only neccessary for the X and Y signals.Tthe W component is not inverted since it is a reference and neither is Z, as the dome is a hemisphere with no down aspect.

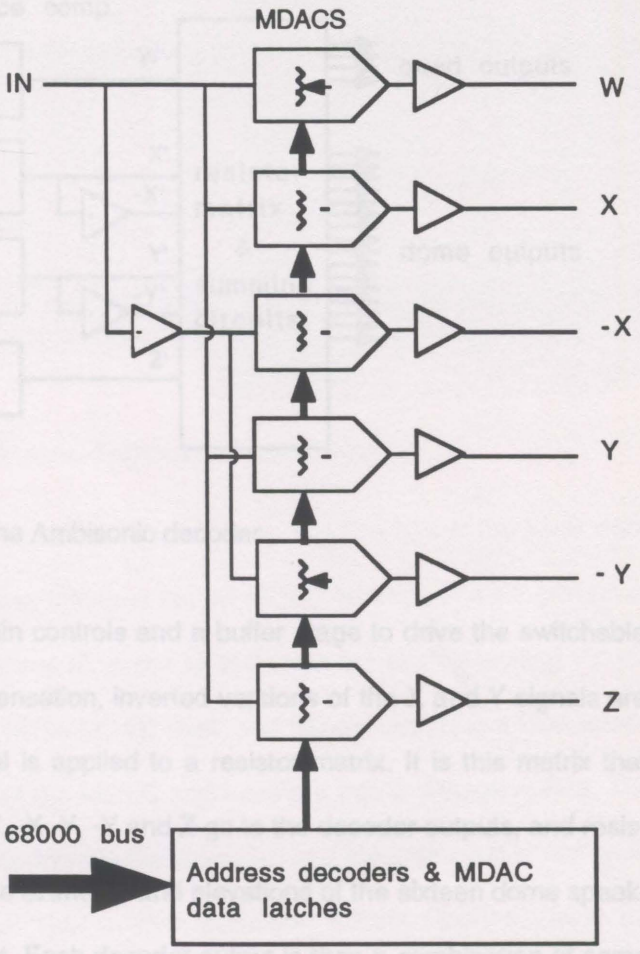


Fig.3 The channel card

At this point, the Ambisonic B-format output from the on-board mixer following the channel cards is identical with that produced by the Soundfield microphone processor box. This means that either or both may be used in studio production and recorded on multitrack tape.

For the moment, let us ignore the studio in the signal chain and assume that it merely gives us a B-format output. The next block is the **decoder** [24]-[25], which generates the psychoacoustically correct signals to suit a given loudspeaker layout (Fig. 4). In our case it actually produces two sets of outputs, one for the sixteen speaker dome system as well as a conventional four channel quad set, for use in a quad production studio. This means that the decoder can be used in a studio compositional environment (one can add a fifth speaker at the zenith driven from the corresponding dome output if desired) as well as in the dome, with very roughly equivalent results. Switching is provided for distance compensation and selection of shelf filter gain to suit horizontal or 3D layouts. The W, X, Y and Z inputs have gain controls, and four bargraph displays are used for test-tone calibration and level monitoring.

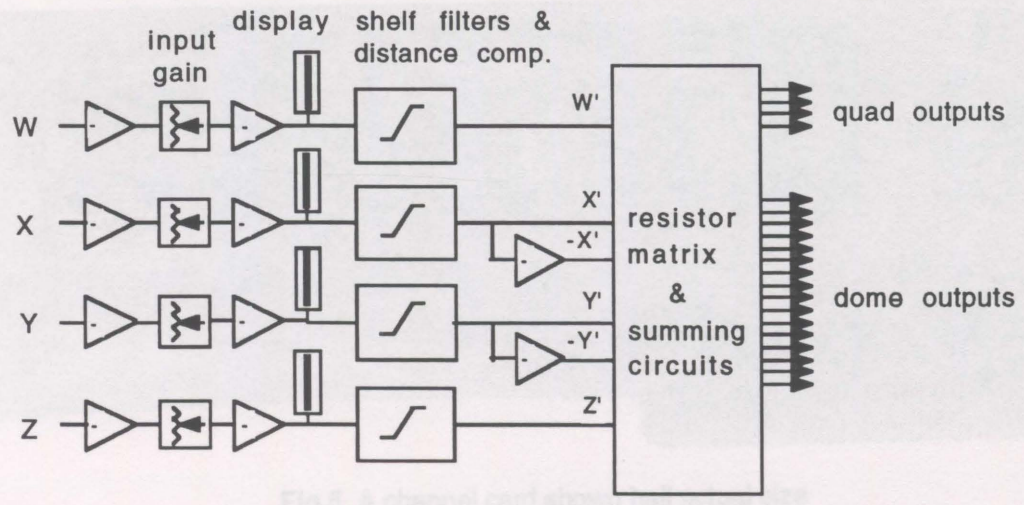


Fig. 4 The Ambisonic decoder

The decoder consists of input stages, gain controls and a buffer stage to drive the switchable shelf filters. After optional distance compensation, inverted versions of the X and Y signals are derived and the modified B-format signal is applied to a resistor matrix. It is this matrix that actually decides what proportions of W, X, -X, Y, -Y and Z go to the decoder outputs, and resistances were carefully calculated to suit the azimuths and elevations of the sixteen dome speakers plus of course the regular quad layout. Each decoder output is then a combination of some of the six signals, summed with a standard inverting op-amp stage.

Once the speaker signals have been derived, we have a D-format output customised to the given playback space. As with stereo, it is neccessary to ensure that the amplitude and phase of these signals is strictly preserved all the way to the speakers.

5. CHANNEL CARD DETAILS

The channel card produces B-format compatible outputs from a single mono input, controlled by the 1616. X and Y signals are combined from X, -X, Y and -Y outputs by a simple mixer in the 1616, that also accepts component signals from another channel card. The mixer thus produces a single B-format output, resulting from the scaling performed by both channel cards.

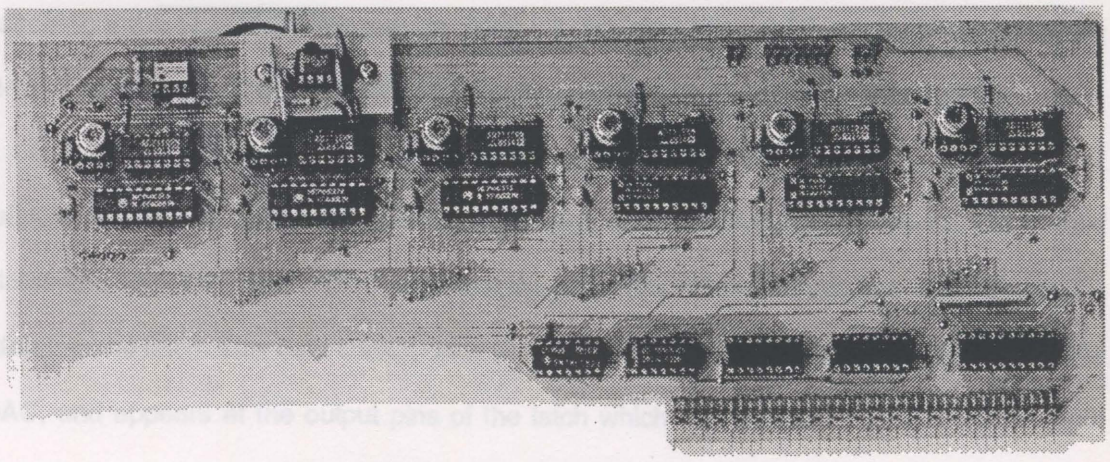


Fig.5 A channel card shown half actual size

In Fig.5 above, we see a photograph of a complete channel card. The five IC's at the lower right handle 1616 bus interfacing and addressing/writing to the six MDAC latches. The six groups of three IC's consist of the following: The lower IC is the MDAC latch for storing scaling values between writes, the upper rectangular IC is the MDAC itself, and the circular IC is a high-speed amplifier for buffering the current output of the MDAC and turning it into a voltage output. The small IC at the upper left of the board is the input buffer amplifier, and the extension board on its right contains a unity-gain inverting stage. This inverter feeds the buffered input signal to the MDAC's that generate the -X and -Y components. The input, six outputs and analogue bipolar power supply are available at the connector at the upper right of the board. The digital power supply and 1616 signals connect to the computer bus using the large connector at the bottom of the board.

