CHAPTER 17

How to Achieve Depth and Dimension in Recording, Mixing and Mastering

I. Introduction

I placed this acoustics lesson in the middle of a book on mastering because the creation of wonderful audio masters requires that some basic acoustic principles be understood. As we enter the era of surround recording and reproduction, many mix engineers are repeating their mistakes from two-channel work—panpotting mono instruments to discrete locations, and then adding multiple layers of uncorrelated stereophonic reverb "wash" in a vain and misguided attempt to create space and depth. It's important to learn how to manipulate the surprising depth available from 2-channel canvas before moving on to multi-channel surround.

It amazes me how few engineers know how to fully use good ol' fashioned 2-channel stereo. I've been making "naturalistic" 2-channel recordings for many years taking advantage of room acoustics, but it is also possible to use artificial means to simulate depth, and there are many engineers working in the pop field who know how to do so. Learn to discern the audible difference between simple pan-potted mono, and recordings which simulate or utilize the reflections from nearby walls to create a real sense of depth. Without such knowledge, your recordings will tend to produce a vague, undefined image; the musical instruments will be obscured and unclear.

Techniques here include using the Haas[1] effect, particularly when implemented binaurally, use of delays and alteration of phase, more naturalistic reverberators, and understanding how to unmask via placement. Also be aware that well-engineered

2-channel recordings have encoded ambience information which can be extracted to multichannel, and it pays to learn about these techniques.

Depth Perception in Real Rooms

Early Reflections versus Reverberation

At first thought, it may seem that depth in a recording can be achieved simply by increasing the proportion of reverberant to direct sound. But the artificial simulation of depth is a much more complex process. Our binaural hearing apparatus is largely responsible for the perception of depth and space, decoding the various early reflections from nearby walls that support and strengthen the sound of musical instruments and voices. First, we must define the terms *early reflections* and *reverberation*. Early reflections consist of the part of the room sound within approximately the first 50-100 milliseconds. There is a great deal of correlation between the direct sound and the early reflections; you can think of the early reflections as being *attached* to the direct sound. In a large and diffuse room, after about 100 milliseconds, enough wall bounces have occurred to make it impossible to hear discrete bounces; this is the onset of random (uncorrelated) reverberation, which we can say is *detached* from the direct sound. That's why it is the early reflections, even more than the reverberation, which largely affect our perception of the depth of the sound, giving it shape and dimension. The ear's decoding ability is such that a few simple well-placed echos actually solidify and clarify the location of the direct sound; this is why a simple, dead, panpotted mono source (without early reflections) is so hard to locate precisely.

Masking Principle/Haas Effect

Recording engineers were concerned with achieving depth even in the days of monophonic sound. In those days, many halls for orchestral recording were deader than those of today. Why do monophonic recording and dead rooms seem to go well together? The answer is involved in two principles that work hand in hand: 1) The masking principle and 2) The Haas effect.

The Masking Principle and Mono versus Stereo Recordings

The masking principle says that a louder sound will tend to cover (mask) a softer sound, especially if the two sounds lie in the same frequency range. If these two sounds happen to be the direct sound from a musical instrument and the reverberation from that same instrument, then the initial reverberation can appear to be covered by the direct sound. When the direct sound ceases, the reverberant hangover is finally perceived. This is why in mixing, we often add a small delay between the direct sound and the reverberation, it helps the ears to separate one from the other, reducing the masking.

In concert halls, our two ears sense reverberation as coming diffusely from all around us, and the direct sound as having a distinct single location. Thus, when music is perceived binaurally, there is less masking because the direct and reverberant sound come from different directions. However, in monophonic recording, the reverberation is reproduced from the same source speaker as the direct sound, and so we may perceive

the room as deader than it really is, because the two sounds overlap directionally. Furthermore, if we choose a recording hall that is very live, then the reverberation will tend to intrude on our perception of the direct sound, since in monaural, both will be reproduced from the same location-the single speaker.

