Methods of Spatialization in Computer Music Composition

Don Bosley

Revised : May 2012
Contents

1 INTRODUCTION 2

2 BINAURAL HEARING: HOW THE BRAIN INTERPRETS SPACE 3

3 FROM MONO TO STEREO AND BEYOND 4
  3.1 The Development of Stereo 4
  3.2 Multi-Speaker Arrays 5
  3.3 Enhanced Spatial Imaging 5

4 REVERBERATION 5
  4.1 A History of Artificial Reverberation 5
  4.2 Recent Models and Methods 7
  4.3 The “Current Future” of Reverb 8

5 MULTIPLE MICROPHONE TECHNIQUE 8
  5.1 General Concepts in Microphone Technique 8
  5.2 Multiple Microphones on Monophonic Sources 9
  5.3 Multiple Microphones on Polyphonic Sources 9

6 “ZONING” SOURCES 10
  6.1 Crossover and Equalization Techniques 10
  6.2 Fourier Transform 11
  6.3 “Zoning” Stringed Instruments 11

7 CONCLUSION: COMBINATIONS OF TECHNIQUES 12
1 INTRODUCTION

In the beginning, the “performance” of computer and electronic music generally involved projects realized behind studio doors and shared via tape playback of the pre-recorded sounds. Recent developments, particularly increases in computing power in the form of portable laptops with equally portable analog to digital audio interfaces, have changed how and when electronic and computer music is made. Processes that used to take hours of machine computation can now be executed in real time and are coalescing methods generally reserved to studios and laboratories with those of live performance. Aesthetically, there are still issues with the laptop as a musical instrument where audiences feel it is in “violation of the codes of musical performance.” [1, pg.101] The merits of laptop performance are not to be questioned here, however a certain group of methods will be analyzed.

One element that computing power has enhanced and helped foster new possibilities with is in area of the enhancement of space, or the illusion of space. Creating space and depth have long fascinated listeners and composers of both acoustic and electronic music alike. Although there was interest in artificial spatialization techniques in the past, the increased computing power available in more recent history has “led to a renewed interest in sound spatialization providing enormous potential for accelerated development of the field.” [2, pg.58] Although hardware developments and other multichannel playback systems have been in continual development for years, there are a new set of possibilities that analog devices could never have achieved.

This paper aims to discuss methods of creating and enhancing the appearance of space in electro-acoustic music. The direction this will take is to explain the historical evolution of techniques, often citing their roots in studio practices and recording arts, and methods for use in computer music applications today. The understanding of the evolution of techniques, can often give insight and help guide a natural progression as to where developments are going. In some cases, the computer has allowed techniques with no precedent or direct ancestry to be created, creating idioms that exist on the computer alone. Along with the historical perspective, the paper seeks to explain the basic technical aspects of each technique.

In the discussions, the idea of sources of sound continually arise. Rather than reiterate the concept keep in mind that a source has the ability to be placed anywhere in a given audio field. By separating a single sound object into multiple components, whatever the technique utilized may be, an amount of independent control is being granted to modify each source. The other factor to consider in reading is that the effects of certain techniques have contextual meaning. In the case of realistic audio reproduction, some effects may be undesirable or even detrimental to the final sound. Using the same effects in creative composition might yield very interesting, and possibly artistically meaningful results in the conceptual framework of a piece.
2 BINAURAL HEARING: HOW THE BRAIN INTERPRETS SPACE

During the 1950s composers were creating both musique concrete and artificially generated sounds using tape splicing and electronic techniques. In particular Pierre Schaeffer and Karlheinz Stockhausen were interested in the idea of using sound position and movement as a compositional tool.[2, pg.59] Since that time, the idea of “space” has become an integral part of all types of electronic music composition. The concept of spatialization precedes computer music by hundreds of years however. —in a number of famous examples.—

The concept of binaural hearing is what drives a great deal of our desire to give depth to recorded and electronic sound. The nature of our hearing in general has driven improvement after improvement, and system after system which attempts to reproduce sound as naturally and accurately as possible. Binaural hearing is based primarily on the fact that most human beings have two ears which collect sound information independently. Two sources of information provide a number of combined stimuli that are almost virtually impossible to interpret using the measurements from one ear.[3, pg.262] Although the mechanics of the ear are well understood, the brain interpretation of these signals is not always as clear. Hearing is generally considered most highly developed, and at times least understood of the senses. The body encodes and interprets sound in a very unique way compared to the other senses [4, pg.9] The overall process that allows for humans to identify, or estimate the direction and distance of a source sound is called localization.