Refering to the schematic diagram (Appendix A), the 1616 computer is connected to the channel card via seven of twenty four address lines, read/write (R/$\overline{W}$), address strobe ($\overline{AS}$), external data acknowledge ($\overline{EXT.DTACK}$) and lower data strobe ($\overline{LDS}$) lines. On a concurrent R/$\overline{W}$ and $\overline{LDS}$ signal, the three address bits A1-A3 decide which one of the six MDAC latches is enabled, via decoder U2 (74LS138) and NOR gates U3 and U4 (74LS02). Furthermore, the concurrence of AS and a valid address as detected by decoder U1 (74LS138) is neccessary for enabling U2. Each channel card has a unique base address as set by a jumper (J1), being $800000H (card 1) or $800010H (card 2). The recognition of a valid address is signalled back to the 1616

by strobing the EXT.DTACK line low via diode D1. Six data bits D0-D5 are buffered by the tri-state line receiver U5 (74LS244) and its outputs are pulled up by resistor network RN1.

Resistors R1-R6 pull up the LE (latch enable) pins of the MDAC latches U6-11. On a valid address and selection of an MDAC latch with the lower bits of the address bus, a six-bit word is written into the latch. This data corresponds to the amplitude scaling required of the appropriate MDAC, and appears at the output pins of the latch which connect to the MDAC D0-D5 input pins.

The particular MDACs (U12-U17) used are actually Analogue Devices AD7118 LOGDACS, capable of attenuating an analogue input signal over the range 0 to -85.5dB in 1.5 dB steps. Internally the AD7118 chip is configured as a 17-bit MDAC, with translation logic to derive the 17-bit attenuation range from a 6-bit binary input. The audio input is applied via a low output impedance inverting buffer U24 (NE5534AN) to the Vin pin. The two MDACs U16 and U17 deaing with the -X and -Y components of B-format require an inverse polarity input. The inversion is performed by U25 (NE5534AN), which is fed by the buffered input. This IC and associated components are on an extension board, representing the only changes to the original channel card neccessary to provide compatibility with B-format. The 1616 normally writes to the X and -X (and Yand -Y) MDAC latches in tandem, with only one of a pair passing the input signal at any one time. The current outputs of the MDACs are converted by high speed FET buffers U18-U23 (AD544) to voltage outputs, with phase compensation provided by C1-C6. The outputs of the buffer amplifiers leave the card via R11-R16.

There are two power supplies used for the channel card. The main five volt supply from the computer powers all IC's except the op-amps U18-25 and extensive bypassing is performed by C11-C25 distributed over the board. The balanced fifteen volt external supply used for the op-amps is also bypassed close to each IC by C26-C39. Following standard practice, analogue and digital grounds are kept separate, apart from the internal connection within each MDAC.

6. CRITICAL EVALUATION

I used the above architecture for the channel card based on MDAC IC's available at the time (1991). The circuit delivers high audio performance with no aurally perceptible noise or distortion. However, transitions between attenuation values are in 1.5dB steps, a cause for concern. This was very much the case in the pre-B-format version of the card, when these steps were clearly audible. In B-format Ambisonics, this effect is greatly reduced to the point of imperceptibility with most program material. This is explained by the fact that individual speaker outputs (D-format) are not isolated unique signals, in fact, all four B-format channels contribute to the total amplitude of any D-format output so we gain some extra resolution. The Ambisonic team at the University of York had a similar observation and noted

> " This may be illustrated by considering the addition of four digital numbers of equal length where the sum has two more bits of resolution than the original numbers".[26]

The lack of resolution is most audible as mild square wave modulation on pure tones when using high MDAC update rates. The maximum attainable speed of movement of a sound depends on this update rate, however the channel card hardware is not the limiting factor. With the low speed 1616 running optimised Forth code, it is possible to achieve in excess of fifteen thousand updates per second. This corresponds to a pair of sources being moved over one thousand times per second. To avoid digital noise caused by lack of resolution, it is advisable to make sure that there are minimal occurrences of steps greater than 1.5dB. For instance, a circular movement pattern using several hundred updates per channel card per revolution is quite acceptable. For optimum results one must find a compromise between maximum speed and best resolution. This could be remedied by a minor re-design of the channel card, using now available MDACs with 0.125dB resolution achieved with eight logarithmically coded bits.

In spite of this possibility, it was judged that it would be more worthwhile to completely redesign the channel card using even newer chips with internal latches and voltage outputs. This would make possible a much better integration, with multiple inputs per card. The 1616 is also not a

very suitable choice of computer in today's context. A Macintosh running at over 20Mhz with NuBus slots is one possibility as a dedicated host for the channel cards. Alternatively, the 1616 and channel cards could be emulated almost entirely in software by fast RISC computers like the Silicon Graphics Indigo. With all these prospects, it was decided that one could live with the resolution problem for a little longer, remembering after all that it is only an evaluation system.

Overall, the channel card conversion to B-format was successful and the Ambisonic approach is certainly a great improvement over previous methods. There is a pleasant precision to the quality of imaging and the ability to control a three dimensional soundfield with just four channels is very elegant and effective. The playback environment needs to have reverberation minimised for best results, and the optimum listening position is central. In the dome reverberation tends to diffuse imaging and acoustic treatment may be the best solution. Otherwise, effective if not completely accurate imaging is perceived by the off-centre listeners, who are after all in the majority. This is true not just for frontal images, but all around, with vertical imaging being generally the most difficult. In the quad studio the acoustic is not the problem, but the side and rear imaging is much more delicate than in the dome. This is a well known limitation of the square speaker configuration. Circular pans around the speakers are much smoother than with conventional pair-wise panpot methods and the reproduction of Soundfield microphone recordings is excellent. These observations attest to the correct calculation and calibration of the decoder matrix parameters, as well as the channel card design and accompanying software.

The Ambisonic system here described has been used in a multitrack production studio at ACAT to spatialize music with some success. When one has multitrack digital recorders like the Alesis ADAT, it becomes possible to record overlapping layers of B-format audio, then mix these down to a composite B-format signal. The "bounce-down" technique also works well when tracks are limited. The final result should be recorded onto four ADAT tracks, as misalignment or inherent phase errors in semi-pro analogue recorders can cause playback image shifts. Overall, the ability to code and manipulate 3D soundfield information with just four tracks, without having to consider the playback speaker configuration, is a huge advantage.

7. SUMMARY

Experiences illustrative of the problems in building a 3D sound projection system have been described, and the current system hardware has been detailed. The operation of Ambisonics relevant to the implementation has been generally explained and the channel card described in detail. An evaluation of this work confirms that the system is functioning accurately and shows the excellent potential and elegance of the approach. Enough experience has been gained to suggest some future implementation improvements and the system has been successfuly used in a compositional context.

