Delay Cat User Manual

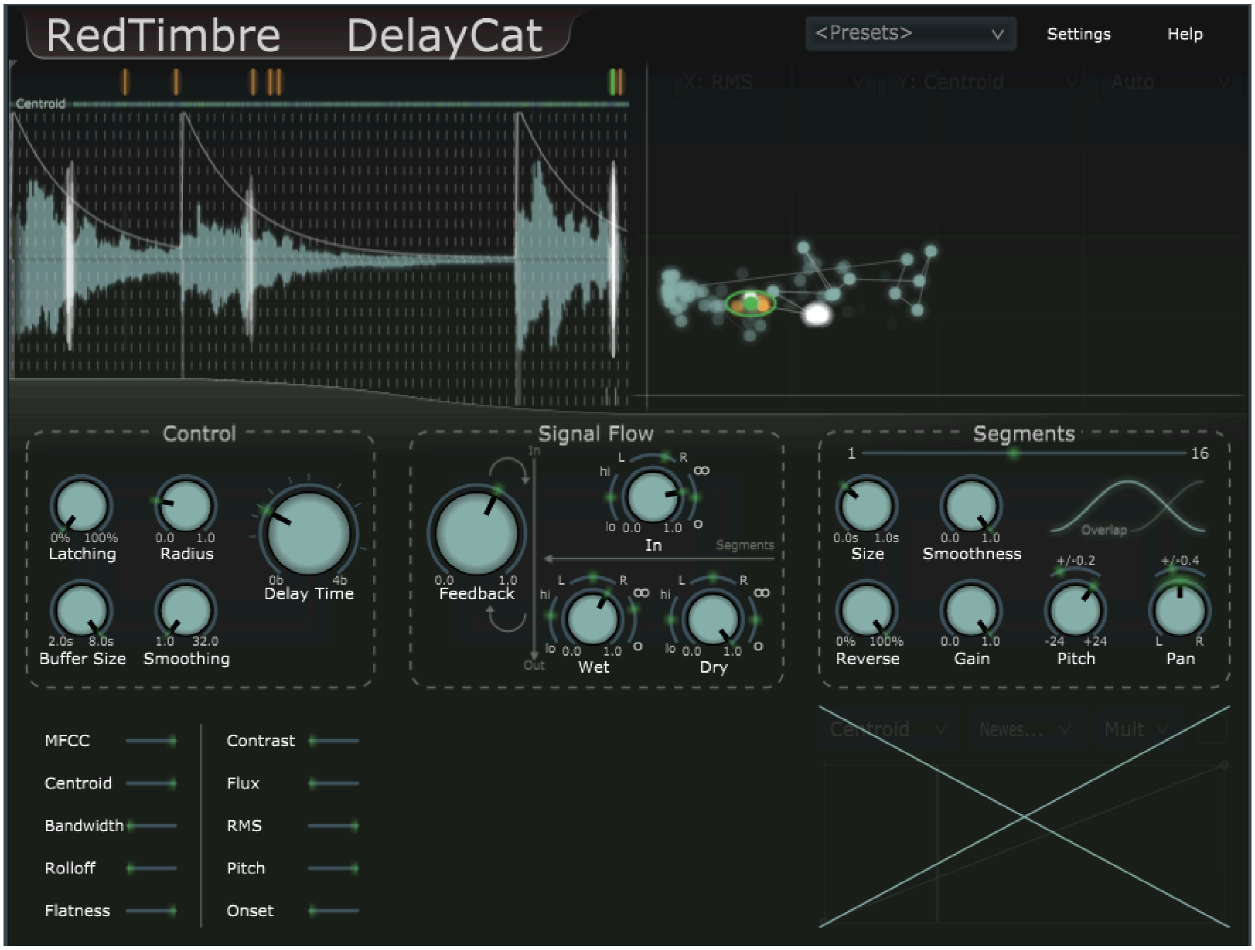


Table of Contents

Introduction:

Delay Cat is a novel delay plugin using a unique Feature Based Delay Line (FBDL) 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 tap the delay line and resample / resynthesize the sound.

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.

