Understanding Multimedia Sound;

The most significant addition a MPC (Multimedia PC) makes to the traditional PC is sound. You can record your voice in order to attach voice annotations to documents. Applications can play sampled audio or music through the MIDI (Musical Instrument Digital Interface) synthesizer.

However, the concepts and terminology of this new technology can be a little intimidating at first. For example, the New File Data Format dialog box shown in Figure A allows you to configure the *audio resolution* of a recording. This dialog box appears when you begin a new recording using the Rack application that comes with a CompuAdd MPC. Terms such as *sample size* and options such as 44.1KHz may be new to you. Furthermore, unless youre a musician familiar with MIDI, youll probably find terms such as *channel number* and *polyphony* foreign.

In this article, well explain how the sound board re-cords and plays sounds. Well also explain audio resolution and the aspects of using sound on a MPC.

Figure A

The New File Format dialog box allows you to adjust the audio resolution of a recording.

How is sound recorded?

First, we need to explain how the sound board records and plays sounds. The sound board uses a technique called *pulse code modulation* (PCM) to digitize sound for storage. To illustrate how this works, well use the analogy of taking snapshots of sound.

Simply put, PCM is a millisecond snapshot of a sounds audio waveform. Using this technique, the sound board takes a large number of these snapshots every second to record sound. Using an analog-to-digital converter (ADC), the sound board converts these snapshots to a digital form. Then, the digital data is stored on your hard disk as a sampled audio waveform (WAV).

To play the sound, the process is reversed. The sound board sends the digital data through a digital-to-analog converter (DAC), then uses its amplifier to play the sound through the speakers.

What is audio resolution?

As weve mentioned, you use the New File Data Format dialog box to configure the audio resolution of a recording. The settings in this box determine the quality of the sound and the format of the resulting digital data stored on your system. Well continue using our snapshot analogy and explain each of the audio resolution terms.

Channels

Mono is one channel of sound, and stereo is two separate channels (left and right). Consequently, to play and record stereo sound, the sound board must have a snapshot of the sound for each channel. This means that instead of playing or recording one snapshot, the board plays or records one snapshot for each channel at the same time. Consequently, this doubles the amount of digital data Windows stores on your system.

Sample size

The sample size determines the *dynamic range*, or loudness, of the sound. Although 16 bits is twice as much data as 8 bits, its not twice as loud. Instead, you can think of 8-bit and 16-bit sample sizes as two volume knobs. The 16-bit knob goes from 0 to 16 and the 8-bit knob goes from 0 to 8.

Although a setting of 0 represents no sound for both sample sizes, the highest settings, 16 and 8 respectively, produce the loudest. The difference is the 16-bit sample size has a greater range of adjustment, or dynamic range, between 1 and 16.

Also, 8-bit sampled sound has more noise, making recordings sound scratchy. Although 8-bit sound is fine for voice recordings, 16-bit sampled sound is better for music. As with stereo sound, 16-bit sampled sound requires twice as much storage space on your system.

Frequency (sampling rate)

The number of snapshots the sound board takes per second is known as the *sampling rate* or *frequency*. Just as more frames per second make the motion in movies more fluid, a higher sampling rate provides a richer sound because there is more sound.

The quality of sound found on music CDs is 16-bit stereo sound sampled at 44.1KHz. However, the MPC specification only requires that the sound board record 8-bit mono sound sampled at 11KHz and play at 8-bit mono sampled at both 11KHz and 22KHz.

Incidentally, when you play music CDs, the CD-ROM player handles all the digital-to-analog conversion. Consequently, the sound board only amplifies the sound from the music CDs. Furthermore, the music plays with the same full CD quality sound of a Compact Disc player.

Storage space for sampled WAV

You might wonder why the minimum specification doesnt call for the full blown 16-bit stereo sound sampled at 44.1KHz. One reason is the data needed to process sound at that quality requires a system that can process and store it.

One minute of 8-bit mono sound sampled at 11KHz requires approximately 660K of storage space. Sampling at 44.1KHz quadruples this storage requirement to 2.6MB. Using 16 bits doubles the data requirement to 5.3MB. Stereo sampling doubles it again to 10.5MB.

Not only does storage become a problem at this resolution, but the system must process 10.5MB of audio data each minute just to produce the sound. Consequently, the upper limit on audio is more appropriate for 8-bit mono sound at 22KHz requiring 1.3MB of storage space for one minute of sound. This tremendous amount of sound data demonstrates why a CD-ROM drive is such a critical part of the MPC.

The MIDI synthesizer

The MPC specification requires that the sound board provide a MIDI in port, MIDI out port, and an internal MIDI synthesizer. Furthermore, the synthesizer must be capable of playing three different melodic voices with 6-note polyphony and two different percussive voices with 2-note polyphony.

Voices refer to individual instruments such as guitar, trumpet, or flute. *Polyphony* refers to the maximum number of simultaneous notes that can be played. For example, three different voices with 6-note polyphony might mean that the trumpet is playing one note, the flute one note, and the guitar the remaining four notes at the same time. The same follows for the percussion instruments except there are only two notes. For example, there might be a cymbal and drum each playing one note. The MPC specification also has a high end synthesizer that calls for nine different melodic voices with 16-note polyphony, and eight different percussive voices with 16-note polyphony.

To play music, MIDI uses a *channel number* from 1 to 16 for each voice, melodic or percussive, played on the synthesizer. This channel number receives the MIDI *note on* and *note off* messages for that voice. Each note on or note off message includes the pitch and loudness of the note. When you play MIDI files, Windows sends MIDI data that contain up to 16 channels of note on and note off messages to the sound boards synthesizer. Unlike sampled WAV audio, a minute of music driven by MIDI data requires only a few kilobytes of storage space.

Although the synthesizer can use only a few voices simultaneously, some sound boards have hundreds of different voices available for playing music. Furthermore, some sound boards use FM synthesis to generate the sounds of the different voices, whereas others use 16-bit sampled sounds of real instruments.

On the other hand, you can use the MIDI in and out ports to connect external MIDI devices such as keyboards, drum machines, and sound modules. This allows Multimedia Windows to record and play music with any external MIDI device.

Conclusion

Sound is the major component of Multimedia Windows. However, when using the new features that utilize sound, you may encounter unfamiliar terminology and options. In this article, weve explained the basics of sound in Multimedia Windows.

Reprinted with permission from The Cobb Group's *Inside Microsoft Windows* journal ©1992. Call 1-800-223-8720 for a free issue.