INTERNATIONAL EDITION

Digital Media Primer

Digital Audio, Video, Imaging, and Multimedia Programming

SECOND EDITION

Yue-Ling Wong

ALWAYS LEARNING

PEARSON

Digital Media PRIMER

Digital Audio, Video, Imaging and Multimedia Programming

Second Edition

Yue-Ling Wong

International Edition contributions
by Shandar Junaid

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Paris Montreal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

audio tracks is necessary and useful in situations such as when you want to mix a voice-over with background music, compose different musical clips, or mix multiple instrumental playbacks where each is recorded as a separate audio. When you are satisfied with the mixing results, you can export the audio as an audio mix-down. ***Mixing down*** means combining all of the tracks with effects. This is analogous to flattening an image in Photoshop. When you open a mixed-down audio piece, there are no separate tracks.

It is advantageous to create or record audio separately and mix them later. This way you can adjust the volume level and apply effects (such as fade-in and fade-out) to each audio independently. For example, if you record a voice-over with background music in one recording, it would require synchronization and control in matching the volume levels of voice-over with the background music all at once. What if later a change in your project requires an extra few seconds of pause within the voice-over and making the voice-over louder? You would have to record the whole audio again, because simply inserting a silence in the voice-over segment will also silence the background music. Simply making the voice-over segment louder will also make the background music louder altogether. However, if the voice-over and background music are recorded separately as different audio files, then all you need to do is to insert a few seconds of silence in the voice-over where it is needed. The volume of the voice-over can be adjusted independently. The background music will be unaffected and can be mixed with the edited voice-over as usual. The background music will still be playing during the newly added silence portion of the voice-over.

5.2.3 Spectral View

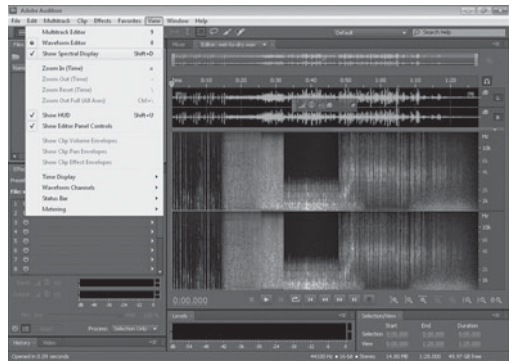
The waveform representation is the most frequently used view in digital audio editing. However, many audio editing programs also have a spectral view available. In the ***spectral view*** (Figure 5.12), the *x*-axis of the display is still time, as in the waveform display. However, the *y*-axis is the frequency of the audio, not the amplitude. The amplitude value is now represented using color. The spectral view can tell you the audio frequency information that you cannot get from the waveform view. For example, as you can tell from the spectral view shown in Figure 5.12, a middle section of the audio is missing sound that is higher than 3,000 Hz.

In general, the higher amplitude is represented by brighter colors. The brightest color of the section that is capped at 3,000 Hz occurs in the range of about 300 to 500 Hz. In the section after it, the brightest color occurs around 4,000 Hz but the 300- to 500-Hz range is very dark (low in amplitude). This means that a sound amplitude in the range of 300 to 500 Hz is very low. You would not be able to discern such frequency information from the waveform view of the same audio (Figure 5.10).

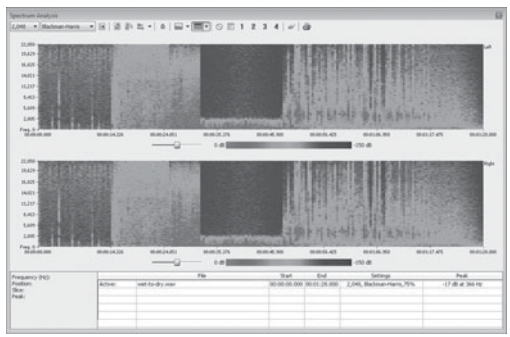
When working in the spectral view, you can still select audio data within a time range. Audition allows you to select the audio data in frequency in addition to time ranges (Figure 5.13). Thus, working in the spectral view also allows you to selectively edit a certain frequency range within a time frame. This is in contrast to selecting audio data in the waveform view, in which, when you make a selection, you select *all* the audio data within the selected time range, regardless of the frequency. Any change applied to the selection affects all frequencies within the selected time frame.

Spectral view is useful in analyzing the frequency range of the audio. It gives you an idea of the most important frequency range of the audio, so that when you need to reduce the audio file size by downsampling, you can estimate the lowest sampling rate without causing serious distortion. This is also useful in situations when it is necessary to isolate the frequency range responsible for an audio anomaly.