REFERENCES

[1] Gareth Loy, "About AUDIUM - A Conversation with Stanley Schaff", *Computer Music Journal* Vol.9 No.2 (1985)

[2] Fritz W. Winckel and Manfred Krause, "Acoustical and Electroacoustical Arrangement for the Dynamically Focussed Room", *Journal of the Audio Engineering Society* Vol.20 No.3 pp. 198-206 (April 1972)

[3] Peter H. Heringa, Ben H.M. Kok and Yves Dekeyrel, "The Acoustics and Sound System for Hemispherical Dome Projection", *Journal of the Audio Engineering Society* Vol.35 No.3 (March 1987)

[4] Larry Klein, "Why Stereo Doesn't Work", *Radio-Electronics* (April 1987)

[5] Y. Makita "On the Directional Localization of Sound in the Stereophonic Field" *EBU Review Part A - Technical* No.73 pp. 102-108 (June 1962)

[6] H.A.M. Clark, G.F. Dutton and P.B. Vanderlyn, "The Stereosonic Reproducing and Recording System", *Journal of the Audio Engineering Society* Vol.6 No.2 pp. 102-117 (April 1958)

[7] J.G. Woodward, "Quadraphony - A Review", *Journal of the Audio Engineering Society* Vol.25 No.10 pp. 843-854 (October 1977)

[8] Michael Gerzon, "Surround Sound Psychoacoustics", *Wireless World* Vol.80 No.12 pp. 483-486 (December 1974)

[9] Terry Mendoza, "Ambisonic Sound", *Electronics Today International (Aust.edition)* pp.58-59 (March 1974)

[10] S.P. Lipshitz "Stereo Microphone Techniques: Are the Purists Wrong?", *Journal of the Audio Engineering Society* Vol.34 No.9 (September 1986)

[11] P.B. Fellgett, "Ambisonic Reproduction of Directionality in Surround Sound Systems", *Nature* Vol.252 pp.534-538 (December 13 1974)

[12] C.A. Puddie Rogers, "Pinna Transformations and Sound Reproduction", *Journal of the Audio Engineering Society* Vol.29 No.4 (April 1981)

[13] G. Theile, and G.Plenge, "Localization of Phantom Lateral Sources" *Journal of the Audio Engineering Society* Vol.25 No.4 pp.196-200 (April 1977)

[14] P. Jeffrey Bloom, "Creating Source Elevation Illusions by Spectral Manipulation", *Journal of the Audio Engineering Society* Vol.25 No.9 pp.560-565 (September 1977)

[15] Michael Gerzon, "Ambisonics in Multichannel Broadcasting and Video", *Journal of the Audio Engineering Society* Vol.33 No.11 (November 1985)

[16] P.B. Fellgett, "Ambisonics. Part One: General System Description", *Studio Sound* pp.20-22, 40 (August 1975)

[17] P.S. Gaskell, "System UHJ: A Hierarchy of Sound Transmission Systems", *Radio & Electronic Engineer* Vol.49 No.9 (September 1979)

[18] D.G. Malham, "Ambisonics - A Technique for Low Cost, High Precision, Three Dimensional Sound Diffusion", *ICMC Glasgow 1990 Proceedings*

[19] Ken Farrar, "Soundfield Microphone - 1", "Soundfield Microphone - 2" *Wireless World* (October 1979)

[20] AMS Industries plc "ST250 Stereo Microphone User Guide Issue 1.3" (January 1991)

[21] J. Howard Smith. "The Sound Field Microphone", *dB* (July 1978) pp. 34-36

[22] Paul D. Coleman, "An Analysis of Cues to Auditory Depth Perception in Free Space", *Psychological Bulletin* Vol.60 No.3, pp.302-315 (March 1963)

[23] John M. Chowning, "The Simulation of Moving Sound Sources", *Journal of the Audio Engineering Society* Vol.19 No.1, pp.2-6 (January 1971)

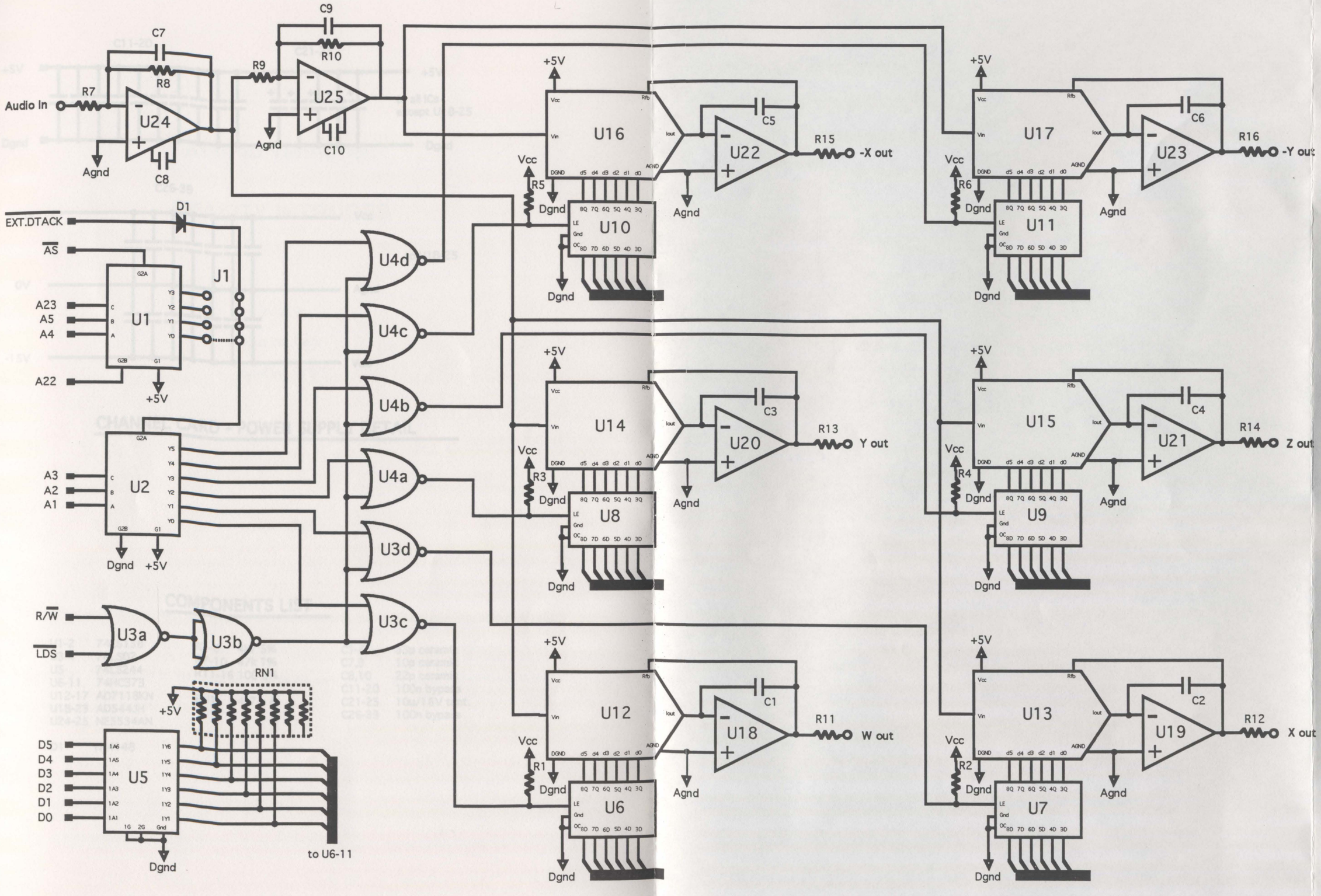
[24] Michael Gerzon, "Multi-system Ambisonic Decoder: 1- Basic Design Philosophy", *Wireless World* (July 1977) pp.43-47