Overview:

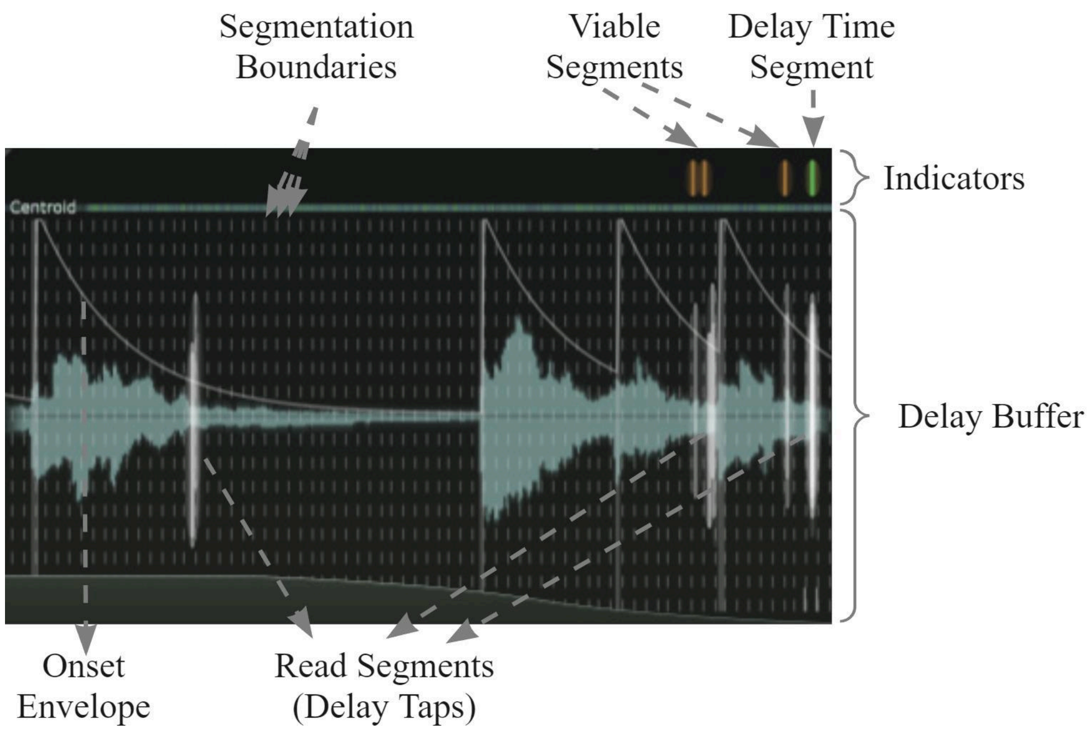
Generally, the plugin can be viewed as a delay line or as a concatenative synthesizer, and should be thought of as both. The architecture is a blend of these two things into a unique system. The main interactive graphics of the plugin represent these two views, providing complementary information to each other, while the rest of the plugin houses parameters controlling the expression of this system.

Specifically, the plugin is broken into distinct sections: Header, Plots / Graphics (Delay Line View, Feature Space / Concatenative Synthesis View), Parameters (Control Parameters, Signal Flow Parameters, Segment Parameters), Feature Weights, and Parameter Context / Graph.

Each of these will have its own proceeding section.

Delay Line View:

