

Mixing: Techniques and Philosophy

Mixing is frequently the first thing we associate with the recording studio. It is how the sounds recorded are assembled into a coherent and artistic combination that often exceeds the sum of its parts. Much like chefs, when given the same ingredients come up with different dishes; mixers have personal approaches that produce different mixes from the same basic tracks. While there is undeniably an artistic component to mixing, there are scientific explanations about why some mixing techniques work so well. We must consider our auditory system performance if we wish to discover the fundamental considerations when combining numerous single tracks into a stereo mix.

Due to the operating mechanism of the cochlea, sounds are divided into frequency responsive regions known as critical bands. These frequency bands generate a single output from all the sound components that fall within the range of each critical band. Within each band, sounds combine to excite the basilar membrane but only the louder components in each band are perceived. Softer elements are hidden, or masked, by the louder sounds. When we hear combinations of sounds through the cochlea, some constituents of the combined sound are eliminated from our awareness by the masking phenomenon. Since the relative levels of the components vary with time, some elements may appear and disappear as the other components change. This leads to a fundamental problem when mixing many individual tracks of sounds with similar frequency content.

Often overlooked, the primary way to reduce the competition for awareness lies in the arrangement of the song. There is a tendency to record as many tracks as we can and then to try to “fix it in the mix.” By recording only the minimum necessary to convey the song, mixing is greatly simplified. This requires pre-recording thought about what actually needs be included in the mix, a potentially difficult task and a job for an arranger. Since this is frequently not how recordings are accomplished, mixers may not always have the easy job of just setting levels and panning for a well-defined project.

If there are competing tracks that must be included, there are several methods of separating individual sound elements; separation in space (panning), separation in frequency content (equalization) and alteration of dynamic range to reduce the effects of masking (compression/limiting.) Even with this rather limited palette, great mixes are possible. The key is knowing how these techniques behave in combination. A complete understanding takes years of experience, but exploring how each possibility might change the mix is a good place to start.

The simplest method of separating sounds in a stereo mix is by panning them to different apparent locations between left and right. Differences in relative amplitude and time-of-arrival are powerful cues about the location of the “phantom image” created by the balance of sounds delivered by the left and right speakers. (The apparent location of these sounds is called a phantom image since there is no real sound source at that location.) Delaying an identical signal in one speaker versus the other moves the phantom image strongly towards the non-delayed speaker. Most mixing boards only provide amplitude panning, however, adjusting only the level to the left and right speakers. This does create a sense of placement, but it is not the complete picture we get in nature where time-of-arrival information coincides with the amplitude differences. Relative amplitude is the weaker of the two cues, nonetheless level panning does allow perceived separation of competing sound sources.

Some older mixing consoles had the ability to place sounds in only three places – left, right or center. An argument can be made that using just these three place assignments leads to an uncluttered mix with an enhanced sense of clarity. There is a competing tendency to try placing each element in a slightly different place in the panorama, creating a rich sound field of many phantom sources. In order for this to be convincing, the playback system and the listening room must provide unambiguous imaging of virtual sources. While this is often pos-

sible in a recording studio, it is far less common in the listening environments of most average music listeners. Good imaging is perhaps more easily provided by headphones or in-ear drivers, with the *caveat* that the image seems to appear inside the head rather than out in space. If our mix is to serve the widest range of listeners, a simpler approach to panning is recommended. Because time-of-arrival cues are absent from level-panned sources, the less ambiguity in placement, the better.

Another powerful technique used to separate sound sources in a mix is frequency balancing through equalization. Many sounds, especially musical instruments, have similar frequency ranges. These sounds are differentiated by the sound of the onset transient, but often the sustained sounds contain very similar frequency components. One way of artificially separating these sounds is to boost some frequencies and reduce others in a complementary fashion. This creates an audible difference that helps to separate the individual instrument contributions. This technique can involve high-pass and low-pass filtering or it can employ peaking filters that boost a range in one source and cut the same range in the other. When we look at a visual representation of the filters, we tend to make significantly visible peaks and notches but most often very slight boost and cut create a more natural-sounding separation. It is better to do this by ear than by eye. That is universally true of equalization for it is our auditory perception that counts, not our visual system.

