Delay Cat User Manual

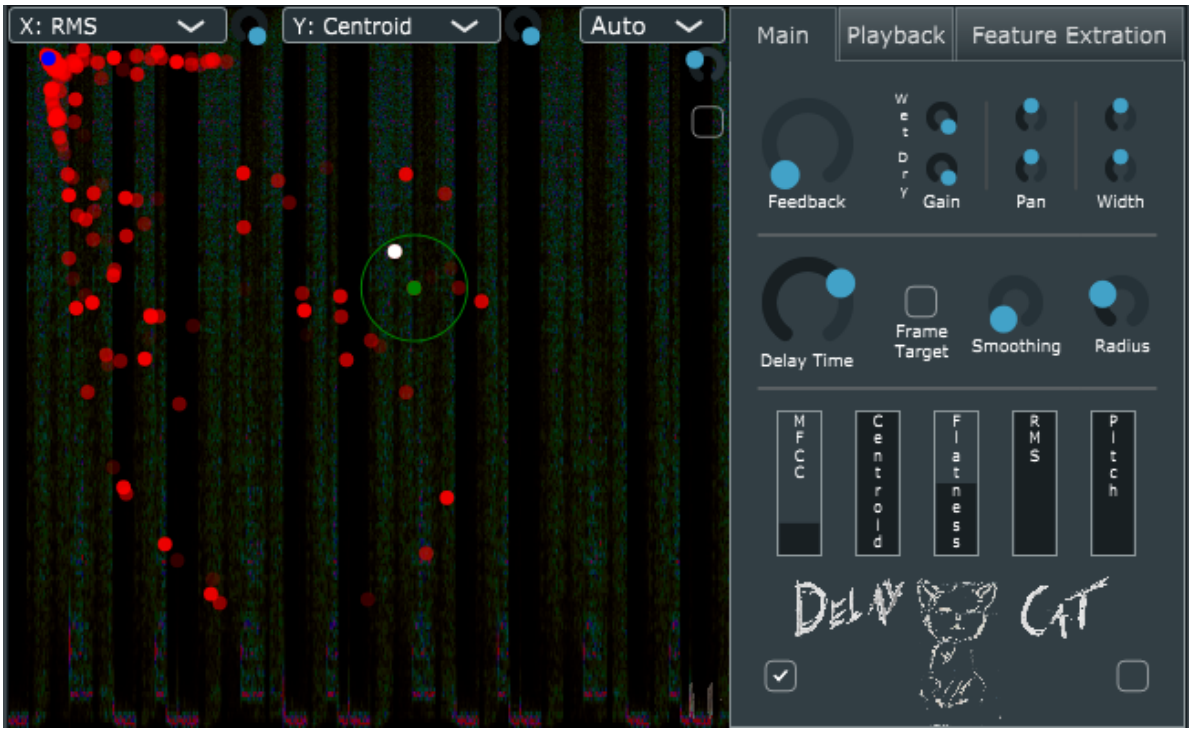


Table of Contents

Table of Contents 1

Introduction: 2

General Tips: 2

Overview: 3

Concatenation View: 4

Main Tab: 5

Playback Tab: 6

Feature-Parameter Mappings: 7

Presets: 8

Feature Extraction / MIR: 9

Introduction:

Delay Cat is a novel delay plugin using a unique Feature Based Delay Line architecture to expand the capabilities of a delay line and power your sound design. The Feature Based Delay Line leverages Music Information Retrieval feature extraction analysis of the input audio to intelligently control the behavior of the delay line and techniques of Concatenative Synthesis to expand the sound design capabilities of the delay line.

Specifically, this means that all the audio in the delay line is segmented and organized according to its characteristics, forming a “feature space” for the contents of the delay line. Segments are then selected and layered according to their position in this feature space, forming the “wet” signal of the delay as they move through the delay line. The position of each segment in the feature space can also be used to control various different functions of the architecture, including a myriad of effects such as pitch and pan.