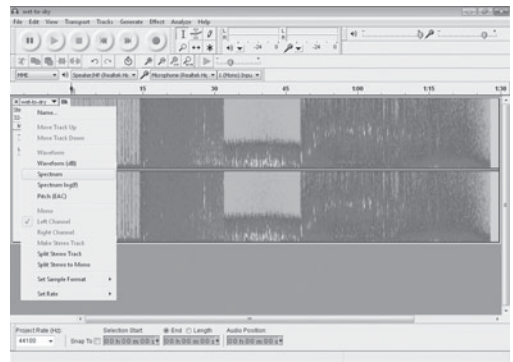
Lowest sampling rate without serious distortion: Recall the implication of the Nyquist theorem (Chapter 4)—the sampling rate of the audio must be at least twice the audio frequency. The audio with the highest amplitude usually is the most important part of the audio. If the frequency range that has the highest amplitude is around 4,000 Hz, then resampling your audio to at least 8,000 Hz.



(a)



(b)



(c)

Figure 5.12 Spectral view of audio in Figure 5.10. (a) Adobe Audition CS5.5: `View > Show Spectral Display`. (b) Sony Sound Forge 10: `View > Spectrum Analysis` and click the Sonogram button. (c) Audacity 1.3 Beta: `Spectrum`. This image can be found on the insert.

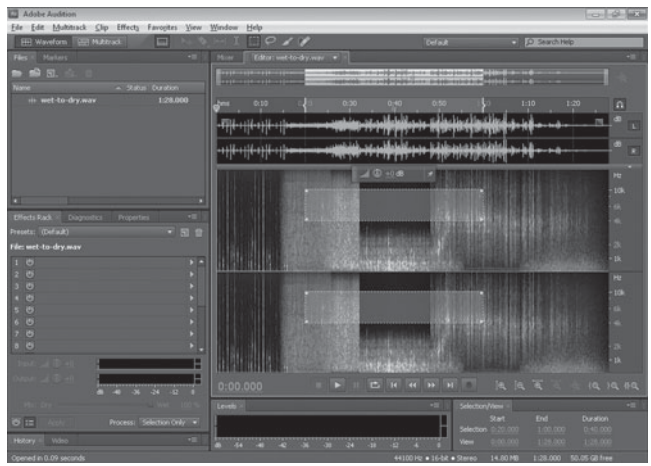


Figure 5.13 The frequency range of about 4,000 to 10,000 Hz within the time frame between 20 second and 1 minute is selected (as indicated by the pale rectangle) in the spectral view. This image can be found on the insert.

5.3 BASIC DIGITAL AUDIO EDITING

In digital imaging, the scanned or digitally photographed image almost always needs retouching and editing, whether it is for cleaning up the image or artistic manipulation. This is also true for digital audio. Making connections across media with respect to the editing process will help you understand both the need for and the process of editing. So let's draw parallels from the common digital image editing processes.

You can edit an image by copying and pasting; adjusting the tonal value; cleaning up dirt, dust, and scratches; resizing the image; reducing its color depth; and applying filters, such as blurring or embossing. The basic ideas of these processes are transferable to digital audio editing (Table 5.1). In both digital image and audio editing, the changes will be applied to the selected area. If there is not a selection made, the changes will be applied to the whole image or audio. These basic audio editing operations are described in the following sections.

TABLE 5.1 Parallels between Basic Digital Image and Digital Audio Editing

Basic Digital Image Editing	Basic Digital Audio Editing
Reassemble image content by cutting, copying, and pasting	Reassemble an audio waveform by cutting, copying, and pasting
Clean up dirt, dust, and scratches	Noise reduction
Adjust tonal value	Adjust volume
Resize the image	Resample the audio
Reduce the image's color depth	Reduce the audio's bit depth
Apply filters for special effects	Apply filters for special effects, such as reverb and pitch changes
If you want to save your file as JPEG for the Web, wait until the last step because JPEG uses lossy compression	If you want to save your file as MP3 for the Web, wait until the last step because MP3 uses lossy compression

