Configuring spectrographic views

To create a new spectrogram, spectrogram slice, or selection spectrum view, click on the appropriate New View button in the view toolbar (Figure 5.3) or choose a view type from the View > New menu.

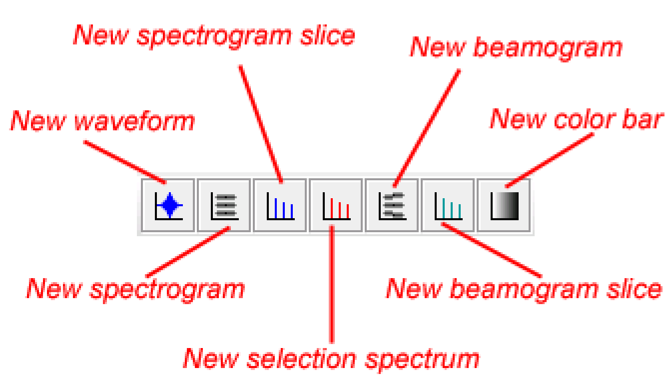


Figure 5.3. The New View buttons, in Raven's view toolbar.

A dialog box appears, containing parameters for configuring the requested type of spectrographic view (Figure 5.4). The dialog boxes for configuring spectrogram and spectrogram slice views are identical, except for their titles. The dialog boxes are identical because both view types calculate a spectrogram of the entire sound; the only difference between spectrogram and spectrogram slice views is in how the data are displayed (see "How the spectrographic views are related" on page 110). The dialog box for configuring a selection spectrum view is the same, except that it lacks the Averaging parameter.

The remainder of this section explains each of the parameters in the configuration dialog box.

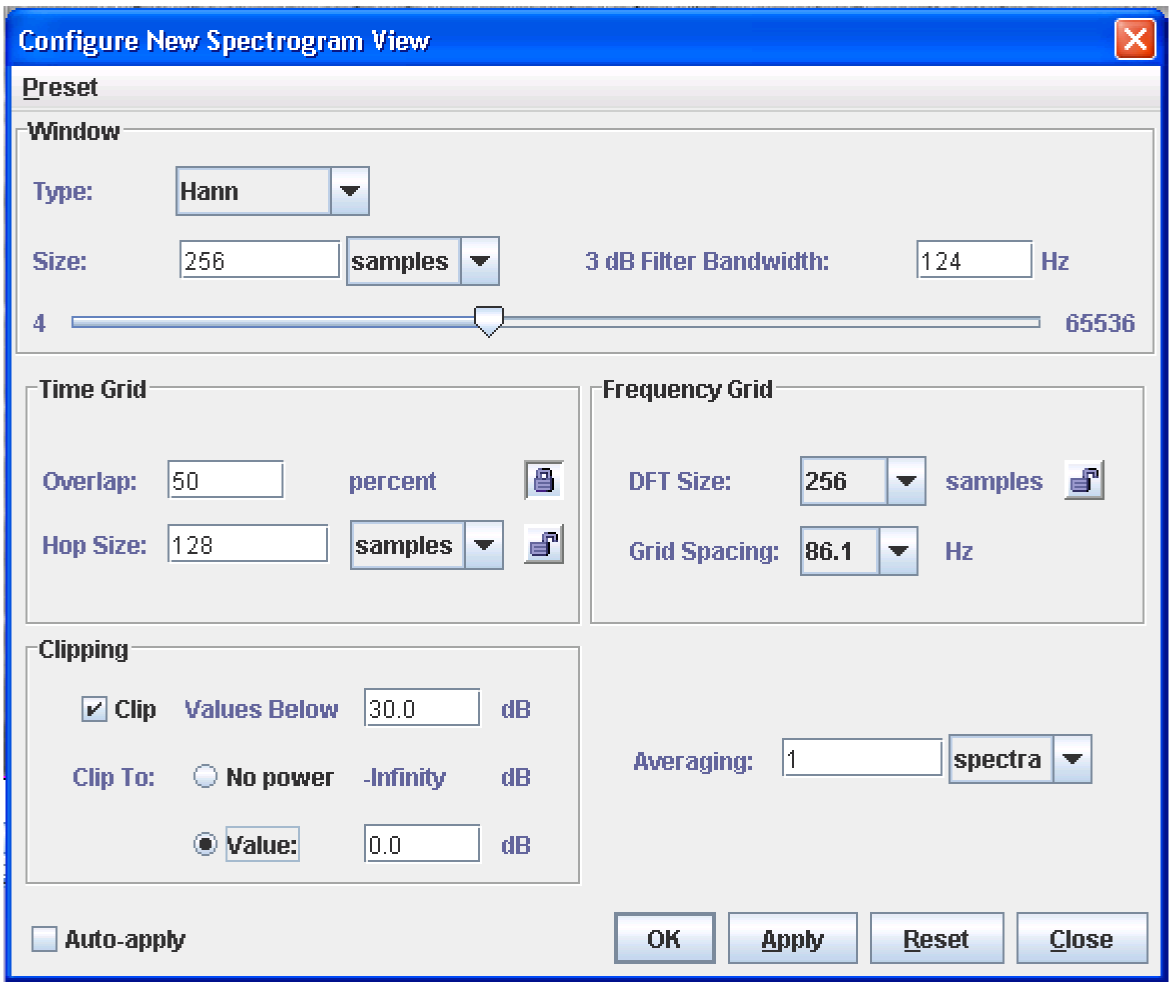


Figure 5.4. The Configure New Spectrogram dialog box.