[25] Michael Gerzon, "Multi-system Ambisonic Decoder: 2 - Main Decoder Circuits" *Wireless World* (August 1977) pp. 69-73

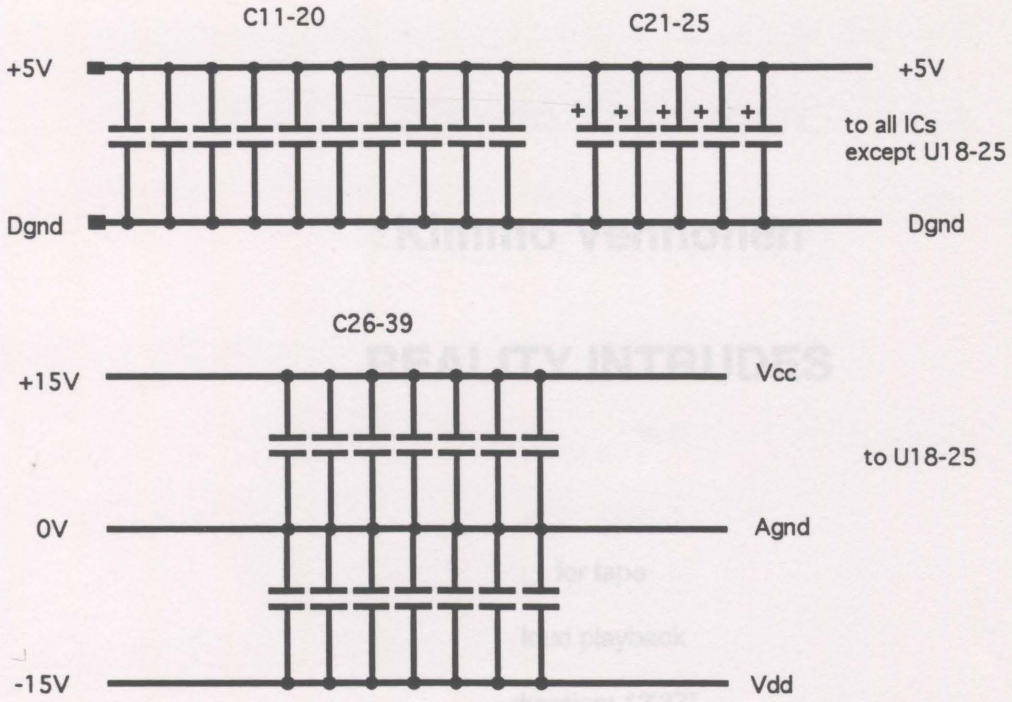
[26] D. G. Malham, "Computer Control of Ambisonic Soundfields" *Audio Engineerinig Society Preprint* No.2463 (82nd Convention March 10-13 1987 London)

[27] David Worrall, "A Music and Image Composition System for a Portable Multichannel Performance Space: A Technical Overview", *Chroma Journal of the Australian Computer Muisic Association* Vol.1 No.3 pp.3-6 (December 1989)

[28] David Worrall, "Streamer: A Real Time Performance and Composition Software Environment", Proceedings of New Music Conference, August 1990, Brisbane Australia.
[29] Michael Gerzon, "General Metatheory of Auditory Localisation" *Audio Engineering Society Preprint* No. 3306 (March 24th 1992)



APPENDIX A CHANNEL CARD - MAIN DIAGRAM



CHANNEL CARD - POWER SUPPLY DETAIL

COMPONENTS LIST

U1-2	74LS138	R1-6	4k7 5%	C1-6	33p ceramic
U3-4	74LS02	R7-10	47k 1%	C7,9	10p ceramic
U5	74LS244	R11-16	100 5%	C8,10	22p ceramic
U6-11	74HC373			C11-20	100n bypass
U12-17	AD7118KN	RN1	4k7 x 8	C21-25	10u/16V tant.
U18-23	AD544JH			C26-39	100n bypass
U24-25	NE5534AN				
D1	1N4148				

Kimmo Vennonen

REALITY INTRUDES

for tape

loud playback

duration: 13'37"

PREFACE

This work is submitted in part fulfilment of the requirements for the Graduate Diploma in Music (Computer Music). The stereo version is presented on a Digital Audio Tape included with this submission. An analysis is included as a separate assessable component elsewhere in this submission. A programme note follows.

Kimmo Vennonen

REALITY INTRUDES

for tape

loud playback

duration: 13'37"

Programme Note

Reality Intrudes was conceived over 1992-94 at ACAT, using a process best described as "handcrafted accretion". There is a complete reliance on sampling "found sounds" from recordings or from our environment, including radio waves. Extensive use is made of treatments and mixing of many layers of sound. Duration is explored in various ways and space is recognised as an important compositional element. A new awareness of the didgeridoo was an important influence, here replaced by a sample of mains hum.

Reality Intrudes could be heard as a work that catches us in our everyday, restless world of associations. As it moves through several stages, we have the chance of entering a calmer, more unified state of being.

Kimmo Vennonen

REALITY INTRUDES

for tape

loud playback

duration: 13'37"

ANALYSIS

PREFACE

Following is an analysis of my work *Reality Intrudes*. This paper is submitted to fulfil part of the requirements for the Graduate Diploma in Music (Computer Music). It should be noted that *Reality Intrudes* is also being submitted for the same award. A tape of the stereo version of the piece is therefore included as part of this submission, and will be useful reference material.

In this analysis I will discuss the work from aesthetic and technical viewpoints. After an introduction, I will discuss its structure, postponing some of the technical issues. This has been done so that the initial perception of the piece is uncoloured by the following discussions. The general reader will then find some helpful background material to these issues. The final sections of this paper present full details of the contemporary techniques that are instrumental to this current realization. Also included is a separate section providing explanations of terms used in this analysis.

Kimmo Vennonen

REALITY INTRUDES

for tape

loud playback

duration: 13'37"

ANALYSIS

CONTENTS

PREFACE i

1. INTRODUCTION 1

2. BROAD STRUCTURE OF *REALITY INTRUDES* 2

3. MICROSTRUCTURE 3

4. METHODS 6

5. TREATMENTS 8

6. MATERIAL 10

7. MIX ISSUES 12

8. CONCLUSION 15

INFORMAL EXPLANATION OF TERMS USED 16

Kimmo Vennonen

REALITY INTRUDES

for tape

loud playback

duration: 13'37"

ANALYSIS

1. INTRODUCTION

Reality Intrudes was conceived over 1992-94 at ACAT* , out of a process best described as "handcrafted accretion". The intention was to create a work using widely disparate sounds and sonic environments, whilst aiming for a unified effect. Transformation and transition were key compositional issues, as the piece operates primarily in the realms of timbre and space. A further and crucial dimension is added by treatments (timbral, dynamic, spatial and temporal) and mixing, heard to the fullest effect in the Ambisonic surround realization. In that version the sound is enveloping, literally three dimensional, and movement in space is explored as a further compositional tool.

