PRIMER ON PC AUDIO

This document provides an introduction to various issues associated with PC-based audio technology. Topics include the following:

- Introduction to PC-Based Audio
- Introduction to Audio Files
- WAV Files
- Comparison Chart of WAV-File Formats
- MP3, WMA, and Ogg Vorbis Files
- Audio CDs
- Transfer of Audio from LPs and Cassettes to CDs

More detailed information on MP3 files, WMA files, and Ogg Vorbis files can be found on High Criteria's web site (www.highcriteria.com).

Introduction to PC-Based Audio

Today, almost every PC comes equipped with a sound card and associated speakers. The speakers and sound card, in conjunction with the computer operating system, facilitate the playing of audio sounds (for example, music, speech, bird calls, and so on) so the end user can hear the sounds.

The sounds created by the speakers are actually waves of compressed air. The human ear is capable of distinguishing these waves, provided that their frequency resides within the audible spectrum. By their nature, these sound waves are continuous. Each wave has a discrete value of pressure associated with it and each of these possible values generates a different sound.

Before a computer can deal with sounds, the sounds must be converted to a series of ones and zeroes. This process is called digitalization. To understand how digitalization works, let us imagine a microphone plugged into a computer sound card. The membrane on the mic moves back and forth according to the current air pressure. The microphone circuitry transforms this movement into a continuous and real-valued electric signal. The sound card in turn has a circuitry called an Analog-to-Digital converter or ADC. The ADC can convert this electric signal to a sequence of ones and zeroes. This conversion is done by taking periodic measurements of the electrical signal, and converting it to an integer value. The frequency at which the measurements are taken is called a sample rate, and the size of the integer (in bits) is called a sample size. The result of each measurement is called a sample.

The more samples per second, the greater the accuracy that will be achieved in reproducing the sounds when they are replayed. The sample rate is measured in the number of thousands of samples per second and is expressed as a specific thousand of samples per cycle (second). A sample rate of 44.1 kHz is used for most CDs; 48 kHz is used for DVD. Some proposed future audio devices could go as high as 96 kHz. A sample rate of 8.0 kHz yields less than one-fifth the samples of CD-quality recording, and it is more suitable for dictation-type applications than for recording music. Of course, a high sample rate means more data must be saved, and the file size and disk space requirements increase accordingly.

The sample size indicates the size of the scale that will be used for each sample measurement and will usually be either 8 bits or 16 bits. A scale between 0 and 255 can be represented in 8 bits; a scale between 0 and 65,535 can be represented with 16 bits. Therefore, a 16-bit sample size provides much greater granularity and can represent the correct pressure much more accurately than an 8-bit sample size. CD-quality recordings are made with a sample size of 16 bits, and recordings such as dictations are made with a sample size of 8 bits. Some newer audio devices are capable of working with a sample size of 24 bits.

The last fundamental characteristic of digital sound is the number of channels. To reflect the spatial properties of sound, it must be measured simultaneously in more than one place. For example, two microphones are generally needed to obtain stereo sound. Each independent measurement point produces one channel. In the vast majority of cases, the number of channels is either one (mono sound) or two (stereo sound). New systems, such as Dolby 5.1, can use up to 6 channels.

To summarize, each digital sound has three fundamental characteristics: the sample rate, the sample size, and the number of channels.

Compression

Even in today's world of high-speed Internet connections and 80-GB hard drives; audio files cannot be regarded as tiny or small. Competing compression methods have been developed to address various requirements—compressed music and compressed speech being two common examples. Each compression method has its own advantages and disadvantages. One comparison often made is the size of a resulting file weighed against the quality of the reproduced sound. Since the quality of the reproduced sound is most often a very personal judgment, you must always keep in mind that your own experience with the various compression methods should be the most useful guide in assessing which compression method to employ.

Compression methods can be broadly categorized as either lossless or lossy. With lossless compression, the sounds are reproduced exactly as they were originally heard. Because reproduction is exact, lossless compression methods cannot compress a file to the same degree that a lossy method can. In a Windows environment, the majority of compression methods use lossy techniques.