Window type — Each data record is multiplied by a *window function* before its spectrum is calculated. Window functions are used to reduce the magnitude of spurious "sidelobe" energy that appears at frequencies flanking each analysis frequency in a spectrum. These sidelobes appear as a result of analyzing a finite (truncated) portion of a signal. A window function can reduce these sidelobes by “tapering” the portion of the waveform that appears in each window. Window functions are discussed further in Appendix B, “A Biologist’s Introduction to Spectrum Analysis”.

Raven provides six different window functions: Blackman, Hamming, Hann (sometimes called Hanning), Kaiser, rectangular, and triangular (sometimes called Bartlett). Each window function is characterized by the magnitude of the sidelobes relative to the center lobe. The difference in decibels between the center lobe magnitude and the magnitude of the largest sidelobe is called the *sidelobe rejection* (Figure B.10 on page 342). In a grayscale spectrogram, differences among windows in sidelobe rejection result in different amounts of gray “fringing” above and below black or very dark areas.

For a given window size, different window functions will result in different filter bandwidths (see “3 dB Bandwidth” on page 117). In terms of a spectrogram, this means that the vertical thickness of a horizontal line representing a pure tone will depend on which window function is used.

Figure 5.5 and Figure 5.6 illustrate the effect of different window functions on spectrogram and spectrogram slice views of the same signal.

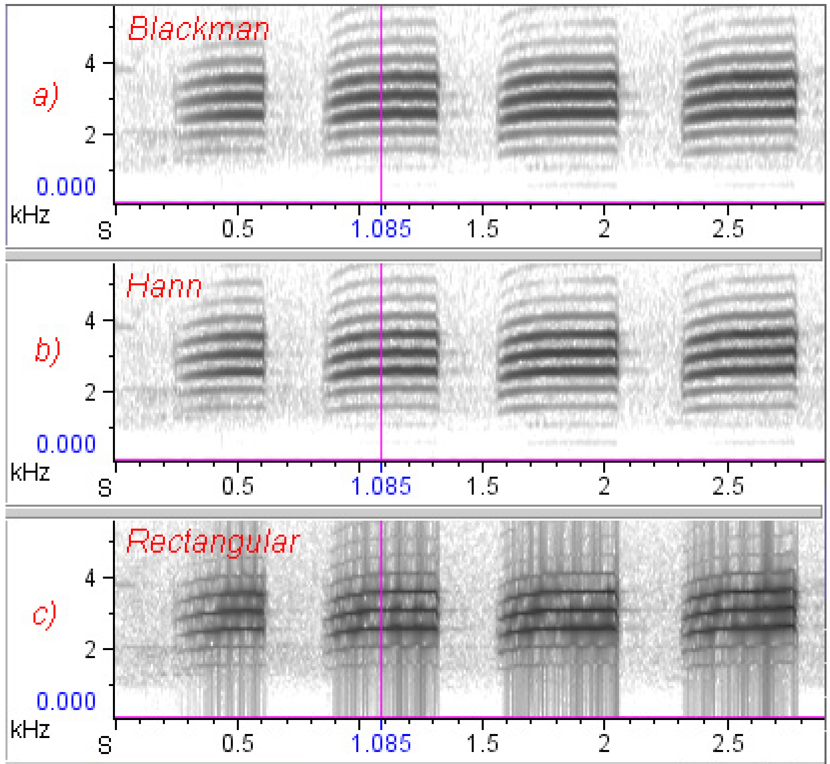


Figure 5.5. Effect of choice of window function on spectrograms. The signal is a series of calls from a red-breasted nuthatch, digitized at 44.1 kHz. All three spectrograms have the same window size (= 512 points, 11.6 mS), hop size = 5.8 mS (frame overlap = 50%), and frequency grid spacing = 86.1 Hz (DFT size = 512 samples). 3 dB bandwidths: **(a)** 141 Hz, **(b)** 124 Hz, **(c)** 76.2 Hz.

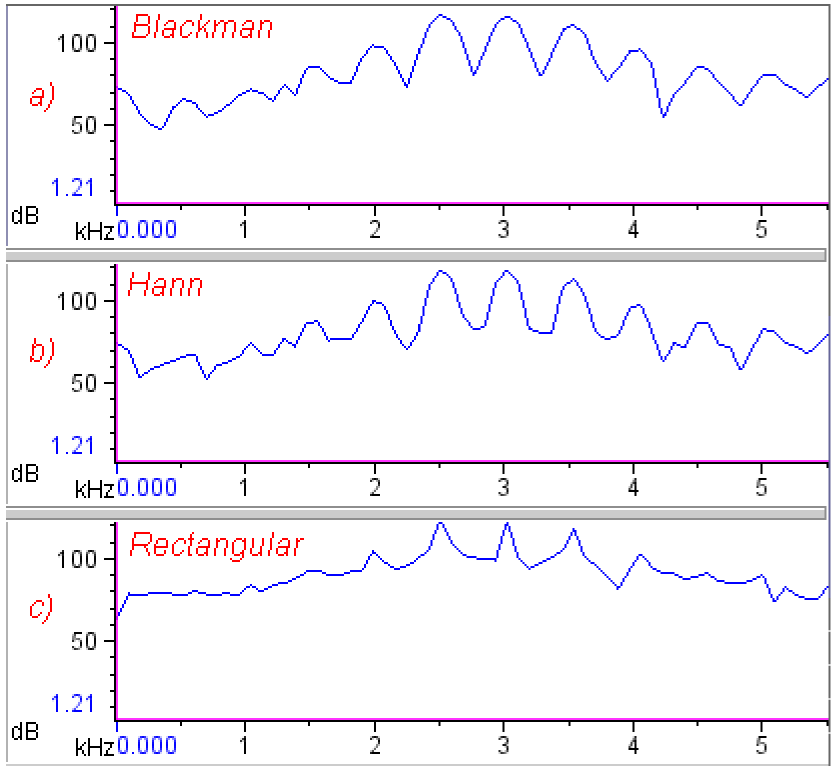


Figure 5.6. Effect of window function on spectrogram slice views. These spectrogram slice views were made at the point indicated by the position marker in Figure 5.5, midway through the second call. All three spectra have the same window size (= 512 points, 11.6 mS), and frequency grid spacing = 86.1 Hz (DFT size = 512 samples). 3 dB bandwidths: **(a)** 141 Hz, **(b)** 124 Hz, **(c)** 76.2 Hz.