“In localizing a sound source we assign a two dimensional direction to the sound source and we estimate how far away the source is extent tells us something about the environment within which we are listening to the source.” [4, pgs.2-3]

Different theories suggest that the brain interprets localization information by learned association, while it is also argued that certain components are instinctive or genetically encoded. Sound is quite different for example than say visual stimuli, because sound is actually the conversion and interpretation of gradients in air pressure - a form of kinetic energy. These pressure changes in the air act on a set of “transducers” contained within the ear. In order to determine the origin of a sound, a number of cues are utilized including the following: interaural (including a number of sub-categories), spectral, distance, and dynamic. In reverberant environments, the precedence effect is also a cue factor.[5, pgs.10-16] Interaural cues are used to locate horizontal components of a sound. ITD, or Interaural Time Difference is the measure of the difference in arrival time of a sound from one ear to the other. The level difference between the two ears is called IID, or Interaural Intensity Difference.

The IID, and other factors are also affected by what Carlile refers to as the Head Related Transfer Function, or ‘HRTF.’ [4, pgs.38-45] The HRTF, even with recent developments in technology, is still relatively difficult to accurately measure. Subtle differences in HRTF can yield major differences in the way sound is perceived. Companies such as Sennheiser, have made attempts to study these subtleties, and have developed mathematical models of “average” HRTFs. This has been accomplished by taking physical measurements of random samplings of individuals heads, and correlating the measurements with how each subject responds to a set of aural stimuli. [6] The ears themselves do very little other than conduct these stimuli to the brain, where they are combined and interpreted.

Spectral cues generally indicate the elevation of a signal.[5, pg.13] Signals which strike the top of the head are often perceived as monaural. Spectral cues also tend to come from waves bouncing off of the chest and striking the pinna of the ear. The mixing of the direct sound, along with sound bouncing from the pinna, create enhanced peaks at certain frequencies, and dip points in other portions of the spectrum where frequencies are interacting.

Distance and dynamic cues also have some relationship in localization of the source of a sound. Distance cues are based on a combination of loudness, and the ratio of direct energy to the energy level of reverberant signal(s). Many of the cues that humans receive based on distance and dynamics are compared or related to previous experience with categorization of similar sounds. Having never heard a tuba, it would be difficult for an individual with their eyes closed to judge the physical distance the tuba is from them - especially in a heavily reverberant space. As mentioned, it is believed that distance and dynamic cues are learned reactions, more so than other types of cues. The loudness cue is hindered by very reverberant spaces because sound is distributed in a non-uniform manner. Dynamic cues usually integrate the movement of the head in order to
locate the source of a sound.

The precedence effect, which seems to have a basis more in psychoacoustics than physics, is another important function in localizing sound. This is particularly true in a reverberant environment. Zwicker defines the precedence effect as “the law of the first wave front.” It simply states that our ears perceive a source location as the first sound that reaches us, even if another is louder. The example given is as follows:

“The law of the first wave front has great importance in room acoustics. The localization of the sound wave of musical instruments is in the direction of the stage, even if the level of sound reflected e.g. from a balcony is higher than the level of the direct sound wave.” [3, pg.293]

It is clear that the brain naturally perceives sound in three distinct dimensions, with different information being collected by each ear. When re-creating sound, or synthesizing an electronic sound from nothing, playing on the localization process can be an effective tool in stimulating the listener. It is also necessary in some cases.

There are a number of other effects worth mentioning that happen as a result of changes in direction and distance. Air acts as a natural filter, but with a sharper cutoff for high frequencies than for low. Because of this, distance cues differ when sound is outdoors versus a finite or enclosed space. Over great distances low rumbles will still carry, while higher frequency sound information is lost. Doppler shift is a change in pitch that happens when the distance of a source varies from the listeners position. The common example of Doppler effect is the apparent shift in pitch heard as a train passes a stationary listener.[7, pg.464]

3 FROM MONO TO STEREO AND BEYOND

3.1 The Development of Stereo

Sound recordings revolutionized not only the way sound is heard, but the way sound is experienced. From the advent of Edison’s cylinder, all the way through modern compact discs and portable music players, the development of technology has modified how, where, when, and why we listen to music. This has been complimented by technologies like broadcast media, which have allowed for music and other forms of communication to be taken out of one physical space, and travel virtually any distance into another. In some cases, particularly with the transfer of audio data over the internet, there has also been a drastic change in who certain content is available to.

The first methods used for recording contained all information on a single channel, referred to as monophonic, or ‘mono’ for short. In 1933 Bell Labs introduced a method for binaural playback called “two eared listening.” The concept was only logical given the nature of our binaural hearing, however the world was not ready for stereo during the height of the Great Depression.[8, pg.98] The post World War II technological boom yielded great innovations and refinements. The technological leaps made during the war leaked into everyday facets of people’s lives, from electronic components, to materials developed for military purposes being put into commercial and widespread home use. Audio was no exception, and with financial prosperity and disposable income among many (but not all) post WWII Americans, a demand for quality sound reproduction inside the home developed.