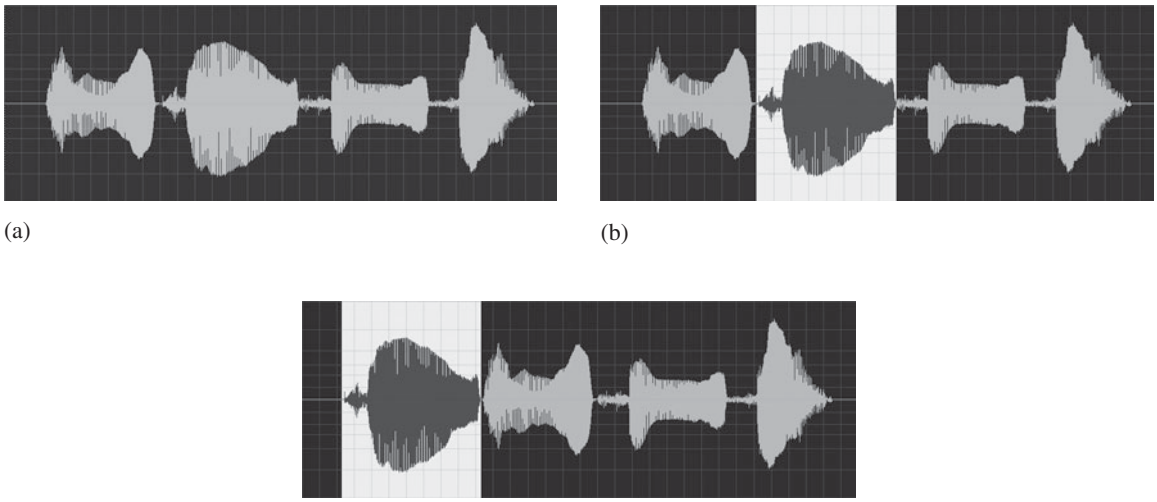
5.3.1 Reassembling a Waveform

The audio waveform can be selected and then applied with operations such as Cut, Copy, or Delete. To select a segment of a waveform, simply drag the cursor across the waveform and you can confirm the audio content of the selection by playing it.

Because the waveform is a visual representation of the audio amplitude in time, it is often possible to locate a sound segment in the waveform by looking for pauses or distinctive amplitude changes. You can edit out a specific word or pause from the audio by selecting it and then using the Cut or Delete operation. You can also click on a particular time on a waveform and paste a waveform segment that corresponds to a particular sound or speech segment. Figure 5.14a shows a waveform of the audio speech of the words one-two-three-four. The second word is cut and pasted before the first word (Figure 5.14c). The audio now becomes two-one-three-four.

Using the Cut or Delete operation, you can also trim unnecessary extra time off the beginning and end of the recording.

Reassembling Waveforms in Adobe Audition Screen-capture movie demonstrating the selecting, cutting, and pasting of audio waveforms in Adobe Audition.



(a) (b) (c)

Figure 5.14 Reassembling an audio waveform: (a) An original recording of speech saying "one-two-three-four." (b) The second word, "two," is selected. (c) The word "two" is cut and pasted before the first word.

5.3.2 Changing Volume

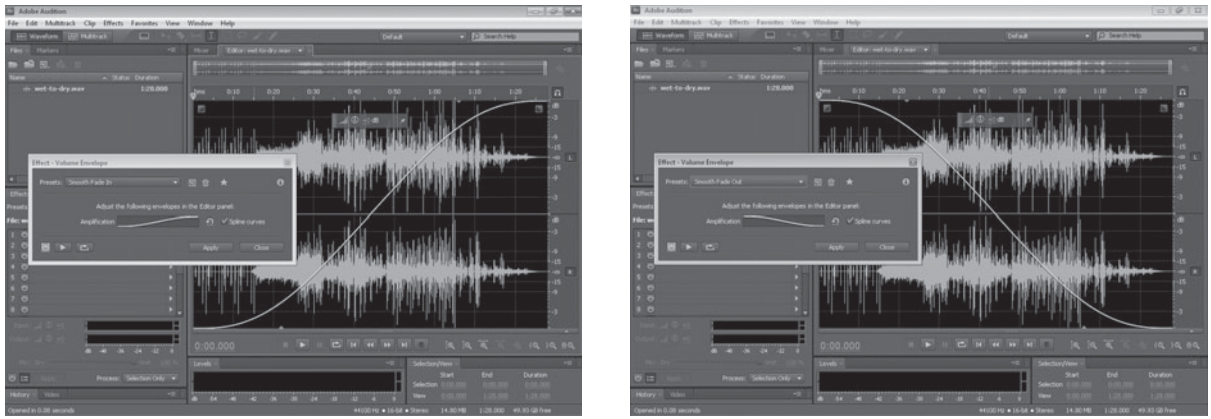
Adjusting the volume or amplitude of an audio is necessary when the recorded volume is too loud or too low for the intended use of the audio in a project. Another situation where volume adjustment is needed is when assembling multiple recordings into a single sound track. In this case, you often need to adjust the volume of each recording to give a consistent volume level across the recordings. There are two common menu items for simple volume adjustment: amplify and normalize. ***Amplify*** lets you specify amplification in dB (decibels) or percentage. ***Normalize*** lets you set a peak level for a file or selection, and it will amplify the entire file or selection so that the highest level is the peak level you specify. Use normalize instead of amplify to ensure the amplification does not result in clipping. For example, if you normalize an audio to 100%, you achieve the maximum amplitude that digital audio allows.