The appearance of sidelobes in spectra of finite-length signals, the use of window functions to reduce their magnitude, and differences among the various window functions are discussed further in Appendix B, "A Biologist's Introduction to Spectrum Analysis".

Window size The Window Size parameter controls the length of each data record that is analyzed to create each of the individual spectra that together constitute the spectrogram. You can specify window size either in number of samples from the digitized signal, or in time units (seconds or milliseconds) by choosing the preferred unit from the drop-down menu. The default unit is samples. If you specify window size in seconds or milliseconds, Raven uses the number of samples that most closely approximates the window size that you enter.

The maximum value of the Window Size parameter depends on whether the DFT Size parameter is locked, as discussed in "Frequency grid spacing and DFT size" on page 122. When DFT Size is unlocked, Window Size can be set to a maximum of 65,536 samples (= 2^{16}). When DFT Size is locked, the maximum value of the Window Size is equal to the DFT Size.

Window Size slider control

Adjacent to the Window Size field is a slider control that provides an alternate means for changing the window size. Sliding the control to the right increases the window size. The control is logarithmic: the farther the slider is moved to the right, the more the window size changes in response to a given movement. The window size slider is useful primarily when the Auto-apply checkbox is checked (see “Apply and Auto-apply” on page 129). When Auto-apply is turned on, Raven recalculates the spectrogram immediately as you adjust the slider, allowing you to instantly see how changes in window size affect the tradeoff between time and frequency resolution.

> Use of the Window Size slider with Auto-apply turned on may result in unacceptable delays in redrawing the spectrogram with longer signals and/or slower computers.

Beta (Kaiser window only) For the Kaiser window, you can set an additional parameter, called Beta, to values between 0 and 20. For a given window size, higher values of Beta result in larger filter bandwidths and smaller sidelobes.

3 dB Bandwidth 3 dB bandwidth is the filter bandwidth of the individual analysis filters in the filterbank simulated by the short-time Fourier transform (STFT) with the selected window type and size (see Appendix B, “A Biologist’s Introduction to Spectrum Analysis”). Specifically, the 3 dB Bandwidth field displays the width (in Hz) of the main lobe of the spectrum of a sinusoid at the point where the power is 3 dB lower than the maximum power in the spectrum (Figure 5.7).

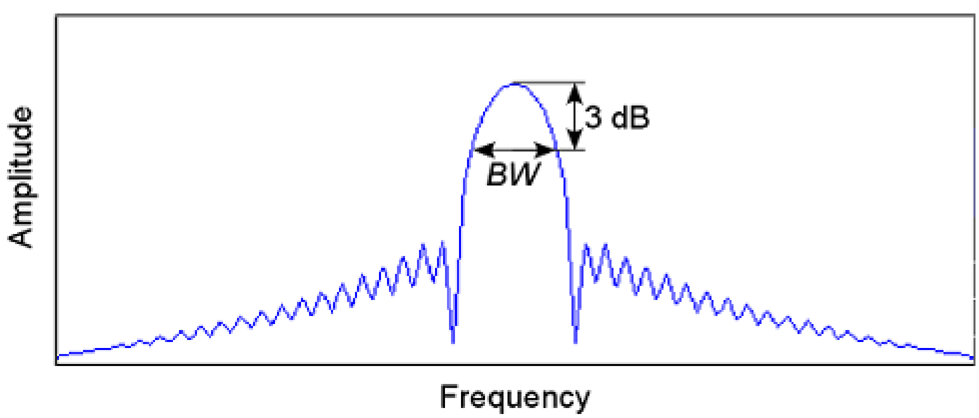


Figure 5.7. Spectrum of a pure tone sinusoidal signal. The 3 dB bandwidth is the width *BW* (in Hz) of the spectrum's main lobe at the point where the power is 3 dB less than the maximum power in the spectrum.

When you change the window size or the window type, the 3 dB Bandwidth field is immediately updated to display the corresponding bandwidth. For a given window type, improved time resolution (shorter windows) inevitably results in poorer frequency resolution (larger bandwidths). You can edit the 3 dB Bandwidth field to specify a desired value directly. When you press <Enter>, click on another field in the dialog, or click OK or Apply, Raven will choose the window size that results in the closest available approximation to the 3 dB Bandwidth value you entered. For further discussion of the tradeoff between time and frequency resolution in spectrograms, see Appendix B, "A Biologist's Introduction to Spectrum Analysis".

Choosing the window size

In a spectrogram, where you are typically interested in frequency variations with time, the "best" choice of window size depends in part on the nature of the signal, and on what features you are most interested in observing or measuring. If you are most concerned with precise frequency measurements, you will probably want to choose a large window size (hence better frequency and poorer time resolution). If you want better time resolution, choose a shorter window size; the bandwidth will then be larger (poorer frequency resolution; Figure 5.8).