This is one explanation for the incompatibility of many stereophonic recordings with monophonic reproduction. The larger amount of reverberation tolerable in stereo becomes less acceptable in mono due to the physical overlap. As we extend our recording techniques to 2-channel (and multichannel) we can overcome masking problems by spreading artificial reverberation spatially away from the direct source, achieving both a clear (intelligible) and warm recording at the same time. One of the first tricks that mix engineers learn is to put reverberation in the opposite channel from the source. This helps unmask the sound, but can produce an unnatural effect.[2] As we get more sophisticated, we discover that instead of hard-panning the source and its mono echo or reverb return, using multiple delays or stereophonic early reflections can yield a far more cohesive, natural effect. The presence of the stereophonically-spread early reflections also serves to clarify the location of the dry source. In a sophisticated stereo mix, engineers take advantage of variations on these themes to produce variety and space in the recording.

The Haas Effect

The Haas effect can help overcome masking. In general, Haas says that echoes occurring within approximately 40 milliseconds of the direct sound become fused with the direct sound. We say that the echo becomes "one" with the direct sound, and only a loudness enhancement occurs; this is what happens in a real room with the earliest wall and floor reflections. Since the velocity of sound is approximately one foot per millisecond, 40 milliseconds corresponds to a wall that's 20 feet distant (assuming a flat wall perpendicular to the angle of the direct sound).

A very important corollary to the Haas effect says that fusion (and loudness enhancement) will occur even if the closely-timed echo comes from a different direction than the original source. However, the brain will continue to recognize (binaurally) the location of the original sound as the proper direction of the source. The Haas effect allows nearby echoes (greater than about 10 ms. and less than about 40 ms. delay) to enhance and reinforce an original sound without confusing its directionality. The maximum definition of the source's directionality will occur using the longest delay possible that is not perceived as a discrete echo.

The Magic Surround

We can take advantage of the Haas effect to naturally and effectively convert an existing 2-channel recording to a 4-channel or surround medium. When remixing, place a discrete delay in the surround speakers to enhance and extract the original ambience from a previously recorded source! No artificial reverberator is needed if there is sufficient reverberation in the original source. Here's how it works:

Because of the Haas effect, when the delay and source are correlated (e.g., a snare drum hit) the ear fuses them, and so still perceives the direct sound as coming from the front speakers. But this does not apply to ambience because it is uncorrelated—the ear does not recognize the delay as a repeat, and thus ambience will be spread, diffused between the location of the original sound and the location of the delay (in the surround speakers). Thus, the Haas effect only works for correlated material; uncorrelated material (such as natural reverberation) is extracted, enhanced, and spread directionally. Dolby laboratories calls this effect *the magic surround*, for they discovered that natural reverberation was extracted to the rear speakers when a delay was applied to them. Dolby also uses an L-minus-R matrix to further enhance the separation. The wider the bandwidth of the surround system and the more diffuse its character, the more effective the psychoacoustic extraction of ambience to the surround speakers.

Haas In Mixing

There's more to Haas than this simple explanation. To become proficient in using Haas in mixing, you can study the original papers which discuss the various fusion effects at different delay and amplitude ratios. During mixing, remember the 1 foot per millisecond relationship, and see what happens with carefully-placed and leveled delays in the 12 to 40 millisecond range. You will discover that they can enhance an instrument's clarity and position all due to psychoacoustics: the ear's own decoding power.[3] In fact, Haas delays are far more effective than equalization at repairing the sound of a drumset which was recorded in a dead room, for example. Furthermore, multiplying the delays until they simulate the complex early reflections of real rooms can greatly improve our stereo mixing technique. More than a few delays is beyond our ability to do on a simple mixing board, and for early reflections we must use computerized simulations found in devices such as the TC Electronic, EMT, and certain models of Sony reverbs. The latest algorithm from TC, currently only available in the System 6000, is quite astounding.

Haas In Mastering

We often receive recordings for mastering which lack depth, spatiality and clarity because the mix engineer did not mix the early reflections or reverberation well enough or loudly enough. But since the mix has already been made, adding artificial reverberation can muddy the sound. This is why an ambience extraction technique should be employed instead. My K-Stereo processor, model DD-2, can enhance the depth of existing stereo mixes by extracting and spatially-spreading their inherent ambience.