For a more in depth description of the architecture, there is a research paper on the topic.
dafx23.create.aau.dk/index.php/proceedings/

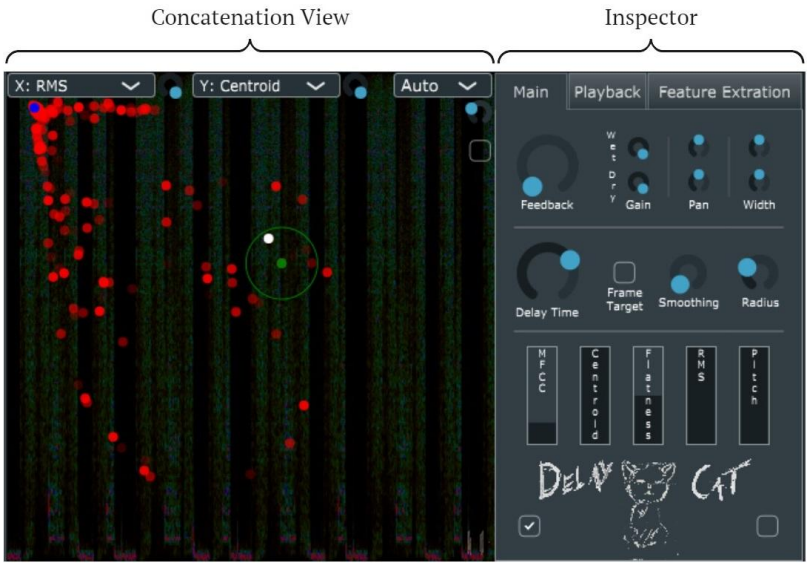
General Tips:

Hover over any parameter to get a short tooltip description of what it does.

Right click on any parameter to view context parameters and graphs (if applicable).

Overview:

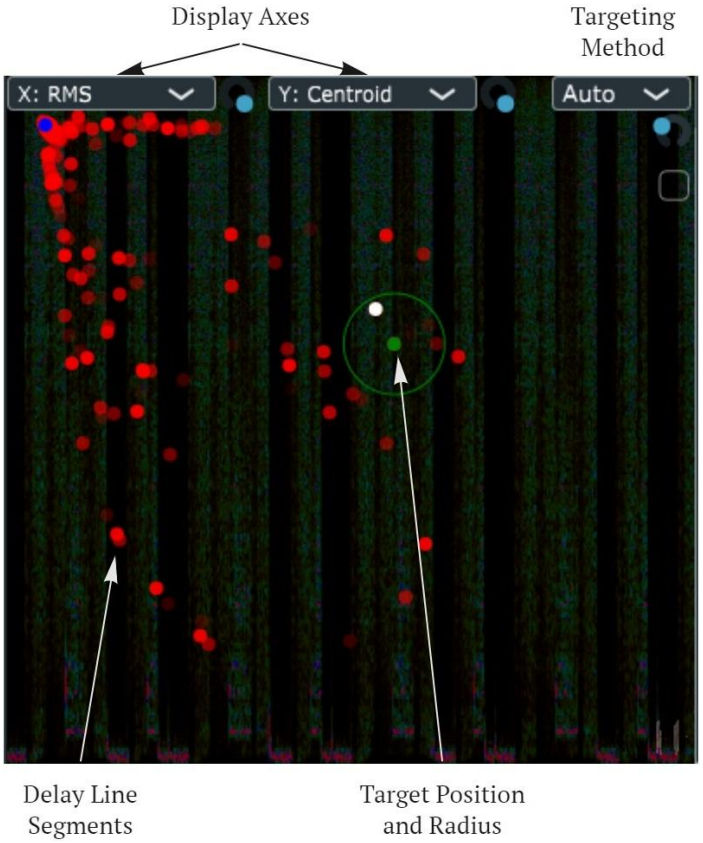
The plugin is broken into two distinct sections: Concatenation View and Inspector.



1) **Concatenation View**: This is the main interactive graphical display for the Feature Based Delay Line, representing the feature space of the delay line.
2) **Inspector**: This is the inspector, in which parameters are found, grouped and organized into tabs.

Concatenation View:

This view represents the current state of the Feature Based Delay Line visually, in a two-dimensional representation of the feature space.



Plotted in this view is the audio inside the delay line, analyzed and segmented as it enters the delay line. Represented is the audio in the delay line (red), the target position and radius (green), and the currently selected segments, which are being sampled for playback (white).