There are two extended applications of volume adjustment. The ***fade-in*** effect suppresses the volume at the beginning and gradually increases to 100% within the designated duration for the fade-in. Similarly, the ***fade-out*** effect lowers the volume level gradually to zero within the fade-out period.

The ***Envelope*** graph allows you to control the volume changes over time using a curve. The *x*-axis of the Envelope graph represents time and the *y*-axis represents amplification level. There are often presets available for Envelope graphs. You also can define your own curve. As you see in Figure 5.15, the fade-in and fade-out effects can be created by the Envelope graph.

5.3.3 Noise Reduction

Recordings often contain background noise, such as computer fan noise or hissing from the microphone. The two most useful features for noise reduction found in Adobe Audition are Hiss Reduction and Noise Reduction. ***Hiss Reduction*** can be used to reduce hiss from the source, such as microphone or audio cassette tape. Hissing noise is often characterized by a certain frequency range. The basic idea of how hiss reduction works is that a threshold



(a) (b)

Figure 5.15 Examples of Envelope graphs (Adobe Audition): (a) Fade-in. (b) Fade-out.

amplitude for a frequency range, called the ***noise floor***, is defined—as preset or user-defined. If the audio amplitude within that frequency range falls below the threshold, the audio of that range will be greatly reduced. However, audio outside that frequency range and audio of any frequency that is louder than the threshold will remain unchanged.

Noise Reduction (Adobe Audition) or ***Noise Removal*** (Audacity) is another effect you can use to reduce noise that is constant throughout a waveform. This feature performs a statistical analysis of a pure noise segment that is selected by the user. Then it performs noise reduction on the whole audio based on this noise profile. Because it reduces noise based on a noise profile generated statistically, not a frequency range, the Noise Reduction can remove a wider range of noise. The downside of this effect compared to Hiss Reduction is that you need to have a sufficiently long segment (preferably more than 1 second) of pure noise for the statistical analysis to generate a representative profile for the noise. To get a pure noise segment, you can start the recording with a second or more of silence. That silence can be used as the Noise Reduction Profile for the final audio recording.

Applying Noise Reduction in Adobe Audition Screen-capture movie demonstrating the application of the Noise Reduction effects to an audio recording in Adobe Audition.

5.3.4 Special Effects

There are many special effects or filters you can apply to an audio. For example, many audio programs let you alter the pitch of a sound, alter the speed of its playback, and add echo or ***reverb*** (Figure 5.16). There are a wide variety of preset reverb effects, from auditorium to bathroom. You can also customize the reverb settings.

5.3.5 Downsampling and Reduction of Bit Depth

It is advantageous to record and work with audio at CD-quality levels (i.e., 44,100-Hz sampling rate, 16 bit, and stereo) or higher. However, depending on the project, you may need to downsample the audio to lower the file size before sharing your audio. In Adobe Audition, you can alter the sampling rate, bit depth, and channel number by choosing `Edit > Convert Sample Type. . .`. Save the altered audio file as a different name so that you can keep the original, higher quality audio file for any future editing.

Figure 5.16 Applying reverb to an audio in Adobe Audition CS5.5.

See Chapter 4 for a detailed discussion on optimization of audio files and effects of sampling rate, bit depth, and number of channels on the audio file size.

Another way of lowering file size is to save the audio file as a compressed file format, such as .mp3. MP3 files use a lossy but effective compression by applying human perception of sound. This means that some audio data are removed and the audio data are approximated, but the data removed are basically imperceptible to humans. Because MP3 uses a lossy compression, you should avoid using MP3 files as the master copy for further editing. The file that you plan to use for editing should be saved using a lossless compression method, if possible.

5.3.6 General Steps of Digital Audio Recording Touch-up

Even the most controlled studio recordings require touch-up before further manipulation of the audio. The general steps of touching up digital audio recordings are:

1. **Noise reduction:** Remove background noise by applying the Noise Reduction effect.
2. **Trimming:** Remove extra time at the beginning or end of the audio recording. You also may remove unwanted long periods of silence within the audio.
3. Adjusting volume level to the desired level.

Audio Manipulation (Lab) Practice recording your voice, removing background noise, reassembling the waveform, and applying audio effects.

The audio should be touched up before applying any special effect, such as reverb. Otherwise, the background noise and the unwanted silence will be mixed with the effects and become harder, if not impossible, to remove.

5.4 MULTITRACK BASICS

The basic steps of working with a multitrack session are as follows.

Step 1 Place clips on tracks. You can always add and remove clips at any time during the process.