By the mid-1950s stereo tape machines were readily available, and found use among audiophiles and casual listeners alike. In 1957, the standardization by the Recording Industry Association of America of stereo records began the widespread acceptance as the format for home use, which over the course of the next few years would cause the reel-to-reel tape machine to be all but phased out in home use.[8, pg.98] Stereo, although commonplace today, is important because it represents the first step in widespread use of techniques that mimic human hearing. Stereo is achieved by balancing the level of a signal in each channel, called ‘panning’ or ‘pan’ for short. The sound itself doesn’t move, but as the level is raised in one speaker it proportionally drops in the other, giving the illusion of movement. By adjusting the pan a location within the stereo field can be created. Stereo is essentially a play on the interaural time and intensity differences described previously.
3.2 Multi-Speaker Arrays

Multi-speaker arrays have existed for years, and have found common use among consumers in the form of Surround Sound Systems. Roads outlines a number of basic recommendations that can enhance the use of multi-speaker arrays: Use at least a quadraphonic sound projection system, left-right configuration in the back channels reversed. Situate the loudspeakers at opposite corners in an elevated position, which will give the sound a vertical dimension. It is important to give each performer their own amplifier and loudspeaker unit, and to keep it near to the performer so that electric sound radiates from the same point that a performers acoustic sound does - this creates association between the performer and electronic sound, rather than a disconnection. Assemble an orchestra of loudspeakers on-stage which creates an array of individually localized and diverse sounds within the space, similar to the concept of a traditional orchestra.[7, pgs.455-457]

Speaker elements do not have to be placed such that their energy radiates all in the same direction. Spherical configurations of speakers, as well as combinations of speakers placed in vertical and horizontal planes, offer simple implementations of complex alternative sounds. Composers such as Stockhausen, Xenakis, and Boulez, have all taken advantage of speaker systems which contained elements suspended from the ceiling giving the appearance of overhead components.[7, pg.453] The Sennheiser companies Auro 3-D system implements vertically elevated speakers in conjunction with ear level speakers to enhance the appearance of vertical cues. [6]

3.3 Enhanced Spatial Imaging

Although multi-speaker arrays are an excellent solution for enhancing the degree of space contained in an audio field, they are not always practical or available. A solution for enhancing the perceived depth of traditional stereo amplification systems are spatial processors, which enhance the stereo imaging of mixes. [9, pg.251] One example is SPAT, a real-time modular spatial-sound-processing system developed by IRCAM and Espaces Nouveaux for the Max/MSP environment [10,pg.2] SPAT uses different characteristics of human hearing and models them in order to achieve greater immersion in sound fields through simple playback systems.

Ambisonics, ambiophonics, and cross-talk cancellation, all deal with wrapping sound around the user, giving the illusion that a sound is emanating from beyond the actual position of the speaker source points. Cross-talk cancellation in stereo applications evaluates information in each channel, identifies components unique to each channel, and removes portions of each. The result is that the left speaker only plays what the left ear should be hearing, and the right speaker plays what the right ear should be hearing.[5, pg.30] Successful cross-talk cancellation is based on the listener being correctly placed in the ‘sweet spot’ of the stereo field, and also on accurate measure of the HRTF. In a real time performance setting, cross talk cancellation could be problematic for computational and processing delay reasons, but because there is no way that everyone in an audience could possibly be seated in the sweet spot. Reverberant spaces also serve to mix the information that computational power had just been invested in separating.

4 REVERBERATION

4.1 A History of Artificial Reverberation

In a brief discussion of psychoacoustics, Max Mathews states that “reverberation is important to musical quality.” [11, pg.177] He elaborates on this statement by comparing a circuit to an anechoic chamber and stresses the ‘dry’ sound that naturally results from creating sound utilizing such an environment. Reverberation, often referred to as ‘reverb’ for short, is a natural phenomena that occurs in any physical space. Attempts at constructing and manipulating space have been made for over two millennia in order to enhance, or perceivably enhance the character of acoustic sounds in some way. From the Greeks to more recent attempts by architects like Xenakis (who also happened to be a composer), people have designed environments which alter the character or atmosphere that sound, both musical and non-musical, exist in. This has been accomplished through the manipulation and use of construction materials, geometric shapes, and a host of other factors carefully studied over hundreds of years.

Natural reverb is generally described in three distinct stages. These are direct sound, early reflections, and reverberations.[12, pg.171] Direct sound is the sound that the ears receive directly from the source. Early
reflections are signals that have only bounced off of one wall, the ceiling, or the floor - often called first order reflections. Most early reflections occur within a maximum of 100 milliseconds, although many concert halls have times significantly less than that. The third portion of the signal is reverberation which refers to the remaining reflections or ‘tail’ of the signal. The overall function is an inverse logarithm, whose total time is generally calculated as the period it takes for a signal to drop in level by 60 decibels.[13, pg.197]

The earliest simulations of reverberation came in the form of echo chambers, which used electronic equipment to capture natural reverberation. The reverb effect is accomplished by sending a signal to a loudspeaker inside of the chamber which usually contained highly reflective surfaces. A microphone at the end opposite of the speaker would pick up the reflections and send them back to source, where they would be mixed with the dry signal. Echo chambers were effective and fairly natural sounding, but relatively inflexible, took up a great deal of space and were highly dependent on the combined frequency response of the speaker, microphone, and chamber.