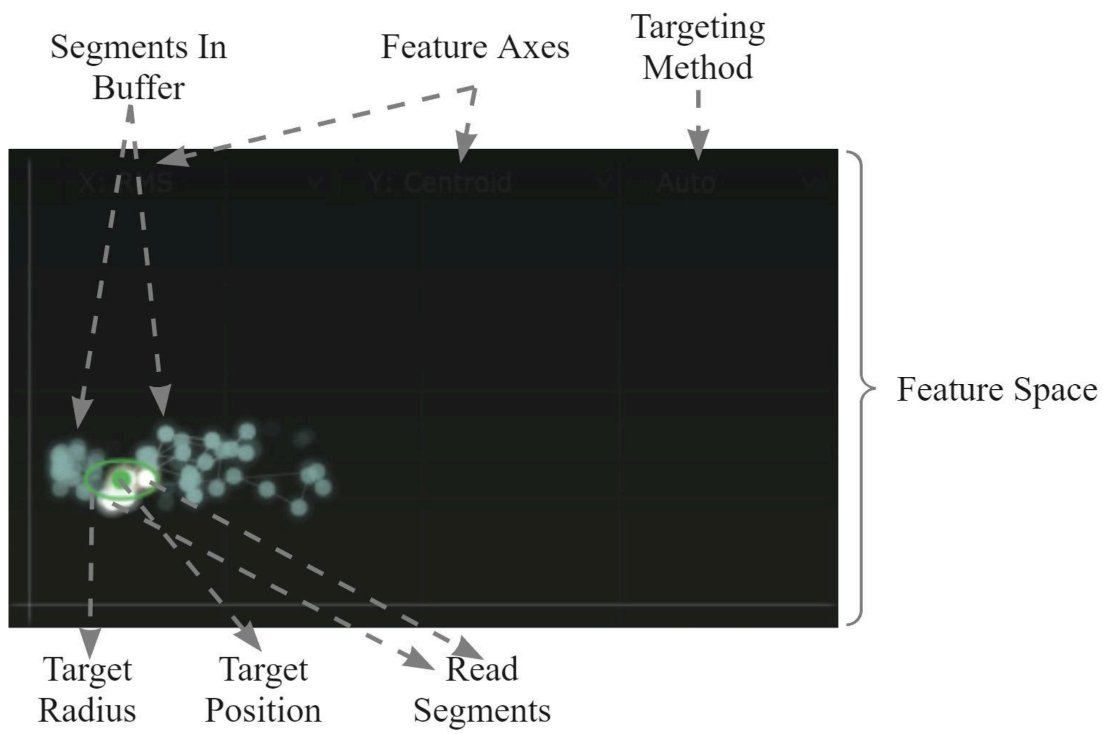
This view represents the current state of the Feature Based Delay Line buffer visually, as well as data related to the contained audio and operations processing the contained audio, such as segmentation boundaries, read segments (delay taps) and more.



1. **Segmentation Boundaries**: represent the segment boundaries determined by the *Segment Size* parameter.
2. **Onset Envelope**: this envelope, triggered by onset detection and optionally plotted over the audio, can affect many operations in the FBDL.
3. **Delay Time Indicator**: this indicator represents delay position determined *Delay Time* parameter. This position acts as a *reference delay time* in the FBDL which determines a *reference segment* and *reference feature set* to determine *viable segments* for selection.
4. **Viable Segments**: these indicators (orange) represent *viable segments* for selection, meaning a *read segment* (delay tap) may be placed at one of these positions in the buffer. Notably, these *viable segments* will be sonically related to the *reference segment* by a degree determined by the *Target Radius* parameter. Note: there is typically an implied viable segment behind the delay time indicator.
5. **Read Segments (Delay Taps)**: these lines (white) represent the current positions being sampled from the delay buffer. The strength corresponds to the gain and the vertical position corresponds to the pan. Note that these are clustered around the viable segment positions (the ones on the left are from recently viable segments).

<u>Feature Space / Concatenative Synthesis View:</u>

This view represents the feature space (n-dimensional qualitative space) of the audio contained in the FBDL buffer, as well as data related to resynthesis (concatenative synthesis) such as *target position*, *target radius*, *read segments*, and more.



1. **Feature Axes**: two features can be selected to plot in the graph. Currently (barely visible) in the graphic it is set to *RMS* (loudness) on the x axis and *spectral centroid* (brightness) on the y axis. The full vector of features can be selected as well as a best fit plot based on principal component analysis.
2. **Targeting Method**: a targeting method of *auto* or *manual* can be selected, which determines how the target position will move through the space. *Auto* moves the position automatically based (primarily) on the *Delay Time* parameter, while *manual* is up to the user. Further methods are possible but not currently provided.
3. **Segments in Buffer**: Each point represents a segment of audio in the delay buffer, with its position encoding its characteristics along the given feature axes. The faint lines represent connections in time between segments.
4. **Target Position**: the current position used to select *read segments* for synthesis.
5. **Target Radius**: the current radius used to select *read segments* for synthesis.
6. **Read Segments**: the segments currently being sampled from. Notably, these segments will be related sonically but not necessarily in time (as we see them spread out in the delay line view).

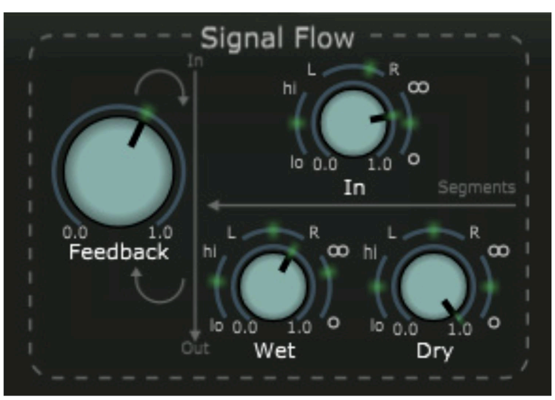
Control Parameters:

These parameters determine the high level control of the system, mostly consisting of "targeting features," or features that determine target position and control the selection of segments used for synthesis.