On the most general level, *Reality Intrudes* is designed to be heard as a work that catches us in our everyday, restless world of associations and slowly delivers us to a calmer, more unified state of being. At various stages in the piece duration may be perceived quite differently. I have always intended the work to speak not only to the intellect, but also the emotions and the body. It is my personal opinion that most successful musical works share these three concerns.

---

* Australian Centre for the Arts and Technology, Institute of the Arts, Australian National University, Canberra.

Reality Intrudes is music concrete, which originated with Pierre Schaeffer in the 1940's and became a tradition. Since then this technique/style has increased in influence to the point that nowadays its methods are an integral part of many contemporary music genres. A prime identifier of music concrete is the conception of the sound as an object in itself, dictating or suggesting possibilities to the composer. In this type of music "found sounds" are often the compositional building blocks, and modern samplers and digital technology are considerable aids to realizing music concrete. Music becomes a plastic art, analogous to the potter's clay. One is able to work with the "stuff" of sound itself on the most microscopic and macroscopic levels.Yet another popular formulation of music concrete composition is the studio as the instrument. Tape recorders, effects processors, samplers and microphones are combined synergistically, to create and transform sound. This is a far cry from the older conception of the studio as a place with good acoustics, where one goes to "record" a pre-existing piece of music.

2. BROAD STRUCTURE OF *REALITY INTRUDES*

The work has three main sections, a Prelude, an Interlude and a Coda, although often these are interpenetrating and it is inappropriate to carry this description too far. However, for the sake of further discussion I will identify the following sections, etc. with approximate time points as measured from the start:

[a]	Prelude	0'00"
[b[	Section 1	0'30"
[c]	Interlude	1'47"
[d]	Section 2	4'01
[e]	Section 3	10'20"
[f]	Coda	12'53"
	(end)	13'37"

3. MICROSTRUCTURE

During this discussion component sounds are identified by simple descriptors, not necessarily indicative of origin or design process. The use of terms like high frequency, low frequency, mid-bass etc. is parlance borrowed from sound engineering, and is much more appropriate than trying to assign more musical terminology to an essentially sonic domain. These terms are presented in bold face on their first occurrence and listed in the explanation of terms section.

The Prelude is in the form of a crescendo from silence. The very first sound is a steady buzz, which almost immediately breaks up into an intermittent, fast **chaotic** texture. Related layers of the same basic sound process attach themselves to the initial "seed", causing an increase in density and loudness. In the final moments, a **mid-bass** drone appears from the layers of chaotic textures. This drone is in fact formed from the original source material for the beginning of Section 1, and functions as a link into a very different sonic environment. At the end of the Prelude all components have reached their final volumes, and no further buildup is possible without a new sound.

A colloquial description of Section 1 could be that it is "a high speed roller-coaster ride that comes to a dramatic end". The beginning is marked by an immediate increase in loudness, caused by a distorted sound analogous to a stylus skating across a record. As this peaks, another loud sound breaks in, moving downwards in pitch initially, then settling on a constant pitch and timbre. This, together with repeating and complex repetitive figures in the **mid-range,** is one of the characteristic sounds of Section 1. The drone that briefly appeared at the end of the Prelude re-enters, this time seeming very much in the **low-bass** region and heavily distorted. Some of the main compositional elements having thus been introduced, the section continues using these sounds in various ways, sometimes almost unrecognizably transformed.

Almost halfway through Section 1 the distorted bass drone makes its final loud entry, and a **high frequency** sound similar to the chaotic texture in the Prelude "rides" above it. This marks

a turning point. The high frequency texture resolves itself into a series of rapid high frequency **loops**, and the overall pace remains very fast, almost headlong. This "snowballing" effect can not be sustained for long, and some loops begin to have a definite downwards movement in pitch as well. A sustained sound emerges, causing a "drag" effect and a general slowing and slurring over time. After a few moments of this, it turns into a swelling, enveloping **reverberation**, as if one had been "ejected" into a large space.

The Interlude is heralded by a loud, low bell sounding over the preceding reverberation.This bell is really the unifying motif of the Interlude. The disruption and pace of Section 1 fades into the **distance** with the decay of the bell, and a six beats of a low drum in the **foreground** introduce several seconds of complete silence. Here is the calm that counterpoints the chaos of previous events, the opportunity for a breath and a moment's repose. Then in the far distance a response is heard, and the foreground drum repeats its simple motif. Other drums join in, and two elements characteristic of Section 2 are presaged: a very low bass drone for several moments, and a distant horn. The bell, horns and drums "converse", joined by the reappearance of the drone, this time with higher harmonics, **distortion** and then a sound reminiscent of guitar feedback. This sound is a crucial link into Section 2. An expectation has been built up of imminent change and this is now answered by a low bell. The distortion has "infected" the bell, and gradually all is subsumed by this "fiery" distortion. The piece reaches its nadir at the end of this Interlude.

Quite soon the distortion changes, acquiring an undulation in volume, and the drone "pushes up" from the low bass region. Most of the distortion clears, exposing the rich harmonic structure of the drone. After one last distorted bell, it becomes clear that new territory has been entered. A mid-bass harmonic quality happens in the drone, accompanied by two wind-like **resonances** in the midrange. The primary one is undulating, related to the flow of bass energy in the drone with occasional high resonant peaks. The other is much more transient, but is also related to the bass movement. The drone including associated resonances is really the central stream of Section 2 out of which everything arises and to which everything returns. There is a strong

similarity to a didgeridoo drone here, a kind of "timelessness". The **sub-audio** components of the sound can have a powerful visceral effect if the playback level is loud enough and large speakers are used.

New forms appear, characterized by high frequency **noise**-like textures, brass-like sounds or a warbling tone. Often there is a modulation of location or volume occurring. None of these sounds exist for very long, always being replaced by others. There is a general tendency towards greater organization however, and even a few melodic fragments are heard. After some minutes the horn from Interlude is heard again and this marks a turning point. It introduces an accompanying voice "riding" the waves of noise, and there is a definite interaction between the voice and the horn. This event heralds a coalescence, a greater density and loudness of the brassy or noisy textures that follow after the voice is lost in the general flow. Layers of sound begin to overlap more and more, with more interplay between them, yet they are still related to the drone. The rapidly fluttering brassy textures become constantly present, building to a peak on a steady, loud brassy tone.

An element reminiscent of the Prelude and Section 1, a high frequency chaotic sound, breaks through. This turns out to have low frequency components in unison with the drone as the brassy textures are heard again at different pitches. The chaotic sound gradually takes over from the drone and the beginning of the Section 3 is the final disappearance of the drone. The function of the chaotic sound is to **spectral**ly link Section 2 and Section 3, to "grow" out of the textures surrounding the drone and serve as a "catalyst" for the new organization and sounds to follow.

A moment frames a sustained harmonic texture arising from the chaotic sound, which then breaks up again. New sounds are then introduced. A mid frequency radio **static** noise, low horn-like sounds and "crystalline" sustained tones sometimes with reverse **envelope**s or organ-like timbres, are the key components of Section 3. This section has much more stasis than anything preceding - sounds "hang" in space instead of being crammed together, or growing

out of a stream of bass energy. Nevertheless, there are moments of density. The chaotic sound disappears into this new structure. The radio sounds including horn-like textures are the prime direction determining elements, with the crystalline or organ sounds serving as counterpoints. True to its nature, the radio provides occasional small surprises, even a fragment of a voice, within the constant backdrop of noise and static.