The first appearance of a pure electronically simulated reverb begins where many musical developments in the 20th century found their roots - Bell Labs. Bell Labs owned a patent for a device which simulated the time delay imposed on a signal when travelling over long runs of telephone cable. According to Accutronics, when Laurens Hammond invented the Hammond Organ in 1935, early customers complained that the sound was “too dry.” The general sound associated with an organ at the time was the pipe organ, which found typical use in reverb rich spaces. The small, carpeted, low ceiling living rooms of private homes where many Hammond Organs found use did not possess the acoustic properties of rich reverberant space. When looking for a solution to the dry sound of the Hammond Organ, either an employee or Laurens Hammond himself discovered the Bell Labs patent. The device was modified to suit Hammonds needs, and thus the first commercial reverb unit was released. [14]

Today, the spring reverb is common in guitar amplifiers, early modular and portable synthesizers and other devices requiring a relatively inexpensive analog solution. The basic operation of a spring reverb is that sound is turned into an electrical current, which passes through a spring or series of springs contained within a housing, and yields a reverb like result after traveling through the spring. The electricity in the signal causes a driver to vibrate the spring - electrical signals themselves will also cause this phenomena. The spring will allow some energy to pass forward, and other portions of the energy will bounce backwards towards the source. This energy is then eventually pushed forward again by a combination of incoming signal and electrical attraction. Different frequency bands also travel at different speeds through springs. All of these factors add up to give spring reverb the sound it has. However, spring reverb does not generally sound like naturally occurring ambience; it sounds like spring reverb.

EMT (Elektromesstechnik), introduced the first plate reverb in 1957, the model 140. [15] Plate reverbs were for a long period, the most popular reverb device used by professional studios. [9, pg.248] But, as they are physically quite large, relatively heavy (over 400 pounds) and are sensitive to external vibration, they continue to be used mainly in studio-only applications. Plate reverbs rely on techniques that are similar to both the spring reverb, and the echo chamber described above. Plate reverbs are a frame with metal foil plates suspended by springs. A driver receives the electrical signal from a send and vibrates the first foil plate. The second plate is connected to a transducer element, similar to a microphone, and picks up the vibrations that the first plate emanates. The plates themselves never make physical contact and a damping blanket is used to control the vibration of the first plate; this can also change the length of time it takes the sound to decay. The resulting signal simulates a more natural sounding reverb than spring, and takes up less space than an echo chamber, but still lacked certain natural properties. EMT eventually put out a stereo version of the unit which gave the sound a more realistic approximation, with enhanced spatial character, but still left something to be desired.

In 1961, Manfred R. Schroeder (again an employee of Bell Laboratories) released the article “Natural Sounding Artificial Reverberation”, which criticized the mechanical techniques of the day, and outlines the classic concepts in modeling reverb that are still used today. What is particularly interesting is that Schroeder partially developed his model as a method for “the electro-acoustic conversion of concert halls.”[16, pg.10] Feeling that certain spaces were too dry, he sought to add ‘life’ to them and found major flaws in the electro-mechanical methods that were available at the time. The two major drawbacks described are that these devices did not posses a flat amplitude-frequency response, and that the density of echoes they generated were not enough to sound like a true reverberant space. The model he provides is simply put into two essential elements: the first is a short time delay that realistically simulates the initial echoes or flutters of a
room; the second element is a series of delays that continue to feedback into themselves and into each other, in order to simulate the filling of a room, as well as natural logarithmic decay. These networks of delays are called “tapped re-circulating delay lines” or TRD for short. Schroeder’s reverb algorithms took hours to run on the computers of the day.[7, pg.477]

Nothing changed drastically in reverb technology until the release of the first digital reverb unit in 1976. Again EMT was at the forefront, releasing its model EMT 250, which like its plate reverb counterpart, still finds use in studios today. [15] Originally sporting a price tag of over 20,000 dollars, the unit offered a number of very unique features. The unit was based on algorithms designed primarily by Dr. Barry Blesser, who is best known as the founder of the company Lexicon. Unlike its predecessors, which essentially offered a crude form of wet/dry mix, and dampening as their only options, the EMT 250 was able to modify a number of parameters within the reverb algorithm, giving the user a unique amount of control over facets of the sound that was previously impossible.