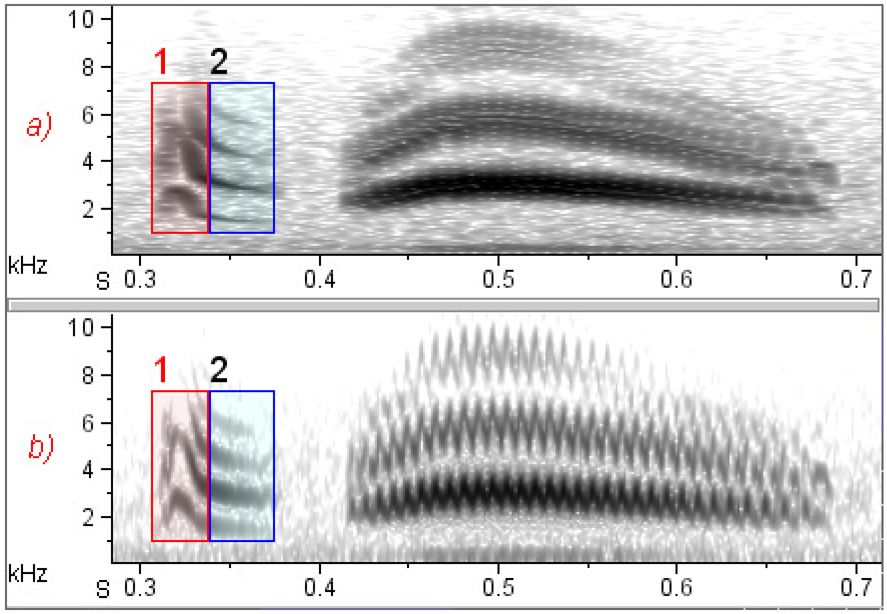


Figure 5.8. Effect of choice of window size on time and frequency smearing in spectrogram views of Cassin's kingbird sound (digitized at 44.1 kHz). For both views, window type = Hann, hop size = 64 samples, Frequency grid spacing = 22 Hz. **(a)** Window size = 800 samples, 3 dB bandwidth = 79 Hz. **(b)** Window size = 150 samples, 3 dB bandwidth = 423 Hz. View (a) has better frequency resolution (note sharpness of the nearly constant-frequency bands in selection #2), but poorer time resolution (note horizontal smearing of the rapid downsweep in selection #1, and in the oscillating frequencies in the second part of the call).

Figure 5.9 shows an extreme example of how choice of window size can change the appearance of a spectrogram. See Appendix B, "A Biologist's Introduction to Spectrum Analysis" for further discussion and more examples of the effect of varying window size and bandwidth.

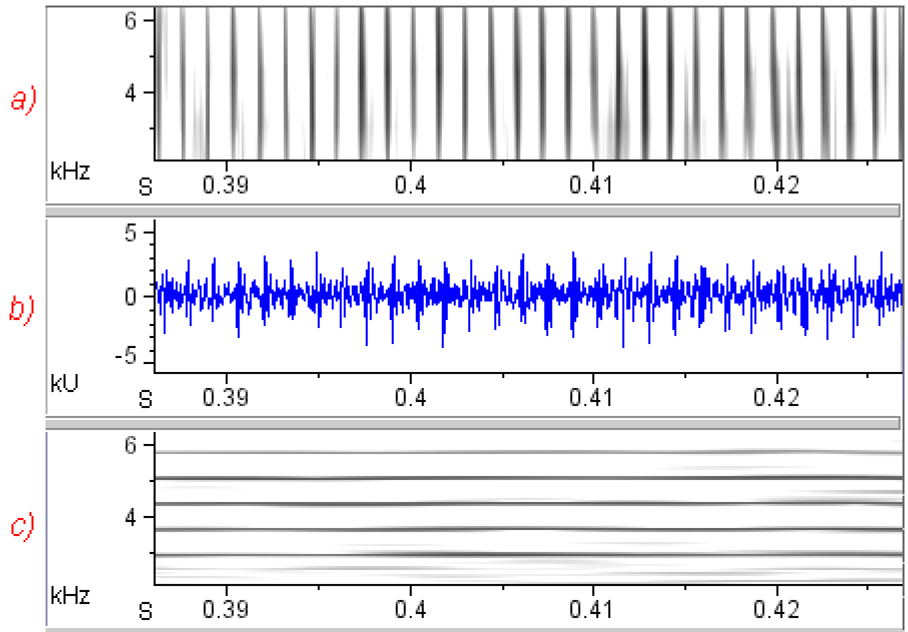


Figure 5.9. Effect of varying analysis resolution on spectrograms. The signal is part of a rapid series of clicks produced by a spotted dolphin, digitized at 48 kHz. The period between clicks is about 1.4 mS, corresponding to a frequency of about 720 Hz (= 1/0.0014). The two spectrograms differ only in window size, and hence bandwidth. In both spectrograms, hop size = .208 mS, window = Hamming. **(a)** Bandwidth = 3121 Hz (window size = 20 points = .417 mS), overlap = 50%. In this representation, each click appears as a broad-band vertical stripe on the spectrogram because the window size is short enough to resolve individual clicks. **(b)** Waveform. When played at normal speed, the signal sounds to a human like a buzz. **(c)** Bandwidth = 61 Hz (window size = 1024 points = 21.3 mS), overlap = 99%. In this representation, individual clicks cannot be resolved because each window encompasses about 15 clicks; instead the click repetition frequency appears as a series of horizontal bands spaced 720 Hz apart (the click repetition frequency).

Time grid: Window Overlap and Hop Size

Hop size[1] refers to the time interval (measured either in samples or in time units such as seconds or milliseconds) between the beginnings of successive windows or records. In an unsmoothed spectrogram (see "Smoothed vs. unsmoothed display" on page 134), the hop size can be seen as the width or duration of the individual cells in the spectrogram (Figure 5.10). Hop size can be smaller than the window size because successive windows can overlap each other. Windows can also be contiguous (0% overlap) or separated by time intervals that are omitted from the analysis (negative overlap).

