

How to use the DC Live/Forensics Dynamic Spectral Subtraction™ (DSS™) Filter

Overview

The new DSS feature in the DC Live/Forensics software is a unique and powerful tool capable of recovering speech from recordings containing loud music or other coherent noise. Until now, a recording that was covered or masked by loud music was basically a lost cause. DSS decoding is designed to make it possible to attenuate this music and uncover the speech.

Basic Approach of the DSS Filter

DSS works by performing a continuous and intelligent subtraction of one audio signal from another. Normally, with a forensic recording containing speech masked by loud audio, you will require a reference recording containing just the audio that needs to be removed. The audio track containing the music or other audio to be removed is called the “Reference Track.”

The DSS Controls

The following four controls are active when operating in DSS mode. All of the other CNF controls are disabled:

1. CNF Mode (Selection Box):
 - A. Mono DSS – Delay Reference
 - B. Binaural DSS – Left Reference
 - C. Binaural DSS – Right Reference
2. Attenuation: Range = 0 to 100 (Tune for a Null in the Noise) Null usually occurs around 50.
3. Delay: Range = 1 to 10 FFT's (Larger distances between microphones need higher settings of this parameter). In most instances, a setting of 1 is effective.
4. Overlap: Range = 10 % to 50 % (Tune for minimum digital artifacts). In most instances a setting of around 40 % is effective.

Recording a Reference Track

There are many ways that you can obtain a reference track. These examples should make this clear:

Real Time Methods*

1. Place two microphones in the venue. Place one near the target conversation and place the other near the source of the background audio source such as a TV, stereo system, jukebox, or a live band. Record these two signals with a stereo tape recorder or computer.
2. Wire the investigator with two microphones. Place one near the investigator's chest and place the other much lower on the investigators body, like down in his or her sock or shoe. Record both signals with a miniature stereo tape recorder.
3. Wire the room with a wireless microphone located near the sound source like the TV, Stereo, jukebox or live band. Wire the investigator with a wireless microphone located near his or her chest. Record both signals with a remote stereo recorder or computer.

Non Real Time Methods* ‡

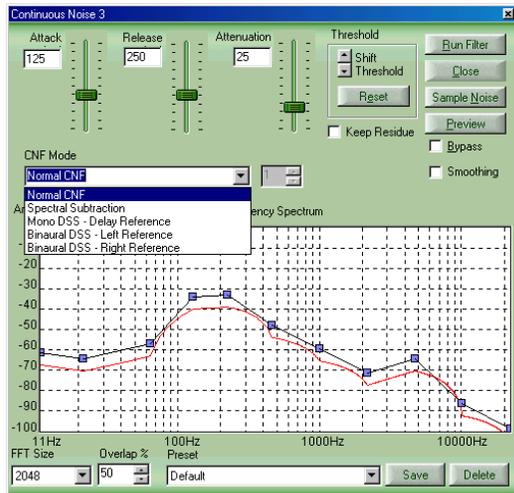
1. Assume that you have a recording made in a bar or similar venue that was recorded with a monophonic pocket tape recorder. The juke box or other interfering music source is covering over the targeted speech. You can go back later with the same recorder and record the same exact song that was being played. This will become your reference track for DSS decoding.
2. You have the same situation as stated above, but you have a second tape recorder on site which is recording only the noisy background environment.

3. You have the same situation as stated above, but you record the same music that had been playing at the venue from a commercial audio CD. This process can be performed in non real time back in your audio lab.

* Note 1: Digital Recorders produce better results than Analog recorders in DSS decoding applications.

‡ Note 2: If the interfering source of audio was a radio or television, many broadcast stations maintain an archive of “airchecks.” You may be able to access the required broadcast “aircheck” recording through either the use of diplomacy or a court order.

Obtaining a reference recording is an important step in removing loud coherent noise sources such as music. Using the Real Time Methods described above, technique number 1 (there under) will produce the best results. In the Non Real Time methods described above, number 2 (there under) will produce the best results since it will rely on a reference recording that closely resembles the noises that you will be attempting to remove from the target signal.



Location of DSS in DC Live/Forensics

There are 3 DSS modes of operation available in the product. To select one, drop the CNF Mode box as shown above. As you see, you can select either the right or the left channel as the reference track.

If you have no reference track recording and cannot re-create one, you can try to use the setting called Mono DSS Delay Reference. This will attempt to attenuate the noise by comparing the audio at an instantaneous point in time and comparing it with a point at some other time before or after the comparison point. This has the effect of allowing the program to create its own reference signal. Note: This method is inferior in comparison to any method utilizing a true reference track.

Creating a Stereo track from two discrete tracks in non real time situations: *

The audio file that you will actually clean up using DSS decoding is ideally going to be a stereo file that you recorded in real time. However, often that is inconvenient and non real time methods must be used. In these cases, one channel of the file will be the recording with the speech you want to recover (the forensic