Chamberlain outlines a model for true stereo input and output, citing that delay times should be different between channels in order to increase the realism of the resulting sound.[17, pg.508] Although units had been previously designed to utilize these features, it is often the case that the design processes, which are time consuming, costly, and detailed, are kept private for commercial reasons. [13, pg.198] Chamberlains model retains the directionality of the original signal while allowing a stereo imaged reverb to unfold. Miller Puckette, essentially adds one more element of weight to the model : the amount of delays in the initial portion of the signal. Puckette uses a mix of stages which result in between 64 and 256 echoes on the onset of a sound before going into the next stage of delay loops. [13, pg.196] Puckettes model achieves an excellent density of initial echoes - one of the two major factors that Schroeder cited as a problem.

4.2 Recent Models and Methods

The next generation of reverberation, or at least the start of a new generation is appearing in computer based plug-ins called convolution reverb. Although SONY released a hardware unit capable of real-time convolution called the DRS-777, it wasn’t until that release of Audioeases Altiverb in 2002 that convolution reverb began to gain widespread use. [18, pg.72] Convolution is a computationally expensive process, especially when dealing with audio signals stored at high sampling rates. So, what exactly is convolution? In audio, convolution is a combination process of both scaling and delaying, whereby one signal is added and multiplied to each element of another signal.

In terms of reverberation, the two signals that are convolved are generally a source signal, and an Impulse Response (IR) of a physical space. The impulse response can be obtained by recording the sound of a short signal with very broadband frequency range occurring in the space to be modeled. Ideal impulse responses are as short as possible (i.e. 1 sample), and contain an equal amount of all frequencies (white noise). Common real world methods for generating an IR include popping a balloon or firing a starter pistol/cap gun. The recorded IR is the frequency response of the space as it evolves over time; it is colored only by the quality of the impulse and the recording equipment used to capture it. The resulting sound generated by the convolution process is a very realistic approximation of the sound of the source being played in the space. Convolution does not have to be limited to reverberation though; the IRs of other physical devices such as microphones, speakers, cabinets, cymbals, etc(use your imagination) can be used to impart the quality of one sound onto another.

The last two forms of reverberation (and sound modeling in general) made possible by modern computing power, are the use of geometric and waveguide modeling. Geometric modeling simulates the physical shape and character of a room, however it generally fails to account for the diffusion of sound rays as they reflect off of surfaces, resulting in an unrealistic sound. [19, pg.24] As a solution, scattering algorithms can be applied to the simulated reflections, as well as incorporation of a traditional set of recirculating filters as outlined by Schroeder.

A waveguide is a computational model that outlines the way in which waves travel through a given medium, particularly the way it diffuses. [7, pg.487] The waveguide approach to reverberation relies on what Roads refers to as a set of bi-directional delay lines. Each delay line is a different length which simulates the different echo times that are present when sound energy is diffused throughout a space. The sound is re-circulated (as with many other methods that simulate reverb) and slowly peels away certain information, whether amplitude, or amplitude at certain frequencies to simulate a natural smooth decay. The waveguide
method is more realistic in simulating the diffusion of rays, however it is far more computationally expensive than geometric modeling, which is already very computationally expensive itself.

4.3 The “Current Future” of Reverb

Although accurate simulations of reverberation are exciting, and often times finds appropriate use, they are practical. The excitement really lies in the prospect of creating spaces that aren’t real, or even physically possible, and computer software allows this to happen flexibly in real time. These unrealistic effects provide a palette of sounds for computer musicians and performers alike to tap into, that will allow them to diffuse sound in the spirit of spatialization, but without the conventions associated with natural sounding reverb.

When utilizing waveguide reverberation, smoothly varying delay lines can generate an effect of moving walls.[7, pg.489] Puckette outlines a process to create an infinite reverb, taking portions of a sound and extending them in time.[13, pg.198] By utilizing a gate function, sound can be let into the reverberator, and once closed the delay line can be set to re-circulate infinitely without fear of over saturation. Granular reverb can be accomplished via convolution and intermingles any number of sonic grains with the source signal to generate an amalgam of simultaneous reverberated sounds. Granular reverberation is similar to the effect of sound in the atmosphere being reflected off of clouds. [7, pg.487] Another real-time application of convolution reverb would be to graft additional reverberant character onto a physical space. By measuring the IR of a physical space, and comparing it to the IR of another space, whether real or simulated, a difference impulse can be calculated. This method, comparable to additive and subtractive synthesis methods, can create a seamless effect that can be accomplished consistently within any space (given certain limitations).

5 MULTIPLE MICROPHONE TECHNIQUE

5.1 General Concepts in Microphone Technique

As sound recording evolved from mono to stereo, many new microphone and mixing techniques arose to satisfy the sonic curiosities of listeners and recording engineers alike. Not only were mixing techniques utilized to achieve natural sounding recordings, techniques were also developed to intentionally use stereo to achieve unnatural effects as well. As technology has progressed and more complex multichannel realms continue to come to life, the spirit of experimentation seems to have disappeared in favor of the ‘safety mixing conventions’ which present the image of recorded music and sound in a “realistic” fashion. The use of multiple microphones can be applied to live performance as well, and its creative applications tend to be in-line with the early spirit of experimentation in mixing multi-channel sources.