1. Hop size was called time grid spacing in versions prior to Raven 1.2.

Window overlap is usually expressed as percent of window size. For example, an overlap of 50% means that each window begins halfway through the preceding window. An overlap of -100% means that one window of data is skipped between successive windows that are analyzed; -300% skips three frames, and so on. The relationship between hop size and window overlap is given by

$$\textit{hop size} = \textit{window size} * (100\% - \textit{overlap}\%).$$

The Hop Size and Overlap fields in the dialog box are coupled so that you can specify hop size either directly, by typing a value in the Hop Size field, or indirectly, by typing a value in the Overlap field. Using the units drop-down menu, you can specify the measurement units for hop size as either samples (the default), seconds, or milliseconds. If you enter a value in the Overlap field that does not correspond to an integer number of samples, Raven substitutes the closest overlap value that does.

Figure 5.10 shows three spectrograms that differ only in hop size.

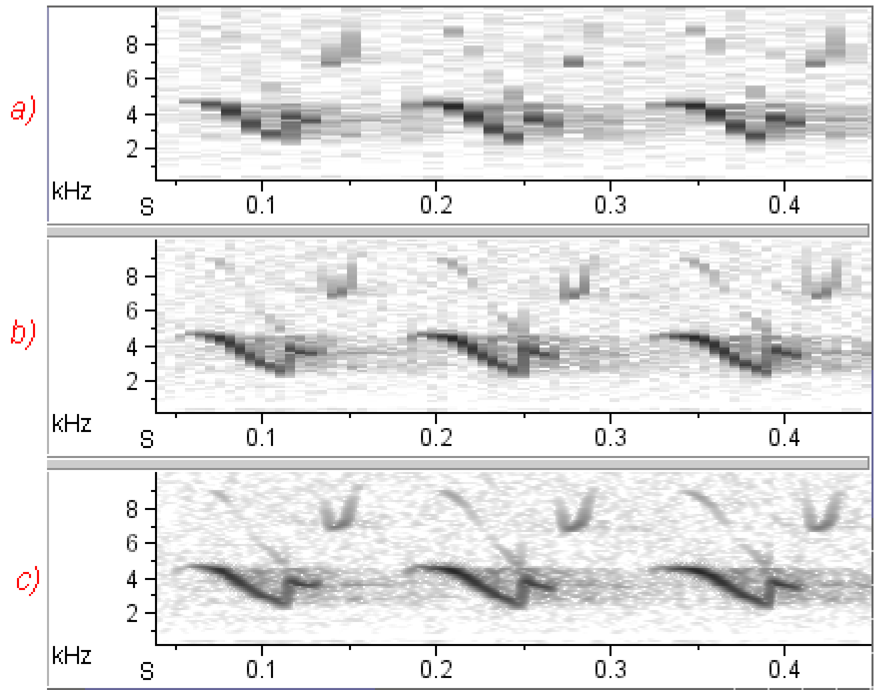
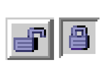


Figure 5.10. Effect of varying hop size in spectrograms. The signal is part of a song of a lark sparrow, digitized at 44.1 kHz. The three spectrograms are unsmoothed and differ only in hop size (window overlap). In all three spectrograms, window type = Hann, window size = 512 samples (= 11.6 mS; 3 dB bandwidth = 124 Hz), frequency grid spacing = 86.1 Hz (DFT size = 512 samples). **(a)** Hop size = 11.6 mS (window overlap = 0%). **(b)** Hop size = 5.8 mS (window overlap = 50%). **(c)** Hop size = 1.1 mS (window overlap = 90%).

A spectrogram made with a negative window overlap ignores some of the available data, and can give an extremely misleading picture of a signal. Negative window overlaps should generally be avoided unless you have some specific reason for wanting to omit some parts of a signal from analysis.



Lock Overlap vs. Lock Hop Size

Next to the **Overlap** and **Hop Size** fields are two buttons, marked with open and closed padlock icons. The button that displays the closed padlock indicates which value— window overlap or hop size— will be locked or held constant when you make changes to the window size. Clicking on either button reverses the state of both buttons.

Frequency grid spacing and DFT size

The frequency grid spacing of a spectrogram (visible as the height of the individual boxes in an unsmoothed spectrogram; see “Smoothed vs. unsmoothed display” on page 134) depends on the sample rate (which is fixed for a given digitized signal) and a parameter of the STFT called *DFT size*.[2] The relationship is

frequency grid spacing = (*sampling frequency*) / *DFT size*

where frequency grid spacing and sampling frequency are measured in Hz and DFT size is measured in samples. DFT size is constrained to be a power of 2 that is greater than the current window size.

The DFT Size and frequency Grid Spacing fields in the Configure Spectrogram dialog box are linked: you can specify the frequency grid spacing either directly by choosing a value from the Grid Spacing drop-down menu, or you can choose a value from the DFT Size drop-down menu. The DFT Size menu displays powers of 2 greater than or equal to the current window size. Larger DFT sizes correspond to smaller frequency grid spacings.

Lock DFT Size

Next to the DFT Size drop-down menu is a button marked with a padlock icon. When this button is unlocked (the default), Raven adjusts the DFT size as you change the window size, in order to maintain a consistent relationship to the window size, subject to the constraint of being a power of 2. For example, in the default spectrogram parameters, the DFT size is the smallest power of 2 greater than or equal to the window size. If you increase the window size from 512 to 513, Raven changes the DFT size from 512 to 1024. If you manually choose the DFT size to be, for example, the second power of two greater than or equal to the window size (e.g., with window size of 512, you set DFT Size to 1024 instead of 512), then Raven will change DFT Size to maintain this relationship as you adjust window size.