Haas' Relationship To Natural Environments

In a good stereo recording, the early correlated room reflections are captured with their correct placement; they support the original sound, help us locate the sound source as to distance and do not interfere with left-right orientation. The later uncorrelated reflections, which we call reverberation, naturally contribute to the perception of distance, but because they are uncorrelated with the original source the

reverberation does not help us locate the original source in space. If the recording engineer uses stereophonic miking techniques and a more lively room instead, capturing early reflections on two tracks of the multitrack, the remixing engineer will need less artificial reverberation and what little he adds can be done convincingly.

Using Frequency Response to Simulate Depth

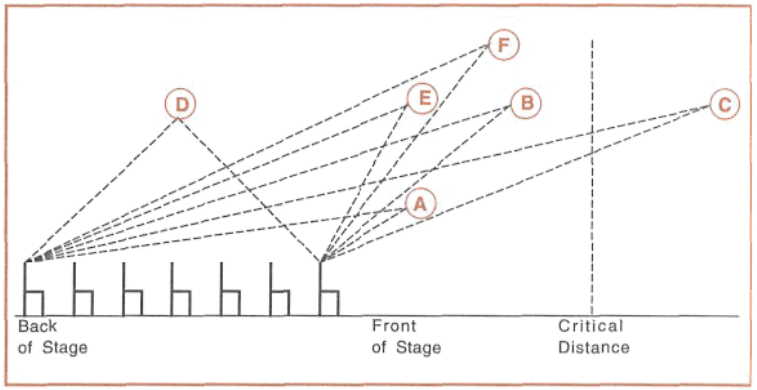
Another contributor to the sense of distance in a natural acoustic environment is the absorption qualities of air. As the distance from a sound source increases, the apparent high frequency response is reduced. This provides another tool which the recording engineer can use to simulate distance, as our ears have been trained to associate distance with high-frequency rolloff. An interesting experiment is to alter a treble control while playing back a good orchestral recording. Notice how the apparent front-to-back depth of the orchestra changes considerably as you manipulate the high frequencies.

Recording Techniques in Natural Rooms to Achieve Front-To-Back Depth

Balancing the Orchestra with only a few micophones (minimalist). A musical group is shown in a hall cross section (see diagram at right). Various microphone positions are indicated by letters A-F.

Microphones **A** are located very close to the front of the orchestra. As a result, the ratio of **A**'s distance from the back compared to the front is very large. Consequently, the front of the orchestra will be much louder in comparison to the rear, and the amount of early reflections reaching the microphone from the rear will be far greater than from the front. Front-to-back balance will be exaggerated. However, there is much to be said in favor of mike position **A**, since the conductor usually stands there, and he purposely places the softer instruments (strings) in the front, and the louder (brass and percussion) in the back, somewhat compensating for the level discrepancy due to location. Also, the radiation characteristics of the horns of trumpets and trombones help them to overcome distance. These instruments frequently sound closer than other instruments located at the same physical distance because the focus of the horn increases direct to reflected ratio. Notice that orchestral brass often seem much closer than the percussion, though they are placed at similar distances. You should take these factors into account when arranging an ensemble for recording. Clearly, we perceive depth by the larger proportion of reflected to direct sound for the back instruments.

The farther back we move in the hall, the smaller the ratio of back-to-front distance, and the front instruments have less advantage over the rear.



At position **B**, the brass and percussion are only two times the distance from the mikes as the strings. This (according to theory) makes the back of the orchestra 6 dB down compared to the front, but much less than 6 dB in a reverberant hall, because level changes less with distance.

For example, in position **C**, the microphones are beyond the critical distance—the point where direct and reverberant sound are equal. If the front of the orchestra seems too loud at **B**, position **C** will not solve the problem; it will have similar front-back balance but be more buried in reverberation.