There are choices in the type of equalizer, from simple high- and low-pass filters to 1/3-octave graphic equalizers. Each has its uses, but in general, the simplest approach that works is best. Digital equalizers are often used now and they are quite flexible. Usually, each frequency band can be turned on or off individually, allowing close tailoring of the filtering without using unnecessary stages of processing. Many times, equalization is used to correct tonal flaws in the original recording, but that is a separate issue with its own demands. When trying to differentiate sounds in a mix, gentler filters are usually sufficient. Unmasking often involves only a few decibels of change to bring out a hidden sound. A one dB boost on one track and a 1 dB cut in another is often effective in creating audible separation of similar sounds. Another consideration in equalization is the bandwidth of the filter, how wide a range of frequencies will be affected. There is a tendency to make narrow but high-amplitude changes, especially when we're looking at the filter shapes in digital equalizer representation. While these do sometimes work, it is generally less obvious to do wider bands with less boost or cut.

A simple trick using a parametric filter can sometimes help find an appropriate combination of boost/cut on two competing tracks. With a fairly narrow Q peaking filter, boost 6 dB or so and move the center frequency up and down through the range of frequencies that are conflicting. Using this technique on both tracks, find the regions that both tracks have maximal energies in common. A small cut on one track with a complementary boost on the other should then increase the distinction between the tracks. This may be repeated to find the next area of overlap and to similarly treat those frequencies. Use the minimum boost and cut necessary, often a few decibels. The technique of moving a narrow peaking filter up and down in center frequency is also a good way of acquainting yourself with the sound of different frequencies in a track. Soon you will be able to identify particular problem frequencies quickly.

The third method of differentiating tracks from each other is to modify their respective dynamic ranges. Most sound sources produce quite dynamic recorded tracks – the dynamic range of the human voice is about 60 decibels, certainly more than can comfortably be reproduced in a normal listening room. Other instrumental sound sources produce comparable dynamic ranges. When combined, dynamic sounds can mask one another in complex and unpredictable ways. By reducing the dynamic range carefully, some of the undesirable interactions can be reduced.

Compression and limiting are conceptually simple techniques but they are not so simple in practice. It is easy to squash a track but it takes some experience to reduce the dynamic range just enough to eliminate a certain unwanted masking without entirely changing the character of the sounds. Gain reduction depends on a variable-gain amplifier and a circuit that estimates the loudness of the signal to produce a control signal than is used

to determine the amplifier gain dynamically. Both the amount of gain reduction and the rate at which the gain changes affect the sound we produce. The compression ratio and the threshold determine the amount of gain reduction. The threshold sets the amplitude above which gain is reduced and the ratio sets the amount of gain reduction. Ratio refers to the output level increase for a given input increase. Low ratios produce less gain reduction while very high ratios allow essentially no output gain increase and are known as limiters.

In addition to the amount of gain reduction, the speed at which gain is allowed to change affects the sound of the process. Most instruments produce an onset transient and a longer sustained sound that persists until the note is over. The initial onset of a sound conveys a lot of the character of the sound source. If the compressor attack is short, the transient can be reduced significantly, changing the character of the sound. Sometimes this is desirable but more often it dulls the note onset noticeably. An attack of several tens of milliseconds leaves the onset intact and begins to compress the continuous sound once the transient is over. When a note is finished, the gain returns to its original value with a time course. A short release time causes the gain to return quickly and can lead to a pumping effect as gain changes audibly with each note. Longer release times make the return to the original gain less noticeable.

The auditory system has a built-in form of compression that protects the ear from overly loud sounds, the acoustic or stapedius reflex. The middle ear produces this effect by tightening tiny muscles that pull the stapes bone away from the oval window of the cochlea, reducing the efficiency of the sound transfer. This reflex has attack and release times of about 40 and 135 milliseconds, respectively. These can be good starting points to create natural sounding compression.