Delay Time: the reference delay time used for segment selection. Can be in seconds or beats.
Radius: the target radius used for segment selection. Loosely controls the amount of variation introduced into the delayed signal.
Latching %: the percent chance a target position will "latch," or repeat in the same position.
Smoothing: the smoothing factor (averaging) applied to the movement of the target position.
Buffer Size: the size of the buffer available for segment selection. The "memory."

Signal Flow Parameters:

These parameters control the flow of audio throughout the system, including input, output, and *feedback*. Each signal stage (with the exception of feedback) also includes panning, stereo width, and filtering parameters.

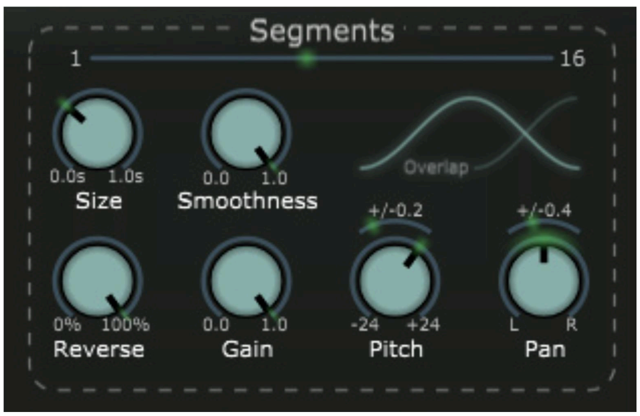


In: signal fed into the delay line (pre-analysis).
Dry: unaffected signal (bypass).
Wet: signal output from the FBDL (after feedback). Synthesis of all read segments (delay taps) into "delayed" signal.
Feedback: signal fed back into the FBDL (before wet signal output).

Segment Parameters:

Parameters related segmentation and synthesis of segments from the FBDL, including *segment size*, *number of segments* (layers), *windowing*, and more. Also includes a segment graphic displaying the gain window applied to each segment and (optionally) the overlap.

Some parameters (pitch and pan) included a variance slider which sets a range of values which will be spread evenly across all concurrent segments, allowing for chorus effects.



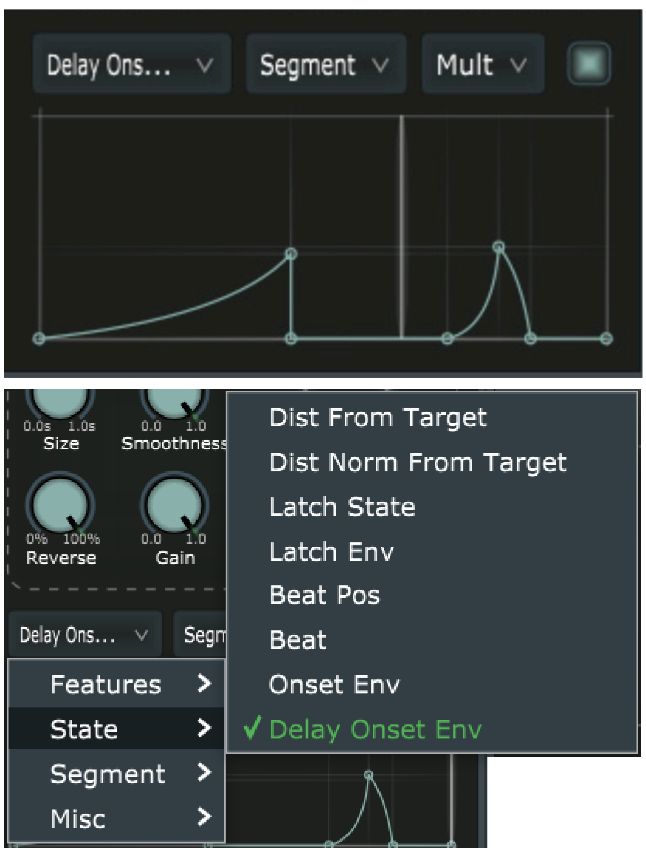
Segments: number of segments (layers or delay taps) currently read from the FBDL.
Size: size of segments written into the buffer. In seconds or beats, with a best fit option based on onset detection.
Smoothness: the windowing parameter controlling the attack and decay of the window applied to each segment.
Reverse %: chance that a segment will be read in reverse.
Gain: gain applied.
Pitch: pitch applied (semitones).
Pan: pan applied.