When the DFT Size padlock button is locked, the DFT size (hence frequency grid resolution) is fixed, and will not change when the window size changes.

> Because the window size cannot exceed the DFT Size, the maximum value you can specify for Window Size, either by typing a value, or by moving the slider control, is limited to the DFT Size value when DFT Size is locked.

Clipping level

The Clipping Level parameter allows you to specify a “noise floor” below which any amplitude value is altered, in order to reduce or eliminate the

---

2. The parameter that Raven calls DFT Size is sometimes called FFT size in other programs. FFT stands for *fast Fourier transform*, which is a particular algorithm used to compute the *discrete Fourier transform* or DFT. Size is a characteristic of a particular DFT, not of the FFT algorithm used to compute it.

effect of the noise. After you enable clipping, you can modify it using two parameters: (1) the power level below which values will be altered, and (2) the power level at which to set the altered values.

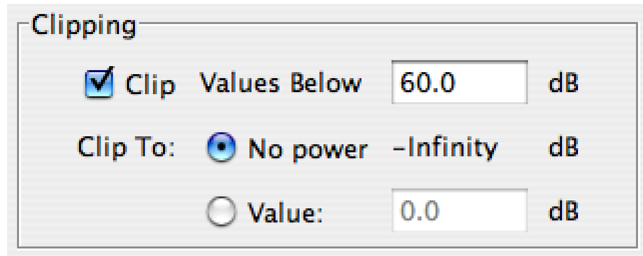


Figure 5.11. The Clipping dialog box.

In a spectrogram view, if you choose to clip to a value of -Infinity, everything that is below your noise floor will appear as white. If you clip to a different value (for example, 0 dB), everything below your noise floor will appear as very light. These differences will result in different appearances once your spectrogram is smoothed, so you may wish to vary your parameters and examine the results as you explore. See Figure 5.12 for an example.

In a spectrogram slice view, if you choose to clip to a particular dB value, then the result will be a smooth slice view without any dropouts; but if you choose to clip to -Infinity dB, the result will contain many dropouts and will be more difficult to read. See Figure 5.13 for an example.

There are several reasons to perform clipping. First, because of the finite precision of the digitization process, a digitized sound always contains some error and has a limited dynamic range. For signals digitized with 8-bit samples, the dynamic range is limited to 48 dB; for 16-bit samples, the dynamic range is limited to 96 dB.[3] Therefore any power value in an 8-bit spectrogram that is more than 48 dB below the highest peak in the signal must be noise introduced by the digitizing process, and should be disregarded.[4] The noise floor can also be useful for removing noise that was present before the digitizing process (for example, from a recording with low-level wind or other broad-band noise).

A more pragmatic reason for noise clipping is that very small power values show up on a log scale as large negative dB levels (because the

---

3. The dynamic range of a digitized sound is 6 dB/bit.
4. If the highest spectral peak in a signal is smaller than the digitizer's maximum output level, the dynamic range between the peak and noise introduced by digitizing will be less than 6 dB/bit. However, it can never be more than 6 dB/bit.

logarithm of zero is negative infinity). The noise floor allows Raven to ignore very small power values.

Finally, the noise floor can also be used to eliminate spectral sidelobes (which show up as gray fringes around strong signal components in spectrograms).

If the noise floor is set too low, excessive noise will be displayed in the spectrogram or spectrogram slice along with the signal. If it is set too high, portions of the signal will not be visible. You may need to experiment with different clipping levels in order to find a value that produces a satisfactory display.

As an alternative to clipping, you can alter the look of a spectrogram without changing the underlying spectrogram data by altering the brightness and contrast of the view. Note, though, that this method does not change the noise floor in the spectrogram data stored in memory. If you are only using spectrograms for visual examination and display, then the distinction between brightness/contrast and the noise floor of the spectrogram data is unimportant. However, if you plan to do any quantitative analysis (e.g., correlations) using the spectrogram data, remember that the only way to change the noise floor is to recalculate the spectrogram, specifying a different Clipping Level in the spectrogram dialog box.

If you want to raise or lower the noise floor in a spectrogram (either in the display or in the data itself), you must recalculate the spectrogram with a higher or lower Clipping Level.

Figure 5.12 and Figure 5.13 show spectrograms and spectrogram slices that differ in Clipping Level as well as in brightness and contrast. Note that if you want to have different clipping levels for your spectrogram and spectrogram slice views, you must unlink those views in their spectrogram parameters.