Near the end of Section 3, there is a swirling organ/crystalline sound that momentarily dominates the radio horns, with a short, subliminal **diffuse** harmony in the low midrange. This is followed by a final, definitive response from the radio sounds, and then we become aware of the diffuse, harmonic sustained sound which is slightly louder this time. This sound functions to subtly reconcile the radio sounds with the crystalline/organ textures and also introduces the Coda.

This begins after the exit of radio static and horns, leaving just the diffuse harmony and mostly the crystalline aspect of the organ/crystalline textures. The Coda is characterized by a decrease in density and volume. The crystalline sounds are "suspended in the air" above the soft diffuse harmony and gradually disappear. The final seconds allow one to hear the harmony on its own for a few moments, a kind of "afterglow", before it too fades into silence.

4. METHODS

The prime technique used in *Reality Intrudes* is digital sampling, closely allied to treatments, often performed within the sampler itself. Sampling is a modern development of tape recorder methods from earlier days, and has huge advantages over the latter. A sound may be extracted from its original context, most notably in the temporal sense (the equivalent of splicing it out), but also in terms of timbre by means of **low-pass**, **band-pass** and **high-pass** filtering. Very often the composer audits raw material with good potential that is inextricably linked to its

context, so the eventual sample ends up carrying part of its original context with it. In effect, the sample can be an object in itself with a "history" and an "aura".

Composing with samples requires a sensitivity to these objects and their attributes, and an acceptance that flexibility is an essential trait when putting samples together. It is like working with wood - the end result is much better if one understands where the grain is and how to work with it. Transformational techniques afforded by modern sampling technology are not infinite in capabilities, and must be used wisely when adapting a sample into a new context. The role of surprise is also important; it is hard to predict just how well something will work when it is reversed for instance, or has its **attack/decay** envelope modified. One must be ready to experiment, and to be informed by occasional "gifts from above".

A final comment about sampling is that it is not necessarily performed by a machine of the same name. It is a way of working more than anything else, and can also be performed with a tape recorder, a CD player or a computer, in fact anything that can store recorded sound and transform it. The converse is also true - one can use a sampler to make music that has nothing to do with the philosophy of sampling outlined above (using generic, factory-made piano samples to replace the acoustic instrument, for instance) .

"Treatments" (otherwise known as "effects") are transformational techniques applied to the raw sonic material. This is apart from the inherent modifications possible by basic sampling, for instance playing a sound backwards or at a different pitch. Treatments are usually additive, in the sense that one imparts a certain kind of "atmosphere" to a sound. They can function in the time domain (for example **delay**, reverberation, **chorus**, etc.), the pitch domain, (**pitch-changing**, **harmonizing**) or in timbre space (**phasing**, **equalization**, **exciting**, **ring modulation**, distortion) or in the field of **dynamics** (**compression**, **gating**, **expanding**). Of course, many treatments have implications in more than one of these domains. Digital technology has made possible the fine control of treatments - the simulation of lifelike reverberation being a case in point.

Treatments can be combined with each other, in **series** or in **parallel**, to create new possibilities, or more real simulations of nature. Even wider horizons are seen when one applies **feedback** under human control to a complex network of treatment processors. The modern day electric guitarists are only scratching the surface. As mentioned before, some samplers incorporate a wide choice of internal treatments, but usually this is not as flexible as connecting **stand-alone** units in a custom configuration. For this analysis, it is not necessary to distinguish the origin of a treatment. Of course it is possible to sample a treatment, which then becomes an object in its own right.

Every sound in *Reality Intrudes* originated in the outside world, either in the form of a recording I made with a microphone, something found on a CD or tape, or even the radio. Nothing is truly synthetic or made to measure. On the other hand, a lot of high-powered processing was sometimes applied to these found sounds, magnifying some detail or giving them "alternate personas".

5. TREATMENTS

To give some idea of the use of samples and treatments, the following may be useful. Some of the more important samples as identified previously are listed next to the treatments used on them. The usage of a sample or (class of related samples) is indicated by [P] - Prelude, [1] - Section 1, [i] - Interlude, [2] - Section 2, [3] - Section 3 and [C] - Coda. Please note that the list below only includes treatments. I will comment shortly on the origin of some of the sample materials and spell out what kinds of transformations they underwent by sampling processes.

Chaotic HF texture [P], [1], [2], [3] - no treatments

Mid-bass drone [P], [1] - distortion, pitch change, equalization, compression, reverberation

Rapid mid-range loops [1] - equalization, distortion

Sustained reverberance [1] - reverberation, equalization

Low bell [i], [1] - reverberation, exciter, equalization, distortion

Horns [i], [2] - reverberation

Drums [i] - reverberation

"Guitar feedback" [i] - distortion, compression, equalization

Low drone [i], [2] - compression, expansion, reverberation, distortion, **high-Q wah-wah**, equalization, complex feedback, pitch change

High warbling textures [2] - reverberation

Brassy textures, sometimes fluttering [2] - reverberation, equalization, **amplitude/pan modulation**

High frequency noise [2], [3] - equalization

Voice [2] - equalization, compression

Radio static [3] - compression

Radio horns and other radio sounds [3] - compression, equalization

Organ/crystalline [3] - equalization, reverberation

Harmonic drone [3] - reverberation, harmonization

It can be seen that compression, equalization and reverberation are the most often used treatments; this is true for many genres of recorded music. Compression is primarily used to make a sound have different attack/decay characteristics or to maintain a more constant level, often necessary for a sound to be heard consistently when combined with dense textures. Equalization can alleviate unwanted frequencies or to accentuate desired ones; often to make a sound "stand out" or make it subtly "sink in" to a mix. Reverberation is used to give a sense of space and distance, or to widen the sonic image. It can also simulate any acoustic, from a concrete pipe to a concert hall.

6. MATERIAL

The chaotic high frequency texture, heard throughout this work, is a vastly sped up sequence of various looped extracts of shortwave radio. The sequence is exceedingly complex, and the speeding up and looping process produced its own artifacts (related to **aliasing**) that completely changed the timbre of the sound. This was additionally aided by pushing the process faster than the sampler allowed, resulting in chaotic repetitive patterns. Some of the results of this microcompositional process were re-sampled and categorized into various convenient classes for easy access. No additional treatment was necessary.

The mid-bass drone in the Prelude and Section 1 originated as a synthesizer sound on a CD. It was sampled and transformed into a continuous, steady loop which could then be passed through various treatments.

The rapid mid-range loops in Section 1 were sampled from a recording I made of the first synthesizer I built at the age of fifteen. That machine was built using a purposely unstable architecture, generating chaotic sounds due to feedback around **phase-locked loops**. This generated extremely unpredictable timbres and rhythms with many odd and even harmonics, varying with every moment. Samples were organized and turned into loops.

The sustained reverberance at the end of Interlude was originally caused by a percussive sound passed through a treatment machine with a very long reverberation setting. The resulting stream of reverberation was sampled and its envelope shaped, then given additional reverberation. This is an example of a treated sampled treatment.

The low bell came from a CD of synthesizer music with distinct Tibetan influences. It was originally much higher in pitch, but playing it an octave and a half lower still retained its particular identity. Horns used in Interlude and Section 2 were recorded in a naturally

reverberant setting, then sampled for easy access.The drums were originally on a cassette tape of music concrete. After sampling, extensive work was involved in removing tape hiss and re-equalizing. The original reverberation and atmosphere remains. The "guitar feedback" is a sample of violin played by the Kronos Quartet, played backwards and lower in pitch, with distortion, compression and equalization treatments.