Parameter Graphs:

Most parameters in the plugin have associated graphs which can be used to control the parameter in an array of advanced behaviors. The source (x axis) of the graph can be set to many different things including feature vector analysis data, onset envelope, beat position, as well as other information about the state of the FBDL, and more. The y axis will either set the value of the parameter or multiply into the value of the parameter.

These graphs can be mixed and matched for fine control of the behavior of the delay, as well as introducing unique and powerful character to the expression of the system.

Mouse over a parameter to preview the graph, and either hover for a half second or click on the parameter to lock in the graph, so it can be edited.

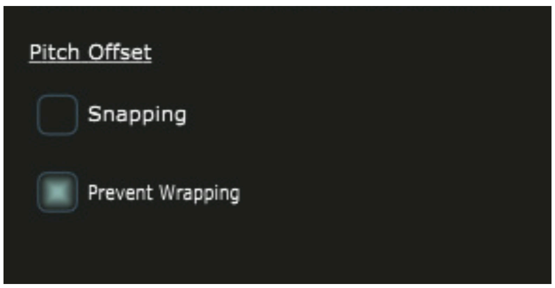


Graph Source: the x axis, which will be sampled to determine the output (y axis) of the graph. Includes the full feature vector, envelopes including onset and latch envelope, and much more. A full list of sources is (will be) available later in this manual. The graphic (left) is currently set to delayed onset envelope, i.e. the time since the last onset at the delay position in the buffer.
Feature Vector Source: the source of the feature vector used by the graph (if the graph is using a feature or other feature related source). When applicable, includes read segment, target, newest segment, and newest frame.
Application Method: method used to apply the graph; either setting the parameter value, or multiplying into the parameter value.

Parameter Context:

In addition to a graph, most parameters will have a couple options which will fall into this area. Examples include, BPM and snapping options, gain, pan, and filter, details, etc.

Mouse over a parameter to preview the context, and either hover for a half second or click on the parameter to lock in the context, so it can be edited.

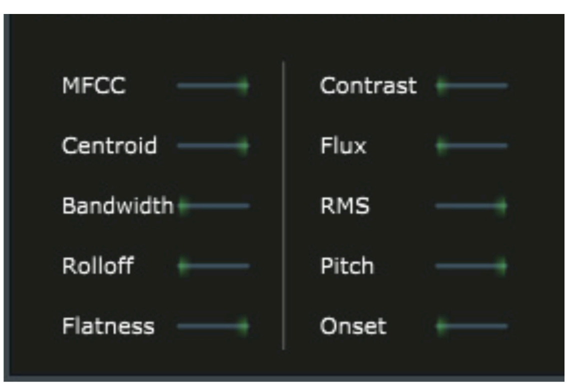


In this example, the segment pitch parameter is selected. The options *Snapping* and *Prevent Wrapping* which relate to the pitch parameter are thus displayed.

Feature Weights:

The feature weights determine the strength of each feature in the segment selection process, while the GUI also serves as a way to view and interact with the different features (try inspecting the context).

Essentially each of these features is a characteristic by which a segment in the feature space might be determined as "too different" and thus eliminated. With all weights turned off, any segment can be selected, and with only a single weight turned on, only that characteristic will be checked.



In this example, we see some features turned fully on and some features turned fully off. MFCC, Centroid, Flatness, RMS, and Pitch will be used for segment selection while Bandwidth, Rolloff, Contrast, Flux, and Onset will not.

A full list of features and their descriptions is written later in this manual.

Feature Extraction / MIR:

MFCC (Timbre):
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 (Brightness):
A measure of the "brightness" of a sound. This feature measures the center of the frequency spectrum of a sound.

Spectral Bandwidth (Noisiness):
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 (Cutoff):
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 (Timbre):
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 (Noisiness):
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 (Change):
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.

RMS (Loudness):
A measure of "loudness." This feature measures "root mean squared" of all frequencies in the FFT analysis.

Fundamental Frequency Estimation (Pitch):
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 (Attack / Decay):
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.