With any multi-microphone set-up, there are many considerations. “Bleed”, which represents how much of each ‘separate signal’ will arrive in other microphones is the first, and particularly for two reasons. The first reason is the directionality of microphones chosen will be greatly affected by how much bleed is desired. Distance, can also effect the need to increase gain resulting in increased background noise and possible leakage from other sources. In the case of electronic modification, this can yield undesirable results.

The second reason is phase cancellation based on microphone type and placement has to be considered. Phase or acoustic phase cancellation occurs when a source signal is not equidistant to the microphones receiving it. [20, pg.109] Phase cancellation as discussed in terms of binaural hearing, can naturally occur and aid in the localization process, however phase cancellation in microphone applications can cause undesired, even crippling results if not controlled. Other factors which can cause phase problems, and amplify phase problems so to speak, are things like reflections off of floors and music stands.

There are two basic solutions which can help, and even solve, phase cancellation. In Practical Recording Techniques, Bartlett and Bartlett outline the rule of “3 to 1”, as a way to avoid phase cancellation. It is recommended that microphones are placed three times the distance from each other than they are from the source.[9, pg.123] For example, if a saxophone had two microphones being placed on it, one right at the bell approximately an inch away, a microphone picking up the keys at two inches would need to be six inches from the other microphone. Another basic solution is to output each microphone from its own amplification source. Although this is not a guarantee for removal of phase cancellation, it can reduce the effects heard by having the sound emanate from different sources. With the coloration of phase, one should also be aware
of the off-axis coloration, which is that some microphones become dull or colored when they are not aimed directly at the sound source.[9, pg.124]

Another characteristic of microphone use is referred to as the proximity effect. By changing the distance of a microphone, the frequency response of the microphone is varied.[20, pg.91] This could be useful in the fact that by placing different microphones of the same type at different distances, signals with different emphasized (and de-emphasized) components could be sourced for placement at different points in the audio field. The proximity effect could also be used as pre-treating for environments where components are going to be spatialized by bands of the spectrum. According to Burroughs (who was the former president of Electro-Voice), “proximity effect may be used to add depth and fullness to a thin voice or instrument.” [20, pg.96]

5.2 Multiple Microphones on Monophonic Sources

Imagine a trumpet being played. Although sound seems to radiate from essentially one point, there are other sounds that although not as strong, that are simultaneously occurring. The light clicking of valves, the subtleties of the sound of the player breathing, or even their lips coming to and from the mouthpiece of the instrument. By using an array of well placed microphones, not only could the sound we naturally associate with the trumpet be captured, but all of the other subtle nuances of the performance could be captured as well. The usefulness is two fold; the other elements could be used as a raw audio for placement through a sound field, or individual processing. The second use is that the extra microphones could also be converted into control streams used to control parameters of other sounds during an electroacoustic performance.

5.3 Multiple Microphones on Polyphonic Sources

Polyphonic in this discussion does have the same contextual meaning as polyphonic in traditional musical terms where any instrument that can produce multiple notes simultaneously (versus a monophonic instrument only capable of producing one note at a time) is considered polyphonic. Instruments capable of polyphony are not necessarily polyphonic sound sources. In order to have a true poly or multi sound source, an instrument must have sources that are physically separated, or even isolated. To give an example, the acoustic guitar is capable of polyphony, but all of its sounds radiate from essentially the same source.1 If a microphone picking up a guitar is shifted the frequency response, or even apparent timbre of the signal may change, however the pitches heard will remain constant. Although a multistringed instrument as well, the piano has an individual hammer and set of strings for each note it produces. As a microphone is moved towards the bass strings away from the higher register, it will favor those pitches. Through careful microphone placement, this provides the ability to treat pitch registers as independent sources.

Drum sets and percussion batteries also represent an interesting prospect in multiple microphone or multiple source technique. The different sources contained within a drum set are physically spaced apart in a manner that is allows each physical sound source to be captured in an isolated fashion. Typical overhead microphone techniques used on drums incorporate either a Spaced A/B pair, or a Coincident X/Y pair. The X/Y pair holds the advantage of no phase issues because source sounds arrive at the microphone elements simultaneously. However, the spaced pair gives increased stereo imaging, but at the risk of phase problems.[21, pg.38]

The microphones that are in common use in studio and other high quality applications are all based on the principle of some force of air acting on a diaphragm. Although differences arise in the nature of the diaphragm, the basic premise remains the same. There has arisen a new generation of thought in microphone usage, particularly among computer musicians in which alternate methods of picking up signals are being utilized. Microphones based on lasers, optical, and other technology are all being manipulated but one common and relatively inexpensive method that is being experimented with is contact and piezo microphones.

