

Convolution Techniques

Barry Truax, Simon Fraser University

Suggested Setup:

- use the free download of SoundHack (soundhack.com freeware)
- open, play and edit the results in a sound editor that provides: standard edit functions, gain control, fade-in and out, pitch transposition, reversal, single channel selection, standard EQ/Filter, add/mix (ideally the editor goes through a high-quality soundcard, not the internal soundcard of your computer/laptop); note: you must delete the file in SH before you can save the edited version and SH files must have an .aif or .wav extension to be opened in Audition; SH itself will only process .aif files except for Ring Modulation

SoundHack Convolution options: (Command C)

- **Normalize:** always use this
- **Brighten:** usually use this unless your sound is particularly strong in the high frequencies (since most acoustic spectra fall off approx. 6-12 dB/oct, the highs will normally become very dull unless you use this option). If the result is too bright, you can always EQ it in the editor.
- **Gain:** 42 dB is the default, fine for impulse reverb, etc, but for auto-convolution (AC) you'll likely have a huge dynamic range, so use 0 dB; if the output is still clipping, then reduce the overall gain of your signal in the editor, and re-do the convolution
- **Window:** this is for the analysis phase, and usually doesn't make a strong difference in the output; EXCEPT that convolving files A*B will be different from convolving B*A when you don't use the rectangular window! The output will sound similar, but the overall amplitude envelope will be different.
- **Ring Modulate:** this is an alternative to Convolution where the signals are multiplied, not the spectra, but note that the two files need to have similar lengths which is not the case with Convolution
- **Moving:** the analysis moves through the impulse file, instead of using all of it (useful for controlling the density of the output in auto- and hybrid convolution)
- **Pick Impulse:** used to select the impulse file; note that with auto-convolution you need to make a copy of your original file (command D) beforehand, as SoundHack needs to work with two distinct files
- **Process:** you then name the output file and select its format; SH offers to name the file as the compound of the two input files with their extension, separated by the * symbol for convolution; however, this can get very long, so I tend to shorten this to something like:

bell*reverb bell*self bell*self*self bell*rain

Note that with auto-convolution and other high dynamic range sounds, you will need to save the file in **24-bit format**, because in only 16-bit, you'll hear strong quantizing noise (sounding like a square wave) in the low amplitude parts.

Convolution Techniques:

(A) **Impulse Reverb:**

This is the classic case where you convolve a dry sound with the Impulse Response (IR) of a space. The result is that the dry input sound appears to be located in that space, at the same distance as the original impulse was from the microphone, coloured by the frequency response of the space, with the characteristic, complex reverb tail. Note that the output file duration is the sum of the durations of the source plus IR, as you would expect with reverb.

If you convolve the dry sound with the IR twice, its apparent distance doubles.

Given that IR recordings are generally done in an empty space with no foreground sounds (e.g. audience), the result of the Impulse Reverb convolution will generally be more reverberant than you may want (or what would happen with an audience present), so it is often good to mix the wet and dry sound together in your mix, which gives the added advantage of changing their amplitude ratio, such that the source appears to move closer and farther away.

Note that if you add Impulse Reverb in some editors, it's up to you to add silence at the end of your source file to allow for the reverb to extend properly and not get cut off (but this is not required in SH).

(B) **Auto-Convolution (AC):**

This is the term for convolving a sound with itself, and in a certain sense, it's the "worst case" scenario in terms of keeping control of the result because it will produce the strongest matching of frequencies between the two files because they are identical. So, the strongest frequencies in your file get VERY strong, and the weakest frequencies essentially drop out, and the dynamic range of the sound increases substantially.

As already stated, use 0 dB gain in SH, but if you know that your input signal is already peaking around the maximum, and has strong frequency components, reduce its level by at least 6 dB before starting (or you'll have to go back to do it later anyway).

There are two distinct types of output with AC depending on your source material:

(1) A single event, or smooth continuous texture: the AC output will double in length and so the sound appears to be stretched. Given the multiplication of the frequency spectrum, the effect will be to filter out low level noise or weak spectral components. Also, any sharp attack in your source will be smoothed out dramatically, even if the brighten option is used. Applying AC again to the result will again double the length and further filter the spectrum to its principal components and smooth the attack even more. Convolvering the AC file back with the original will make it 3 times the original length, and so on. Be sure to remove excess silence from the output file before re-convolvering it, as that will double as well, and it often is prudent to fade out a long decay a bit early.

(2) When a file has distinct and separate events (e.g. speech, or a rhythmic motif), it will not result in what you'd expect as a simple stretching of the sound. Rather, each component sound is convolved with every other component sound and mixed together. At the beginning you'll hear the first component convolved with the first component, and at the end it will be the last one with the last one, but in between there's an accumulation of all other combinations, generally peaking in the middle. The explanation of this effect may be found in the technique of "direct convolution" where every sample of the first file is multiplied by every sample of the second file and added together (a computationally huge

task, hence the efficiency of SH analysing the sound first and simply multiplying the resulting spectrum amplitude values). Or, think of each individual reflection in an acoustic space being added together to produce a smooth reverberation. If the resulting sound density is too great, use a moving window so that shorter sections of the impulse are used.