The low drone is without question the most processed sample in the piece. It began as phased mains hum, which was then turned into a simple loop. This was distorted, equalized, expanded, compressed and reverberated into approximately its current basic timbre. A very **selective high-Q filter** implemented within a wah-wah treatment on the sampler was manually modulated while the loop was playing. In addition to manual control, the filter's **resonant frequency peak** automatically followed the bass amplitude undulations of the processed drone loop, greatly magnifying mid-frequency spectral details. During this live modulation, the pitch of the drone was subtly varied, along with some looping parameters. All these transformations combine to create a very organic sound. A lot of time was spent trying to recreate these effects with a newer and slightly different sampler, but eventually I returned to a recording of the modulated drone made on an older machine (no longer available to me), as it had greater integrity.

There are actually two low drones - the secondary one can sometimes be heard adding to the "energy" of the first, but with a slightly different timbre. It uses additional pitch changing to generate sub-audio **sub-harmonics** and uses a different set of parameters for the high-Q filter. The secondary drone is combined with the first and it arose out of an attempt at recreating the first drone. Subtle **frequency-selective compression** is applied to the composite drone, with the effect of magnifying the attacks of the bass swells.

The high warbling textures heard in Section 2, as well as the high frequency noise are both samples of short wave radio with minimal transformation, apart from different playback pitches. The brassy textures had the same origin, but sometimes received amplitude/pan modulation to impart a fluttering character. The voice heard in Section 2 is an Islamic call to prayer, sampled

from a record. The radio horns in Section 3 came from a radio tuned to mobile telephone and pager frequencies, and were unchanged apart from playback at different pitches.

The organ/crystalline sound is another example of complex transformation, involving the manipulation of sampler looping parameters. A single series of samples over a minute long was made from a CD. Some parts of the original source material used organ-like sounds, while others had the crystalline nature. The composite sample was then played backwards and forwards between different points at different pitches, with added reverberation and equalization. The process of changing loop parameters in real time was not entirely predictable, causing a very improvisational approach, especially when listening to the accompanying radio sounds. The result was found to be unrepeatable, so I had to revert to a recording made with the older sampler for a work-in-progress version of *Reality Intrudes*.

Finally, the harmonic drone at the end of Section 3 and throughout the Coda was a tamboura-like sound from a CD, looped and given extra richness via harmonizer and reverberation treatments.

7. MIX ISSUES

In much contemporary music, the **mixdown** of all component parts is crucial, being responsible for the realization of any potential contained in the compositional ideas, recording and pre-production. The mixdown is not just a trivial technicality in this genre, but requires the highest combination of technique and artistry. *Reality Intrudes* was conceived for a multi-channel surround playback, but it was recognized early on that it must also work in standard stereo, with all of its attendant limitations on clarity and spatial imaging. A good stereo mix can be a vital step on the way to a successful surround realization. Dynamics must be preserved in some parts, while compression is sometimes necessary to guarantee the audibility of sounds at their optimum volumes.

In this realization the control of overall subjective volume is very difficult, due to the rapidly changing timbres. It is well known that the ear perceives sounds of equal absolute intensity but differing timbres at different volumes. For instance, a harsh sound with many odd harmonic components is subjectively much louder than a bass sound with predominantly even harmonics, like the drone in Section 2. Simultaneous sounds in the related frequency bands can cause masking effects as well. Surround sound can ease these problems by distributing sounds in space, making it much easier to discriminate between them.

In the piece there were some samples that needed to be heard at maximum volume, for example the low bells. Likewise, the central drone needed a high recorded volume, because most of the energy was found to be in the low bass region which many smaller speakers could not reproduce well. If equalization was not optimized then the drone would sound very "woolly" on some larger speakers. High frequency samples occuring at the same time needed to be mixed in quite softly, or they would overpower the subtle upper harmonics of the drone, causing a flat, one dimensional result. At other times there were predominantly mid or high frequency sounds occurring, and care was needed to prevent the combination sounding too harsh for too long.

Due to the many overlapping and changing layers of sound, it was found to be impossible to mixdown all the component sounds at once with a satisfactory degree of control. The solution was to pre-mix certain layers so that they "dovetailed" with others in close temporal and spatial proximity. The pre-mixing of sounds was a slow process, performed by hand and repeated until it was optimal in each case. The "shaped" sample was then recorded back onto the tape it came from, replacing the "raw" sample.

Many sounds were already treated in some way (often by reverberation), imparting an individual stereophony to them which the ear interprets as environment or distance cues. In a complex mix these recorded ambience components can be made to "happily" combine to create a three

dimensional illusion. In addition to the standard stereo left-right dimension, high frequencies are perceived at greater heights relative to bass sounds due to **head-related effects**, and reverberance-related distance cues (properly implemented) add depth simulation.

An **Ambisonic** surround mix is of course preferable, since it encodes in just four channels literal three dimensional spatial information, including left-right, back-front and up-down dimensions. In this and other respects it is vastly superior to the quadraphony that was practiced in the 1970's. In Ambisonics the ear-brain combination is not working too hard to extract spatial detail, resulting in less listening fatigue.

Playback can be through an array of six or more speakers surrounding one in a hemisphere or other arrangement. In the sixteen speaker setup of the Polymedia Inc. dome, a good sense of spatial detail is observable. The current Ambisonic mix can only be played back from a four (or more) channel digital recorder, through a decoder for a given speaker array. However, the Ambisonic approach guarantees optimal playback through many possible speaker arrays, as long as correct decoding is used.

Some sounds are set to be enveloping, diffuse or hard to **localize**, while others have precise trajectories drawn by hand or computer. In the surround environment it is possible for various zones to have different reverberation parameters for instance, and a sound may be made to "dance" through these virtual spaces.

8. CONCLUSION

Reality Intrudes is a music concrete piece with much detail, from the micro-level of sample treatments to the overall spatial realization. There are many complex technical issues involved, and the prime issue for the composer is to foresee and deal with them in such a way as to not loose sight of the whole. Working slowly over years is one way to maintain such a perspective, although perhaps it is not optimally efficient.

In this paper we have set *Reality Intrudes* into the context of music concrete, discussing structural and aesthetic details without technical concerns, in order to aid the initial appreciation. After an introduction to sampling, technical methods and source materials were presented, to aid a fuller understanding of the work.

INFORMAL EXPLANATION OF TERMS USED

aliasing - an effect similar to ring-modulation, caused by a unwanted high frequency (or sampling rate frequency) interacting with a wanted signal. It creates inharmonic "dirty" by-products corresponding to sum and differences of the two signals, plus sums and differences arising from the sum and difference interacting with the original frequencies. Modern digital equipment is designed to minimize aliasing, although sometimes one can use a sample rate low enough to cause aliasing as a desired result.