Compressor vs. Limiter

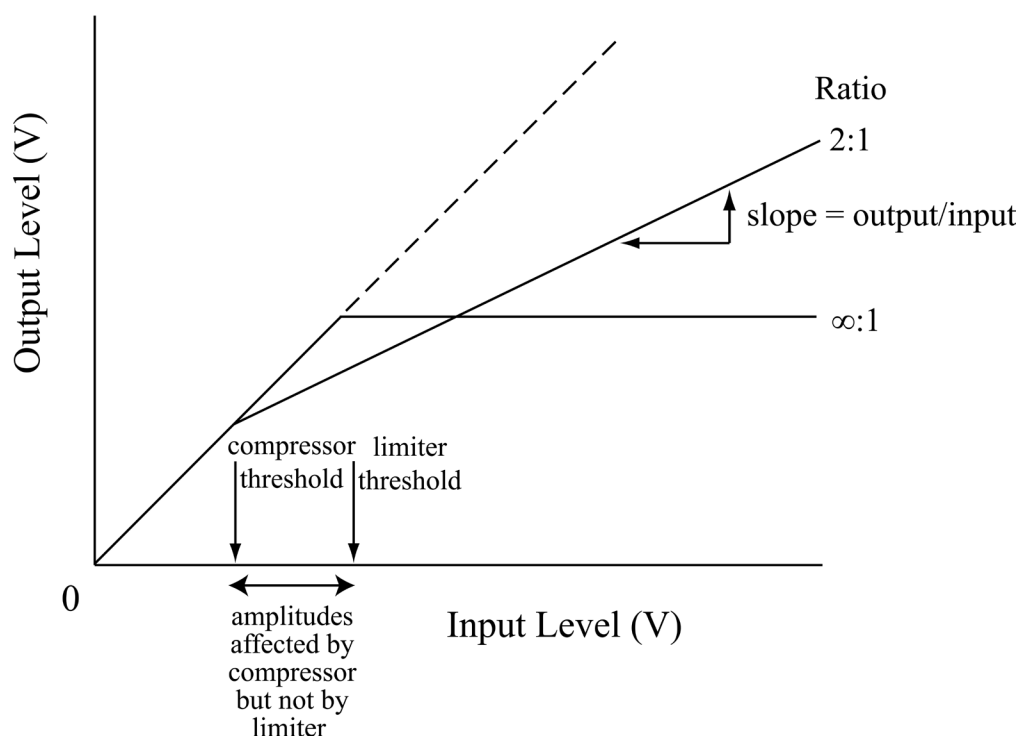


Figure 1: Compressor and limiter cover different input amplitude range for comparable dynamic range reduction.

The question of compression versus limiting can cause confusion. Both reduce the overall dynamic range but they do it in audibly different ways. In order to produce the same overall reduction in dynamic range, a limiter would use a higher threshold than a compressor (Figure 1). The compressor would have a lower threshold and

lower ratio. This means that a wider range of amplitudes would be affected since the lower threshold would cause more compression of lower-level components than the limiter. The limiter would not affect the amplitude balance of these components other than to make them louder *en masse*. While this might seem like a small distinction, it produces audibly different results. When we just want the overall sound to be louder, the limiter accomplishes that by trimming the signal peaks and raising the level of the rest of the signal. A compressor alters the amplitude balances over a wider range of levels and changes the overall sound more noticeably than a limiter. In short, compressors change the amplitude balance of a signal more than limiters, which mainly affect the loudest peaks and leave the rest of the signal unaffected.

In the age of digital audio, there is an additional way to help separate tracks – time-based effects. Delaying a signal and adding it back to the un-delayed signal generates the sound of comb filtering we strive to avoid with multi-microphone recordings. By varying the delay, a sense of coloration and movement can be added to a sound source. Digital reverberation allows a sound to be given a sense of depth or distance from the listener in the sound field. These techniques give us a way of discriminating between dry and effected sounds that can help to combine sounds while maintaining a sense of the individual track contribution to the mix.

As we know, delaying a track on one side of a mix relative to its opposite side shifts the image to the earlier side strongly. While this is not provided by the usual pan potentiometer, it can be used to increase the perceived left/right separation. Delay can also be used to “fatten” a sound by adding delayed versions to both left and right sides of a mix. Delays of less than 35 milliseconds are not perceived as discrete echoes but can be used to make a track seem larger, particularly for centered solo instruments. For example, a 15-millisecond delay added to the left and a 25-millisecond delay added to the right cause a centered image track to sound more prominent in the mix even though the overall volume is not significantly increased. Straight delay can also be used, as it is often on vocals. To keep the repeats in sync with the tempo, delay times can be calculated by dividing 60,000 by the tempo in beats/minute to determine the delay equivalent to a quarter note. This can be increased or decreased to eighth note, triplets, etc. The amount of feedback, or recirculation, determines how many repeats are heard.

Chorus and flange effects generate a similar increase in prominence to tracks for much the same reason. We tend to perceive changes or movement quite strongly and these effects generate a sense of change that makes a sound jump out at the listener. We do need to be careful not to go too far with delay effects, as collapsing the mix to mono can reveal unwanted excesses that do not sound good. The mono button is a good thing to use – check mono compatibility frequently when applying time-based effects.

By using various combinations of these techniques, mixes that have good sound separation can be created. There’s no one way and experimentation over time will increase your effectiveness and the speed with which you can deliver a great-sounding mix.