What is a piezo and how does it work? Piezoelectricity is the formation of an electrical signal or electrical polarity which is a direct translation of an applied pressure.[22, pg.177] As a square is rhombus, but a rhombus in not necessarily a square, piezo elements can be made into contact microphones, but are not always used that way. Piezo elements can come in two basic forms: elements which accept a pressure input called sensors,

---

1There is a special case with “hexaphonic” and other pick-up devices which will be discussed later. The Use of Alternative Microphones
and elements which exert some sort of pressure on another media, commonly air, called actuators. The first of these forms is related but not equivalent to the operation of a microphone and the second type can more closely resemble a speaker, or drive an external element to resemble a speaker. True contact microphones do not utilize any air pressure in order to generate electrical signal, alternatively they transduce the vibrations of a given media directly into electrical energy.[23, pg.121]

Because of the small size of many piezo elements, it is possible to place them in tight spaces, in contact with small objects, and use multiple instances without being physically obtrusive. Imagine having an individual element attached to a mouthpiece, a prominent key, and the bell of a saxophone. The sonic possibilities are endless and the interaction that could be generated using any element as a control signal is extremely powerful but still directly sensitive to the human performance. Even more advanced applications are possible, such as utilizing piezo elements to create a form of acoustically generated convolution.[24, pg.46] Obvious problems include limiting or modifying the resonant characteristics of the instrument. In woodwinds and brass, attaching foreign elements likely affects the overtone content, as well as the players ability to perform certain functions unhindered. This could include making certain registers difficult to play in, as well as limit the dynamic range of the instrument itself, all by slightly changing the way the air columns and materials react to the performers input. Sustain and decay times would almost invariably be shortened as well.

Although piezos have found use in a variety of professional applications, computer musicians seem to be gravitating towards the inexpensive low fidelity type of element. Mark Applebaum, an associate professor of composition on the faculty of Stanford University, has constructed and implemented in performance, a number of instruments he defines as “sound sculptures.” The common functional element they all have is how sound (or vibration) is converted into an electrical signal - piezos. According to Applebaum:

The pickups are 50 cent surplus piezo contact elements; there are eight pickups on the mousetrap, four per stereo side; the duplex mausphon is also a stereo instrument with one pickup on each of its two levels; mini and midi each have one pickup. [25]

The piezo itself is a compositional tool in the experimental phase at this point, and as such people are concerned with the results, rather than things like numbers on a typical specifications sheet. Eric Leonardson, who has also constructed sound sculptures, sums up the state of using a piezo element quite nicely:

Part of the beauty of the piezo contact mic lies in its ability to allow the sound explorer a way to get around the engineering and design challenges that a “proper” acoustic instrument presents, the interrelated material properties of resonance, impedance, mass, etc. Once one has a mixer, amplifier and loudspeakers, the price of a piezo contact mic is negligible, and its ability to act as an aural microscope into unknown sonic yet entirely physical aspects of any object or material is truly exciting, if not amazing. [26, pg.17]

6 “ZONING” SOURCES

6.1 Crossover and Equalization Techniques

A simple solution to separating components of a signal based on register is the use of traditional filtering, particularly the methods associated with what we call equalization. Typical units that accomplish cross-over functions, low, high, band, and stop-band filters can also be particularly effective in the separation of single sources into multiple zones. With traditional equalization, it is difficult to achieve ideal filtering even when implemented digitally. Ideal filter characteristics have brick-wall properties, that is a sharp slope or cutoff at the corner frequency and a flat response in the pass band.[27, pg.55] Generally what occurs when any filtering mentioned is used is one of two things: resonant notches (or peaks), and/or dropout. Resonant notches result in the coloration of the given input at certain frequencies, and dropout results in the loss of energy at certain frequencies. It is possible to implement different filter types in order to achieve smoother response, or greater isolation, but the user needs to be aware that the coloration of the signal will be directly correlated to the strength, or order of the filter.
6.2 Fourier Transform

The Fourier Transform offers an alternative to traditional equalization zoning methods. The Fourier Transform, which decomposes a signal into its component sinusoids, is easily implemented in real-time in many software packages. Discrete Fourier Transforms (DFTs), implemented in algorithms called Fast Fourier Transforms (FFTs) convert time domain signals into the frequency domain, and separate component sinusoids into uniformly spaced frequency bins. The resolution of the bins is based on the sampling rate of the signal, divided by the length of the samples in the FFT, generally a power of two. As an example, a signal sampled at 44.1 kHz, with an FFT length of 512, would yield bins with center width starting at 0 Hz (the DC content of the signal) and increasing by intervals of 86.132 Hz all the way up to the sampling rate of 44.1 kHz.