Using Microphone Height To Control Depth And Reverberation

Changing the microphone's height allows us to alter the front-to-back perspective independently of reverberation. Position **D** has no front-to-back depth, since the mikes are directly over the center of the orchestra. Position **E** is the same distance from the orchestra as **A**, but being much higher, the relative back-to-front ratio is much less. At **E** we may find the ideal depth perspective and a good level balance between the front and rear instruments. If even less front-to-back depth is desired, then **F** may be the solution, although with more overall reverberation and at a greater distance. Or we can try a position higher than **E**, with less reverb than **F**.

Directivity Of Musical Instruments

Frequently, the higher up we move the mike, the more high frequencies it will capture, especially from the strings. This is because the high frequencies of many instruments (particularly violins and violas) radiate upward as well as forward. The high frequency factor adds more complexity to the problem, since it has been noted that treble response affects the apparent distance of a source. Note that when the mike moves past the critical distance in the hall, we may not hear significant changes in high frequency response when height is changed.

The recording engineer should be aware of how all the above factors affect the depth picture so he can make an intelligent decision on the mike position to try next. The difference between a B+ recording and an A+ recording can be a matter of inches.

Beyond Minimalist Recording

The engineer/producer often desires additional warmth, ambience, or distance after finding the mike position that achieves the perfect instrumental balance. In this case, moving the mikes back into the reverberant field cannot be the solution. Another call for increased ambience is when the hall is a bit dry. In either case, trucking the entire ensemble to another hall may be tempting, but is not always the most practical solution.

The minimalist approach is to change the microphone pattern(s) to less directional (e.g., omni or figure-8). But this can get complex, as each pattern demands its own spacing and angle. Simplistically speaking, with a constant distance, changing the microphone pattern affects direct to reverberant ratio.

Perhaps the easiest solution is to add ambience mikes. If you know the principles of acoustic phase cancellation, adding more mikes is theoretically a sin. However, acoustic phase cancellation does not

occur when the extra mikes are placed purely in the reverberant field, for the reverberant field is uncorrelated with the direct sound. The problem, of course, is knowing when the mikes are deep enough in the reverberant field. Proper application of the **3 to 1 rule**[4] will minimize acoustic phase cancellation. So will careful listening. The ambience mikes should be back far enough in the hall, and the hall must be sufficiently reverberant so that when these mikes are mixed into the program, no deterioration in the direct frequency response is heard, just an added warmth and increased reverberation. Sometimes halls are so dry that there is distinct, correlated sound even at the back, and ambience mikes would cause a comb filter effect.

Assuming the added ambience consists of uncorrelated reverberation, then in principle an artificial reverberation chamber should accomplish similar results to those obtained with ambience microphones. In practice, however, this has to be a qualified yes, by assuming not only that the artificial reverberation chamber has a true stereophonic response and is consonant with the sound of the original recording hall, but also that the main microphones have picked up sufficient early reflections for the depth effect to be convincing. Artificial reverberation alone, being uncorrelated, will not help the imaging or produce a focused depth picture.

What happens to the depth and distance picture of the orchestra as the ambience is added? In general, the front-to-back depth of the orchestra remains the same or increases minimally, but the apparent overall distance will increase as more reverberation is mixed in. The change in depth may not be linear for the whole orchestra since the instruments with more dominant high frequencies may seem to remain closer even with added reverberation.

The Influence of Hall Characteristics on Recorded Front-To-Back Depth

In general, given a fixed microphone distance, the more reverberant the hall, the farther back the rear of the orchestra will seem. In one problem hall the reverberation is much greater in the upper bass frequency region, particularly around 150 to 300 Hz. A string quartet usually places the cello in the back. Since that instrument is very rich in the upper bass region, in this problem hall the cello always sounds farther away from the mikes than the second violin, which is located at his right. Strangely enough, a concert-goer in this hall does not notice the extra sonic distance because his strong visual sense locates the cello easily and does not allow him to notice an incongruity. When she closes her eyes, however, the astute listener notices that, yes, the cello sounds farther back than it looks!

It is therefore rather difficult to get a proper depth picture with a pair of microphones in this problem hall. Depth seems to increase almost exponentially when low frequency instruments are placed only a few feet away. It is especially difficult to record a piano quintet in this hall because the low end of the piano excites the room and seems hard to locate spatially. The problem is aggravated when the piano is on half-stick, cutting down the high frequency definition of the instrument.