(C) **Convolving Unique Files (Hybrid Convolution):**

With non-identical files, results will vary widely, so a few considerations are in order:

- two long files will strain the memory limits, but more importantly, will probably produce just an undifferentiated wash of ambient material, so unless you want that, it's best to avoid the situation
- with two continuous textures, a better result is to use a short bit of one with a longer section of the other
- sounds that have a lot of component micro-impulses work remarkably well to preserve definition in the output (e.g. all the sounds used in *Chalice Well*)
- convolving wet and dry sounds produces a natural continuum of spatial presence: dry with dry remain foreground; dry with wet are medium distance; wet with wet are very distant
- convolving $A*B$ will be different from $B*A$ when the analysis window is **not rectangular** (i.e. no envelope); this is hard to explain, but the difference will show up in the overall amplitude envelope of the result. These closely related pairs are often useful in multi-channel compositions on separate tracks because they are uncorrelated and give an expanded sense of environment
- to bring out a specific frequency region of a convolved sound, convolve it again with a sound that's strong in that region rather than applying simple EQ
- when both sound files have rich spectra, the result will be much too broadband and undifferentiated; use the **moving window** option to reduce the complexity, where the shorter the window, the less the overall density in the output, as only a fraction of the impulse file will be used (note: this can be shortened down to the micro range of 100 ms or less, as in *Rainforest Raven*); when using a moving window, the resulting duration is that of the source file plus the duration of the window

(D) **Convolving Stereo Files:**

Stereo files may have dissimilar amplitudes between their left and right channels; these differences get exaggerated in convolution, particularly with AC. Therefore, when cleaning up the resulting file, look at the relative amplitudes in each channel (which can easily throw the location of the sound into either the left or right speaker). Select the weaker channel, find or estimate its peak amplitude and add appropriate gain to that channel (hence the importance of having an editor that allows you to do that on a single channel!).

Alternatively, lower the amplitude of the stronger channel. Remember convolution is a **nonlinear** process where amplitudes are multiplied.

SoundHack has a small but very annoying bug: in the normalization process, a single sample at the peak level may overflow the limit and go maximum negative with a nasty click. However, in almost all cases, if you zoom in to the sample level, you can edit this out seamlessly, but good luck with that! Actually, it's a useful micro-editing technique to develop some skill at using. However, in many cases, simply changing the analysis window will avoid the problem.

Do not expect that a specific stereo image in your source file will remain in the convolved output, at least in terms of left-right placement. However, the sense of depth and reverberance will remain as described above. On the other hand, a relatively uncorrelated stereo file (i.e. without panning) will remain uncorrelated and produce good point-source definition in each speaker.

(E) Things to do with the Output File:

All of the files created in cases B & C may be considered hybrids of your original and so with extended work, it is typical to create a large number of such files. Therefore, an efficient naming convention is useful, and for a larger work, I also create a text file where I make notes on the quality and duration of each file I generate, for future reference during the assembly stage. Families of related files start emerging and these are useful, for instance, in positioning in different channels in a multi-channel mix, particularly if you are trying to create an acoustically coherent soundscape.

As in classical sound object work, a “tree structure” of files may start to emerge, where the output of each convolution leads to further processing of that file, including convolution with itself and/or other files. For instance, pitch transposition (with or without duration extension) may be useful. Time stretching using a granulation program may also be interesting, particularly since a broad-band original sound may have been simplified in its spectrum through AC; these sounds will create less masking of others. I also like to use a resonator at various points in file generation, in order to bring out specific frequency regions in the sound, usually prior to time stretching, as the resonator itself may create some interesting internal dynamics that will become more evident in the stretched version.

Because of the smearing effect of transients, reversing the convolved sound is often aurally interesting and will lack the abrupt decay that happens at the end of a reversed sound with a strong attack. The reversed sound may also mix well with the original sound, a specific case being where the source sound has a short attack and long decay. Its AC version will do the same, but the decay will be even longer. Mixing this sound with its reversed version produces a symmetrical structure that can be repeated or even looped with good effect. In some editors you can try this out with a minimum of operations:

- select all of the file; reverse it and copy to the clipboard; undo the reversal and add the original sound to the clipboard. Note that with a sound that is strong at the start and decays to nothing, you can “add” the two versions together with little fear of clipping, whereas with more continuous sounds you’ll need to “mix” them together with a 50% contribution from each
- a variation on this mixing of forwards and reverse is at a certain point in the process to swap channels so that whatever starts in the right channel ends up in the left and vice versa, creating a nice timbral-spatial movement; the more dissimilar the left and right channels are, the more interesting this change
- multi-tracking several of the related family members in these compound processes, particularly if given separate spatial placements or movements is a good way to create a more complex (or “abstracted”) version of your materials

For information on the theory of convolution, see: www.sfu.ca/~truax/conv.html or the Tutorial on the WSP Database (obtain a guest password at truax@sfu.ca).