In certain software packages it is possible to filter the output of the FFT based on the number of the bins. For example if there were 512 bins, any bin below 100 (or approximately 8520 Hz) could be filtered out one channel, and all information above bin 100 could be filtered out to another channel. The FFT, and the subsequent Inverse FFT which brings the signal back to the time domain, does take some time to implement. The amount of time it takes is dependent on factors like system speed, algorithm, and window length but in most cases the time will be unperceivable to a listener. FFT bins can exhibit spectral leakage between bins, and even between overlapping windows, although generally speaking it will not color the sound the way traditional filtering does. Traditional filtering also offers a linear selection of corner and cutoff frequencies, whereas selecting bins means selecting values in fixed steps, slightly reducing the amount of precision with which zoning can be carried out.

It should be noted that in both the case of FFT zoning, and equalization zoning, that appearances of information will occur in both channels. Most signals are complex and as such contain signal components at many frequencies. Setting a cutoff based on spectral or frequency content will almost always mean that fundamentals and certain pitch information will travel to one zone, while related upper partials and overtones will travel to another. This can actually be quite useful by giving a multi-channel image to a single complex tone.

6.3 “Zoning” Stringed Instruments

For multiple reasons, string instruments represent a unique set of possibilities in zoning and spatialization in general. Instruments capable of polyphony such as single manual keyboard instruments and mallet percussion have only one resonant object that represents each pitch or sound that it can produce. Yes, a piano has three strings for some pitches, but the mechanism that actuates one string actuates all three, so in this case we consider them one object. String instruments are the only instruments capable of a true simultaneous unison (or even three or more of the same pitch), but they have individual objects that are not limited to a single pitch, rather capable of generating a variety of pitch information.

In 1986 Max Mathews modified a violin by placing one contact microphone on each string. He subsequently ran an output from each contact microphone to its own amplifier, which in turn each drove a loudspeaker strategically placed throughout the performance space. [28, pg.4] This technique achieves a sense of depth that extends the source field far beyond the foundation of the violin, essentially making the violin the “size” of the distances between speakers. Even through a two channel system, the source information could be panned throughout the field to give an illusion of spread.

Although adding pickups to acoustic stringed instruments is possible, pickups are an inherent feature of the electric guitar. Guitars have implemented electromagnetic pickups and piezo systems to transduce the vibration of strings into an electric current for years; Armand Knoblaugh applied for a patent on an electric pick-up in 1938, and certainly was not the first to attempt to design or even implement such a device [29, pg.17] The limitation is that a traditional guitar pickup outputs the signals from every string simultaneously. The technology that has propelled the concept of polyphonic output stems from MIDI guitar technology, which utilizes split guitar pick-ups. This is generally for the purpose of sending signals from each string where they can be converted into data for use in a number of applications like synthesis and sequencing, however other commercially available units focus on processing the direct audio signal. Miller Puckette, among others, utilizes “hexaphonic” pick-ups produced by Roland Corporation in order to tap the six individual strings audio signals, and convert them into individual sources.
7 CONCLUSION : COMBINATIONS OF TECHNIQUES

None of the techniques mentioned here are exclusive, in fact in many cases they can be intermingled and give complimentary results. Because of scene memory and other adaptive methods in performance technologies, different techniques could be used at various points in a performance for optimization or effect. Performers can exert a control by repositioning microphones allowing for an instantaneous and broad shift in the dynamic character of the events that follow. All sources can be converted into control signals, used as raw audio, or combined to generate some combination of both.

There are a wide variety of ways to enhance spatialization and depth in electro-acoustic performance, and the computer is the instrument that has handed the power to musicians to accomplish many of these things in real time. As such, power is not automatically equated with artistic sense - just because it can be done doesn’t mean it should, and certain criteria, from audio quality to emotional content, should be applied by the creator in evaluating the results of their action. Using discretion and clear vision, spatialization techniques can add an exciting and constantly fresh aspect to electro-acoustic performance and the creation of electronic music in general.
References


Showcasing spatialization technologies in the new Unreal Audio Engine. Play music from Statue Three, which is located just inside the temple corridor. Auditioning Plugins. By default, the Project is set to use Built-in spatialization plugin. You can change this setting to try the effects of spatialization plugins from other third-party vendors. To audition any of the other plugins, use the following general procedure: Enable the third-party plugin you want to try. What compositional methods or considerations can composers turn to when 3D virtual audiovisual spaces? With this research, practical and particular solutions to both of these problems, artistic and technical, are evaluated and offered through the development of a new software tool and by designing exploratory virtual reality-based compositions as live stage performances. To summarize the contributions of this research 15 Vincenzo Lombardo et al., "A Virtual-Reality Reconstruction of Poème Electronique Based on Philological Research," Computer Music Journal 33, no. 2 (Summer 2009): 26. 16 Ibid., 30. 17 Thom Holmes, Electronic and Experimental Music: Technology, Music, and Culture, 4th ed.