The miking solution I choose for this problem is a compromise; close mike the piano, and mix this with a panning position identical to the piano's virtual image arriving from the main mike pair. I can only add a small portion of this close mike before the apparent level of the piano is taken above the balance a listener would hear in the hall. The close mike helps solidify the image and locate the piano. It gives the listener a little more direct sound on which to focus.

Can minimalist techniques work in a dead studio? Not very well. My observations are that simple miking has no advantage over multiple miking in a dead room. I once recorded a horn overdub in a dead room, with six tracks of close mikes and two for a more distant stereo pair. In this dead room there were no significant differences between the sound of the minimalist pair, and the six multiple mono close-up mikes! (The close mikes were, of course, carefully equalized, leveled and panned from left to right.) This was a surprising discovery, and it reinforces the importance of good hall acoustics and especially early reflections on a musical sound. In other words, when there are no significant early reflections, you might as well choose multiple miking, with its attendant post-production balance advantages.

Miking Techniques and the Depth Picture

Coincident Microphones. The various simple miking techniques reveal depth to greater or lesser degree. Microphone patterns which have out of phase lobes (e.g., hypercardioid and figure-8) can produce an uncanny holographic quality when used in properly angled pairs. Even tightly-spaced (coincident) figure-8s can give as much of a depth picture as spaced omnis. But coincident miking reduces time ambiguity between left and right channels, and sometimes we seek that very ambiguity. Thus, there is no single ideal minimalist technique for good depth, and you should become familiar with changes in depth produced by changing mike spacing, patterns, and angles. For example, with any given mike pattern, the farther apart the microphones of a pair, the wider the stereo image of the ensemble. Instruments near the sides tend to pull more left or right. Center instruments tend to get wider and more diffuse in their image picture, harder to locate or focus spatially.

The technical reasons for this are tied in to the Haas effect for delays of under approximately 5 ms. vs. significantly longer delays. With very short delays between two spatially located sources, the image location becomes ambiguous. A listener can experiment with this effect by mistuning the azimuth on an analog two-track machine and playing a mono tape over a well-focused stereo speaker system. When the azimuth is correct, the center image will be tight and defined. When the azimuth is mistuned, the center image will get wider and acoustically out of focus. Similar problems can (and do) occur with the mike-to-mike time delays always present in spaced-pair techniques.

Spaced microphones. I have found that when spaced mike pairs are used, the depth picture also appears to increase, especially in the center. For example, the front line of a chorus will no longer

seem straight. Instead, it appears to be on an arc bowing away from the listener in the middle. If soloists are placed at the left and right sides of this chorus instead of in the middle, a rather pleasant and workable artificial depth effect will occur. Therefore, do not rule out the use of spaced-pair techniques. Adding a third omnidirectional mike in the center of two other omnis can stabilize the center image, and proportionally reduces center depth.

Multiple Miking. I have described how multiple close mikes destroy the depth picture; in general I stand behind that statement. But soloists do exist in orchestras, and for many reasons, they are not always positioned in front of the group. When looking for a natural depth picture, try to move the soloists closer instead of adding additional mikes, which can cause acoustic phase cancellation. But when the soloist cannot be moved, plays too softly, or when hall acoustics make him sound too far back, then one or more *spot mikes* must be added. When the close solo mikes are a properly placed stereo pair and the hall is not too dead, the depth image will seem more natural than one obtained with a single solo mike.

To avoid problems, apply the 3 to 1 rule. Also, listen closely for frequency response problems when the close mike is mixed in. As noted, the live hall is more forgiving. The close mike (not surprisingly) will appear to bring the solo instrument closer to the listener. If this practice is not overdone, the effect is not a problem as long as musical balance is maintained, and the close mike levels are not changed during the performance. We've all heard recordings made with this disconcerting practice. Trumpets on roller skates?