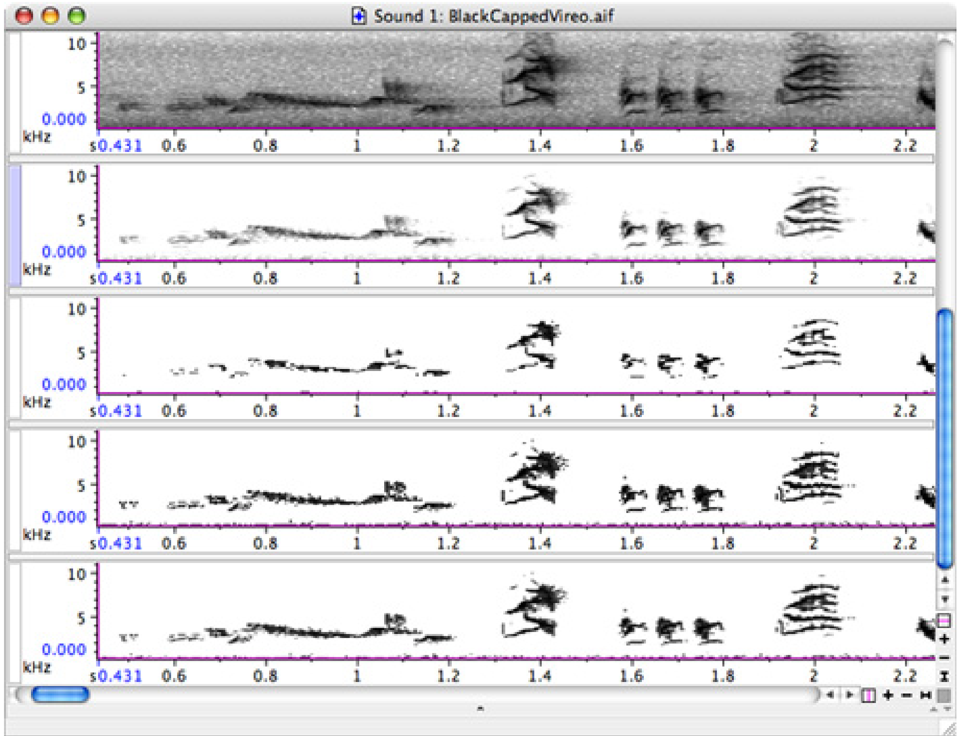


Figure 5.12. Effect of clipping level on spectrograms. The signal is part of a song of a Black-capped Vireo. First row: no clipping level set, standard brightness (50) and contrast (50) settings, spectrogram smoothing enabled. Second row: no clipping level set, brightness set to 65, contrast set to 78, smoothing enabled. The underlying power values have not changed so all measurement values will be the same as those measured in the first row. Third row: clipping enabled, values below 60 dB clipped to -Infinity dB, spectrogram smoothing enabled. Measurement values will differ between this view and the first row. Fourth row: clipping enabled, values below 60 dB clipped to -Infinity dB, spectrogram smoothing disabled. Without smoothing, more of the clipped spectrogram is visible. Fifth row: clipping enabled, values below 60 dB clipped to 0 dB, spectrogram smoothing enabled. Since values were only clipped to 0 dB, more of the clipped values are visible than in the third row.

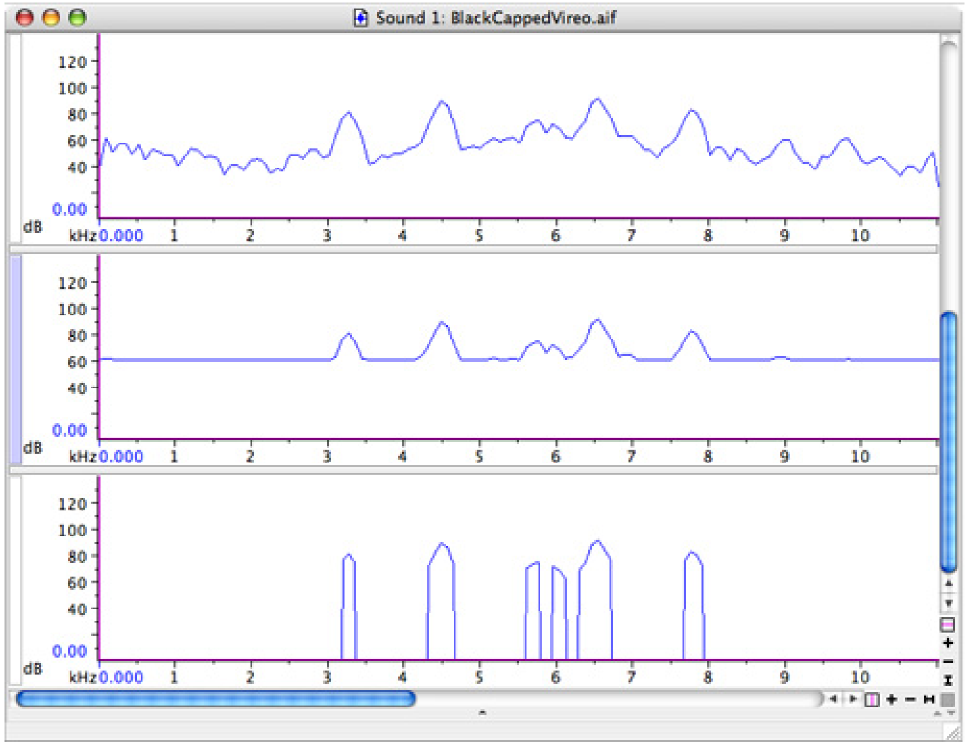


Figure 5.13. Effect of clipping level on spectrogram slices. All three slices are on a single 256-point frame approximately 1.378 seconds into the Black-capped Vireo song shown in Figure 5.12. The signal was digitized with 16-bit resolution, and thus has a dynamic range of 96 dB. The highest power values are near 90 dB, so there is little perceptible noise inserted by the digitization process. First row: no clipping level set. Second row: clipping level = 60 dB, with values clipped to 60 dB. Third row: clipping level = 60 dB, with values clipped to -Infinity dB. Notice the dropouts in the slice view. For this reason, we recommend against clipping spectrogram slice views to -Infinity dB, but for analysis purposes, you may want to clip spectrograms to -Infinity dB.

Spectrum averaging

The Averaging field allows you to specify the number of individual spectra over which Raven should average the power values to obtain the values in each cell of the spectrogram. In most situations, Averaging should be left at its default value of one spectrum.