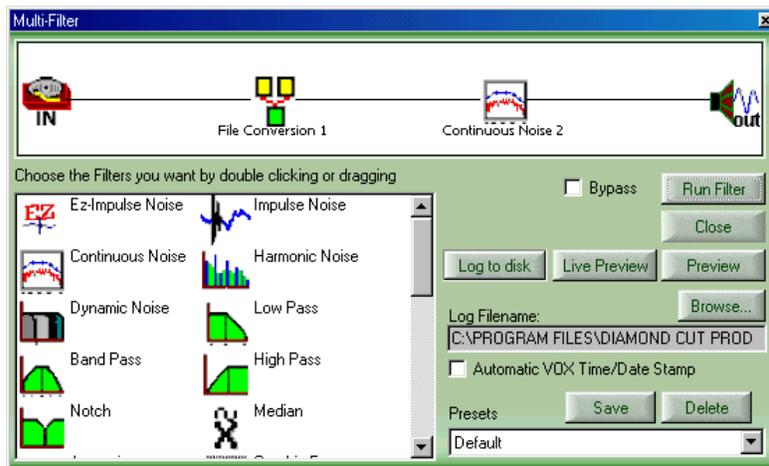
recording) and the other channel will contain just the music or other non-random audio. These two recordings will have to be combined into a single stereo (binaural) recording. The easiest way to accomplish this is to use the File Split and Re-Combine function under the Edit menu. Here's the procedure:

1. Take the two recordings (the forensic recording and the reference recording) and convert both of them into monophonic files if necessary by using the File Converter Filter.
2. Use the File Split and Re-combine feature to merge these two mono files into a stereo (binaural) file
3. Time align these two files by either cutting a piece from the beginning of one of them or insert a piece of silence of appropriate length in front of one of them. Note: Using the Markers and the Time Display feature is quite helpful to precisely measure the time displacement between tracks to calculate how much audio must either be cut or inserted to result in the proper time alignment. The two tracks should be time aligned to within +/- 25 milliseconds of each other for optimum results.

* Note: If the interfering audio came from a live performance, having the live performance re-created by the talent after the fact will not produce a useable reference track for DSS decoding

Using DSS with the Multifilter:

You can use DSS in the Multifilter similar to any other Diamond Cut tool. However, you could also use the File Conversion filter in front of the Continuous Noise Filter as seen here:



The File Conversions filter has a time offset function that can then be used to perfectly “dial in” any remaining time offset so that both tracks are perfectly in sync with each other. This can be done while you listen to the preview and adjust for best noise reduction. Of course, you can also add other filters before or after the ones shown above, as required.

DSS Adjustments:

The controls that are active in the DSS filter are Attenuation, FFT Size and Delay. The Attenuation setting will control the amount of noise reduction that is performed by the DSS filter. You can think of this control as being analogous to balancing the weight(s) on a balance scale. Moving it up will reduce the noise more until you pass through a “null” point in the background music. You need to tune the attenuation control for the most music reduction, which generally will occur around an Attenuator setting of 50, as long as both discrete channels are relatively balanced in amplitude with respect to one another.

The FFT size controls the size of the frequency “buckets” that are being used internally by the filter. Smaller numbers allows for more “self adjustment” of the filter to the mismatched forensic and reference recordings. Larger values produce better frequency discrimination and overall attenuation. We find that settings of 1024 or 2048 generally produce good overall results, but smaller or larger settings should be tried as well.

The Overlap control is generally set to 50%. However, it is worthwhile experimenting with other values of Overlap in order to minimize the introduction of digital artifacts into the final resultant signal.

The Delay control can also help with time alignment mismatched audio channels. A setting of one is typically the best for FFT sizes of 2048 or more and higher settings may give better results with FFT settings of 1024 or less. You can calculate the actual delay time in milliseconds by applying the following formulae:

$$TD = (\text{Delay Setting} - 1) (\text{FFT Size}) (\text{Overlap} \times 10) / \text{Sampling Rate}$$

Wherein:

Delay Setting is an Integer value from 1 to 10

&

Overlap is a value from 10% to 50% (note: in the formula above enter this as a value between 10 and 50)

&

Sampling Rate is a value given in Hertz

&

TD is the resultant delay time given in milliseconds

As always, simply preview the audio and make your adjustments in the DSS filter window and also with the Time Offset slider in the File Conversion filter.

Primary Compensation Issue with the DSS Filter in Real Time Situations

The primary problem encountered when using the DSS filter in real time applications arises from the distance between the two microphones used to make the binaural recording. Because a physical distance in the venue separated the two microphones, the propagation delay (sometimes referred to as group delay) of the signal between the two microphones in the room may need to be compensated for. Since sound travels at 1131 feet / second at 70 degrees F (or 1.131 feet / millisecond), large distances between microphones can cause misalignment between the two tracks of the binaural file. In real time situations, the noise signal on the Target Track will lag the Reference Track at the rate of 0.884 milliseconds per foot. The File Conversion Filter and its Time Offset control can be used to compensate for up to 20 milliseconds of propagation delay representing a distance between the microphones of up to about 23 feet. If more distance had existed between the two microphones, multiple File Conversion Filters can be cascaded in the Multifilter to increase the total compensation time.

Primary Compensation Issue with the DSS Filter in Non Real Time Situations

The primary difficulty encountered when using the DSS filter in non-real time applications arises from the lack of acoustical matching between the Target Forensics recording and the re-created Reference Track. In other words, room resonance, frequency response or natural room reverb may not exactly match your re-created Reference Track. Performing some pre-processing on your Reference Track can compensate for

these acoustical mis-matches. You can use the 20-band equalizer to match the resonance and frequency response of the room. Also, you can use the Reverb to simulate the acoustical reflection characteristics of the venue. These steps will rely on your own sense of hearing to create the match. When you listen to the music on the Target Forensics recording, focus your listening on its musical content. Then try to create that same sound on your re-created reference track using the above-mentioned Diamond Cut tools. Then use this pre-processed Track as your final Reference Track to be applied to the DSS filter.

Dynamic Spectral Subtraction™ and DSS™ are Trademarks of Diamond Cut Productions, Inc. 2003