Delay Mixing. At first thought, adding a delay to the close mike seems attractive. While this delay will synchronize the direct sound of that instrument with the direct sound of that instrument arriving at the front mikes, the single delay line cannot effectively simulate the other delays of the multiple early room reflections surrounding the soloist. The multiple early reflections arrive at the distant mikes and contribute to direction and depth. They do not arrive at the close mike with significant amplitude compared to the direct sound entering the close mike. Therefore, while delay mixing may help, it is not a panacea. To adjust the delay of the solo mike(s) properly, start with a delay calculated by the relative distance between the solo mike and the main mike, then focus the delay up and down in 1 ms. increments until the sound is most coherent and focused and the soloist sounds clearest.

Influence Of The Control Room Environment On Perceived Depth

At this point, many engineers may say, "I've never noticed depth in my control room!" The widespread practice of placing near-field monitors on the meter bridges of consoles kills almost all sense of depth. Comb-filtering, speaker diffraction and sympathetic vibrations from nearby surfaces destroy the perception of delicate time and spatial cues. The recent advent of smaller virtual control surfaces has helped reduce the size of consoles, but seek advice from an expert acoustician if you want to appreciate or manipulate depth in your recordings.

Examples To Check Out

Standard multitrack music recording techniques make it difficult for engineers to achieve depth in their recordings. Mixdown tricks with reverb and delay may help, but good engineers realize that the best trick is no trick: learn how to use stereo pairs in a good acoustic. Here are some examples of audiophile recordings I've made that purposely take advantage of depth and space, both foreground and background, on Chesky Records. Sara K. *Hobo*, Chesky JD155. Check out the percussion on track 3, "Brick House." Johnny Frigo, *Debut of a Legend*, Chesky JD119. Check out the sound of the drums and the sax on track 9, "I Love Paris." Ana Caram, *The Other Side of Jobim*, Chesky JD73. Check out the percussion, cello and sax on "Correnteza." Carlos Heredia, *Gypsy Flamenco*, Chesky WO126. Play it loud! And listen to track 1 for the sound of the background singers and handclaps. Phil Woods, *Astor and Elis*, Chesky JD146, for the natural-sounding combination of intimacy and depth of the jazz ensemble.

Technological Impediments to Capturing Recorded Depth

Depth is the first thing to suffer when technology is incorrectly applied. Here is a summary of some of the technical practices that when misused, or accumulated, can contribute to a boringly flat, depthless recorded picture:

- Multitrack and multimike techniques
- Small/dead recording studios or large rooms with poor acoustics/missing early reflections
- low resolution recording media
- amplitude compression
- improper use of dithering, cumulative digital processing, and low-resolution digital processing (e.g., using single-precision as opposed to double or higher-precision)

In Summary: When recording, mixing and mastering—use the highest resolution technology, best miking techniques, and room acoustics. Process dead tracks with Haas delays and early reflections, and specialized ambience recovery tools. Then you'll resurrect the missing depth in your recordings.

---

1 Haas, Helmut (1951), *Acustica*. The original article is in German. Various English-speaking authors have written their interpretations of Haas, which you can find in any decent textbook on audio recording techniques.

2 Even if unnatural, it can be interesting, nevertheless. Listen to 1960's-70's era rock recordings from the Beatles, Beach Boys, Lovin' Spoonful, The Supremes, Tommy James and the Shondells, and many more, where mono instruments or vocals are panned to one side, and often their reverb return completely to the other side.

3 When adding Haas delays, listen closely in mono, because improper delay ratios can cause comb filtering in mono. A small degradation in mono may be tolerable if the improvement is significant in stereo. Early reflections, due to their more complex nature, are more compatible with mono folddowns than simple Haas delays.

4 Burroughs, Lou (1974), *Microphones: Design and Application*, Sagamore Publishing Company. (Out of Print). Burroughs quantified the effects of acoustic phase cancellation (comb filtering, interference) with real microphones and real rooms, and devised this rule: The distance between microphones should be three times the distance between each microphone and the source of the sound to which it is being applied. This is particularly important to avoid comb-filtering when both microphones are feeding a single channel; when the microphones are feeding different channels (e.g. stereo), the degradation will be much less noticeable in stereo but still be a problem in mono.