Higher Averaging values may provide more satisfactory spectrogram images when more than a few seconds of a signal are displayed. There are two reasons why averaged spectrograms may be preferable at certain time scales. First, if the time scale of a spectrogram view is such that the number of spectra in the visible time span is much greater than the number of pixels in the time dimension of the sound window, then many spectra will not be displayed at all. Some acoustic events that span only a few spectra may not be visible (unless you zoom in to display a finer time scale) because the only spectra in which they appear fall between the pixels shown in the display. By setting the Averaging field to a value greater than one spectrum, you can make visible short-duration events that would otherwise be lost between pixels. Second, spectrum averaging smooths background noise, which can result in a higher signal-to-noise ratio in the spectrogram image (Figure 5.14). At finer time scales (i.e., greater magnification in the time dimension), however, spectrum averaging tends to blur signals (Figure 5.14).

You can specify the amount of data to average in units of seconds or milliseconds (rather than spectra), using the units drop-down menu.

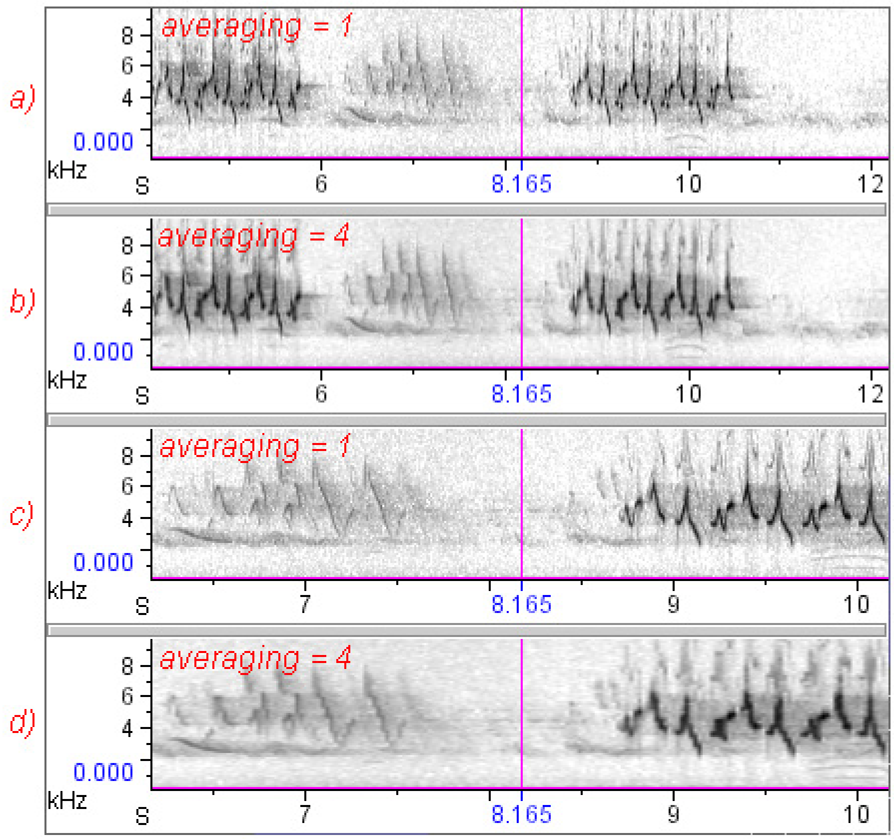


Figure 5.14. Spectrum averaging can yield clearer spectrograms when the interval in view is long compared to the number of spectra in view. All four spectrograms of a nearby common yellowthroat and a distant yellow warbler were made with Window Type = Hann, Window Size = 512 samples, Window Overlap = 50%. All four are linked by time position. **(a)** and **(c)** Averaging = 1 spectrum. **(b)** and **(d)** Averaging = 4 spectra. When viewing a longer time span (views (a) and (b)), the view that uses spectrum averaging provides a clearer image, especially of the faint signal from a distant bird. For the more magnified image ((c) and (d)), the view without averaging appears clearer.

Apply and Auto-apply

If you click the Apply button, Raven immediately calculates and displays the spectrogram, using the parameters currently displayed in the dialog, without closing the dialog. (Clicking OK closes the dialog before calculating the spectrogram.)

If the Auto-apply checkbox is checked, Raven immediately recalculates and displays the spectrogram each time you change any parameter in the dialog, without you needing to click the Apply button. For parameters that you enter by typing in a field (e.g., Time grid spacing or window Overlap), the spectrogram is recalculated when you complete an entry by pressing the <Enter> or <Tab> key, or by clicking another field or control in the dialog.

Spectrogram presets

You can save and retrieve sets of spectrogram parameters using commands on the Preset menu within the Configure Spectrogram dialog. A set of saved spectrogram parameters is called a *spectrogram preset*. To save a preset, choose Presets > Save As... . When the Save Spectrogram Parameters dialog appears, enter a name for the preset, and click OK.

> Spectrogram presets must be saved in the folder `Presets/Spectrogram Parameters/` within the Raven program folder. You can also create additional folders within the Spectrogram Parameters folder by clicking on the New Folder icon within the Save dialog. These folders will appear as submenus in the Preset menu, with each submenu listing the presets in the corresponding folder.

To retrieve a spectrogram preset, choose the name of the preset from the Preset menu. When you retrieve a preset, all of the spectrogram parameters in the Configure Spectrogram dialog are immediately set to the saved values. If you then change some parameters and want to revert to the saved values, click the Reset button or select the name of the preset from the Preset menu again. If you want to save changes you've made under the name of the last preset you loaded, choose Preset > Save "*PresetName*".

Spectrogram views

Significance of the color (grayscale) values

Spectrograms displayed by Raven have a logarithmic power (color) axis. That is, the color (by default, grayscale) values shown in the cells of an unsmoothed spectrogram represent the logarithm of the power at the corresponding frequency for each spectrum in a spectrogram. Hence, the color value is proportional to the power expressed in decibels (relative to an arbitrary reference power).

The numeric values for relative power level associated with each point are displayed in decibels (dB) in the mouse measurement field at the bottom