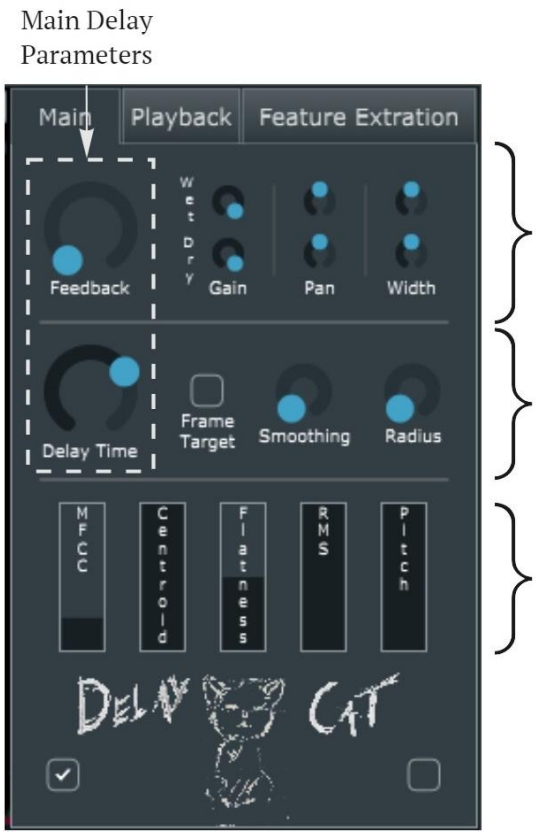
Selecting the display axes allows for different dimensions of the feature space to be plotted on the two-dimensional plane.

The Targeting Method can be set to:

1) **Automatic** (delay based targeting)
2) **Manual** (user input based)

Main Tab:

The Main Tab of the Inspector contains the core parameters of the Feature Based Delay Line (Delay Parameters, Targeting Parameters, Feature Weights) and the Input and Output Parameters.



Main Delay Parameters:
Delay Time: Delay time in the delay line used to determine the target position for segment selection
Feedback: The amount of wet signal fed back into the delay line

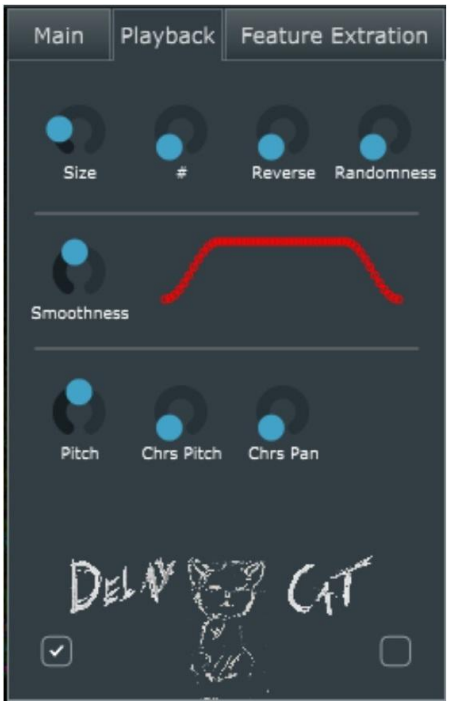
Input / Output Matrix:
This matrix contains the gain, pan, and stereo width for the dry and wet signals.

Targeting / Segment Selection:
Radius: Radius for segment selection
Smoothing: Smoothing factor for target position.
Frame Targeting: Determines if average segment features or individual analysis frames are used for selection.

Feature Weights:
These weights control the relative power of each feature during selection.

Playback Tab:

The Playback Tab of the Inspector contains the segment parameters and audio fx parameters.



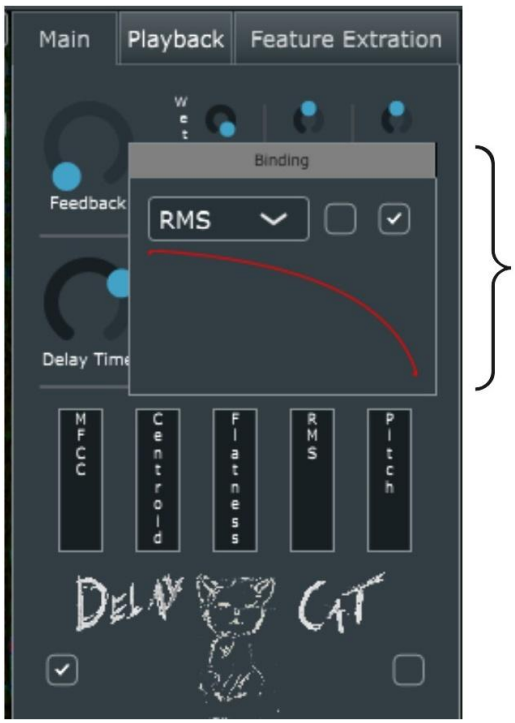
Segments & Reverse / Random:
Segment Size: Size of each segment written into the delay line. Best fit based on onset detection also available (RMB).
Layering: The number of segments read from the delay line simultaneously.
Reverse: The % chance that a segment will be played back in reverse.
Randomness: Amount of randomness.