Ambisonics - a theory of human spatial hearing, and a corresponding system for loudspeaker sound projection and 3D recording. It deals with sound pressure and velocity independently, and is said to be psychoacoustically optimized to our human spatial hearing abilities.

amplitude/pan modulation - changing the spatial position and/or volume of a sound by external means.

attack - referring to the first instant of a sound, how it grows from silence.

band-pass - a filter type only permitting sounds within a certain defined range to pass through.

chaotic - despite an appearance or sound of randomness, chaotic structures can have a very fine and complex internal order. Chaos theory is a modern science that analyses various chaotic processes.

chorus - a time delay treatment skewing a sound by 10 - 30 milliseconds, once or many times simultaneously, resulting in an "ensemble "or "multiplied" effect when combined with the original.

compression - a treatment that, on the broader level, reduces dynamics and keeps the volume of a sound more constant. However it is possible to increase the dynamics of attacks with some compressor settings.

decay - referring to the end of a sound, how it declines to silence.

delay - delay treatments are performed by delaying a sound by microseconds to seconds, then combining the result with the original to produce echoes, for instance.

diffuse - unable to be sharply perceived, spread out in space or lacking fine detail.

distance - a simulated depth conveyed by softness, diffuseness and reverberation is automatically interpreted by us as distance, similar to natural hearing.

distortion - a treatment brought about by changing the shape of electronic signals, usually by restricting amplitude in various ways, for example "clipping". The process causes harmonics, often odd, to appear with the original sound causing a "dirty" or "harsh" quality. Even harmonics add "mellowness" or "warmth".

dynamics - the overall variations in loudness are referred to as dynamics. A sound has plenty of dynamics if it contains both loud and soft extremes. "Dynamic range" is a way of quantifying the difference between loudest and softest sounds.

envelope - the temporal vs. amplitude shape of a sound, including characteristics like attack and decay. One may visualize the individual sound waves being contained in an envelope with a certain shape corresponding to how the sound begins, sustains and ends, or even repeats. Reverse envelopes are particularly striking, due to their lack of similarity with natural sounds.

equalization - a process of boosting or reducing certain frequencies to arrive at a desired spectral shape. This is usually performed by ear, with sophisticated types of "tone controls".

exciting - a form of timbral treatment involving the careful use of equalization and distortion and sometimes delay techniques, to create a "brighter" or "sharper" sound. It often emphasizes high frequency detail and causes a sound to be more discernible when combined with other dense textures.

expansion - the opposite of compression, meaning that the dynamic range is increased.

feedback - connecting the output of a process back to the input, directly or indirectly results in the recirculation of some information or sounds. Positive feedback is ascribed to the case where the overall output keeps increasing in volume or spectral detail for instance, due to a large amount of feedback.

foreground - in this context, a sound without much reverberation and with good definition is perceived as "close up", or in the foreground.

frequency-selective compression - compression that only operates in a defined frequency range or ranges. In combination with equalization, it is possible to apply compression selectively in this way, causing minimal by-products to other frequencies.

gating - a kind of extreme expansion, where the sound is only allowed to pass through when it is above a certain threshold volume.

harmonizing - a pitch-changed copy of a sound is added to the original, causing a harmony. Feedback can cause multiple harmonies from the one input signal.

head-related effects - we hear different angles and elevations of sound sources with different spectral shape and details. This is due to the head partially shadowing the sound, and the convolutions of the outer ear causing spectral changes at different angles of incidence. This mapping of spectral detail to spatial location helps to account for the brain's perception of "high" pitches sounding like they come from a physically high origin.

high frequency - above 5 kHz.

high-pass - a filter type permitting only high frequencies to pass through.

high-Q wah-wah - a treatment using very sharp, variable frequency band-pass filtering changed by the performer. It is often heard with electric guitars. Q refers to the sharpness of the spectral peak caused by the filtering.

localize - to pinpoint the source of a sound is to be able to localize it. Diffuse sounds are very hard or impossible to localize.

loop - on a tape recorder, a cyclic sound caused by splicing the end of an excerpt back to its beginning. A sampler implements a loop by playing from beginning to end, then immediately starting playback from the beginning again. May be reversed as well.

low-bass - below 100 Hz.

low-pass - a filter type permitting only low frequencies to pass through.

mid-bass - approximately 100 - 300 Hz.

mid-range - approximately 500 Hz - 5 kHz.

mixdown - the process of combining multiple sounds (usually on multi-track tape) into a single unified stereo (or other format) mixture. After this has happened, individual sounds can no longer be entirely separated from the whole. The mixdown is one of the final steps in modern music production. "The mix" refers to the end result, not the mixdown process.

noise - an electronics definition is intended, implying the sound possesses all possible (audible) frequencies and is heard as a "hissing", or "rushing" sound. Noise is random in origin and is often thought of as undesirable. Equalization can change the distribution of frequencies and thus the overall sound of the noise.

parallel - an electrical term meaning "all at the same time", or "multiple paths".

phase-locked loop - a circuit that endeavors to keep its frequency of oscillation (or multiple) exactly in tune with an external source. This is not always successful if that source is varying in some way, or if loop parameters are such that instability is likely.

phasing - a time delay treatment that delays a sound by 0.1 - 5 milliseconds and combines it with the original. Odd spectral effects, often deep spectral "notches" and "peaks" can occur due to the addition and cancellation of different frequencies.

pitch-changing - a real-time treatment that transforms sounds up or down by a preset amount.

resonance - a discrete and single frequency emphasized much more more than others is the meaning in this context. Due to the positive feedback involved in making a resonance, it often has a higher volume and "stands out" relative to surrounding sounds.

resonant frequency peak - an exaggerated spectral feature caused by resonance or feedback, usually at a single frequency or very narrow range of frequencies.

reverberation - a pattern of echoes, initially discernible sometimes, then increasing in density to the point of noise while declining in volume. Characteristic of rooms and many natural structures.

ring-modulation - multiplying one sound against another creates inharmonic sum and difference products. This has a "dirty", "metallic" or "weird" effect, like the Daleks in *Dr. Who.*

selective high-Q filter - a filter that has a very sharply defined band-pass response, passing mainly a very narrow frequency range. Q refers to "Q-factor" an electrical term quantifying the sharpness of the spectral peak.

series - an electrical term meaning "chained", or "one after another". In this context not to be confused with a 12-tone series.

spectral - refering to the audio frequency range 20 Hz - 20 kHz, attention is drawn to the use of or distribution of certain bands of frequencies. If two sounds are "spectrally linked", there are common spectral features present.

stand-alone - existing by itself rather than integrated into a whole, self-contained.

static - in this context, noise that is intermittent or "grainy", having some discernible temporal detail. This is a refinement of the everyday term.

sub-audio - at or below the limit of human hearing, typically below 30 Hz. These frequencies can still be perceived by the body, as well as in combination with other frequencies within our hearing range due to modulation effects.

sub-harmonics - frequencies at half, third, etc. of the original, the "mirror image" of the term "harmonics".

Kimmo Vennonen

MIRAGE

for tape

moderate playback volume

duration: 8'45"

PREFACE

This work is submitted in part fulfilment of the requirements for the Graduate Diploma in Music (Computer Music). The stereo version is presented on a Digital Audio Tape included with this submission. A programme note follows.

Kimmo Vennonen

MIRAGE

for tape

moderate playback volume

duration: 8'45"

Programme Note

Mirage is an exploration of a technique I have developed, that I will call complex feedback. Multiple effects processors are combined with a mixer and many patch leads, resulting in enormous timral possibilities with a literally chaotic origin. Sound is almost pulled out from thin air . . . the starting point of everything is silence, and the feedback may be made to disappear into nothingness just as easily.

I interact with the controls on this mountain of equipment, pushing the feedback towards one set of possibilities or another . . . nothing can be guaranteed to sound the same a second time, so everything must be recorded. Some parts of this sonic territory have been traversed often, but there are always other possibilities . . . new ways of walking, new places to go.

There comes a time when exploration pauses and assessment begins. A small, related subset of the recorded sounds is arranged into this current form. This is difficult . . . the sounds have their own life, their own organic way of moving, a resistance to being combined and organised.

Eventually a way is found, that preserves something of this life. I realise that this is important in any creative act . . . nature speaks given the chance.