Smoothness:
The smoothness / window applied to segments played back, plotted (red). Segments can be automatically overlapped during the onset/offset of the windowing (RMB).

Pitch & Chorus:
Pitch: The base pitch offset for segments played back (semitones).
Chorus Pitch Width: The width from the base pitch offset, with which segments played back will be spread within.
Chorus Pan Width: The width from the center pan, with which segments played back will be spread within.

Feature-Parameter Mappings:

Applicable parameters of the feature based delay line can be mapped or bound to the feature extraction analysis in the Feature Based Delay Line with **Graphs**.



Feature-Parameter Binding

Graphs can be brought up by *right clicking* on any applicable feature. This will also display any context settings for the parameter, affecting how it is applied.

The feature set used for mapping can be obtained from either:

1) The average features of the selected segment that the parameters are being applied to. (Specific to each individual segment played back)
2) The Target Position feature set used for segment selection. (Specific to each distinct target position)

Bindings can be used to apply intelligent automation to the parameters, such as feedback control (picture left).

Presets:

Presets store the parameter settings of the plugin state to file. Delay Cat has a handful of internal presets that display some of the functionality of the Feature Based Delay Line to help get started. These presets can be found via program / preset selection from the plugin host (DAW). You can also save and load custom presets to .dc files.

The **Preset Menu** can be found by right clicking the concatenation view, except where there is a segment plotted.

Save To File saves the current parameter preset to a .dc file.

Load To File loads a parameter preset from a .dc file.

Feature Extraction / MIR:

The full set of features included is: MFCC, Spectral Centroid, Spectral Bandwidth, Spectral Rolloff, Spectral Flatness, Spectral Contrast, Spectral Flux, Short Time Energy, Short Time Variance, RMS, Fundamental Frequency Estimation, and Onset / Offset Detection.

The paired down set of features is: MFCC, Spectral Centroid, Spectral Flatness, RMS, and Fundamental Frequency. This feature set is paired down to limit redundancy, while maintaining a wide array of analysis.

MFCC:
A rough proxy for the general timbre of a sound. A breakdown of the FFT frequency analysis into a more select number of Mel frequency bands. Because this feature is a set of frequency bands, all bands are all looked at, and individual bands cannot be mapped to parameters or weighted individually (this is planned).

Spectral Centroid:
A measure of the "brightness" of a sound. This feature measures the center of the frequency spectrum of a sound.

Spectral Bandwidth:
Typically a measure of the "noisiness" of a sound. This feature measures the width of the main area of interest in the frequency spectrum of a sound.

Spectral Rolloff:
A measure of the "cutoff" of a sound, where most of the frequency spectrum is below that cutoff. This feature measures the frequency at which most (usually 85%) of the magnitude of the spectrum is below that point.

Spectral Contrast:
A proxy for the general timbre of a sound. This feature measures the difference between the strongest and weakest frequencies within given sub-bands of the spectrum. Because this feature is a set of frequency bands, all bands are all looked at, and individual bands cannot be mapped to parameters or weighted individually (this is planned).

Spectral Flatness:
A measure of the "noisiness" of a sound. This feature measures how flat the frequency spectrum of a sound is, which corresponds to its "broadband noisiness."

Spectral Flux:

A measure of the “change” of a sound, with peaks typically corresponding to attacks of the sound. This feature measures how much the frequency spectrum of a sound changed from its last frame to the current frame. Used in the calculation of onset / offset.

Short Time Energy:
A measure of “loudness.” This feature measures the sum of all the frequencies in the FFT analysis. Often a rougher estimation than RMS. (This feature may be removed.)

Short Time Variance:
A rough measure of “noisiness.” This feature measures the variance from the mean in the frequency spectrum of a sound. (This feature may be removed.)

RMS:
A measure of “loudness.” This feature measures “root mean squared” of all frequencies in the FFT analysis.

Fundamental Frequency / Pitch Estimation:
This is a measure of the fundamental frequency, or pitch, of a sound. This is a complex computation based on harmonic analysis of the frequency spectrum. (At the moment this feature is sometimes a bit inaccurate.)

Onset / Offset Detection:
This is a measure of when an onset or offset occurs in a stream of audio. This fires when the spectral flux of a signal is measured outside of standard deviation of the last x measurements. The number of standard deviations, as well as how much a signal is used in the ongoing average, corresponding to the “flexibility” of the computation, can be tuned.