Lossy compression techniques actually eliminate some portion of the digital data stream that represents the sounds. The amount and type of data removed depends on the purpose for which the compression method was developed. For example, a particular compression method may have been developed to allow for the playing of popular music. Such a method would attempt to reproduce the music with a great deal of quality. It may try to remove only the portions of the sounds that are near or beyond the hearing thresholds of most individuals. Using this approach, only a small difference in quality can be detected by most people.

A different compression method may be designed to record spoken words. With this type of compression, a much larger quantity of the sound spectrum may be discarded, because it is not relevant for the recording of speech. In this scenario, a much greater level of compression can often be achieved.

In addition to the sample rate, the sample size and the number of channels, the compressed sound has the characteristic called a bit rate. The bit rate tells how many bits per unit of time are needed to store the sound. It is usually measured in Kbps (kilobits per second). For uncompressed sound, the bit rate is always the product of the sample rate times the sample size times the number of channels. For compressed sound, it is lower than the above product and the difference can be viewed as the degree of compression.

Storing the Sound

Digital sound can be stored in many forms. Just to name a few:

- Audio file on a computer's hard-drive
- Audio CD
- Mini-disk
- Digital-audio tape.

In the case of PC-based digital audio, saving to a computer's hard drive is the most important format. The source for PC audio will almost always be an audio file, of which there are numerous types and formats. Many software programs have been developed to play these audio files. The remainder of this document provides users with a basic understanding of some of the audio-file formats and some of the associated audio software.

Introduction to Audio Files

Audio files that are recognized and played on a PC come in many different formats. Different file formats are usually associated with different file extensions. For example, just as an MS Word file might be named filename.doc (with a .doc extension), one format of audio file is a wav file and might be named filename.wav. Other audio file formats include MP3 files (filename.mp3), Ogg Vorbis files (filename.ogg), Real Audio files (filename.ra), and many more.

All of these formats are digital representations of the original sounds. At some point, the sound in these files has been converted from a continuous "analog" signal that may have come from a microphone, record player, or other similar device, to a digital format. This conversion is done with circuitry usually found on the sound card within a PC. This circuitry is called an analog to digital converter (ADC). Similar circuitry is also used to convert the sounds back from digital to analog and is called a digital to analog converter (DAC). The DAC circuit is used to convert a digital file so it can be played back through the speakers on a PC.

As previously discussed, most digital audio files share several attributes that describe the properties of the recorded sound. These properties include sample rate, sample size, number of channels, bit rate, and the compression technique used.

WAV Files

WAV files are the most common standard for recording audio in a Windows environment. However, with the demand for various forms of audio increasing exponentially, many competing formats are available. Commonly, WAV files can be identified by their .wav extension. Within the WAV format, many different compression methods can be used. A number of these compression methods are available as a standard part of the Windows operating system.

Compression CODECs

In a Windows operating system environment, methods for audio compression can be implemented in a special program called a CODEC. There are a number of CODECs that come as a standard part of the various Windows operating systems. CODEC-compression programs commonly come in files with an .acm extension. For example, tssoft32.acm is the CODEC file for a compression algorithm called DSP Group Truespeech.

If a CODEC is installed properly, Windows facilitates the use of the CODEC by any audio program running on the PC. The programs may use the CODEC to encode (record) or decode (play) audio files. The compression method that was used to compress an audio file can be identified with information stored within the audio file. Saving this information allows the correct CODEC to be selected later when the file is decoded for playback.

A number of the CODECs that come as a standard part of the Windows operating system are included for compatibility reasons. Their inclusion allows for interoperability between the Windows environment and other specialized audio systems. Additionally, a number of more general purpose CODECs are included. A subjective overview of these general purpose CODECs is provided below. Users should remember that these explanations are provided on an "as is" basis. If there is any doubt, you should review the available compression methods and make your own judgment about which one is best suited for your application.

Comparison Chart of WAV-File Formats*

The following chart compares different WAV-file formats:

Format	Attributes	MB per Hour	Application
DSP Truespeech	8.0 kHz, 1 bit, mono	4	Dictation
L&H SBC	16 Kbps, 8.0 kHz, 16 bit, mono	9	High-quality dictation or phone calls
MS ADPCM	8.0 kHz, 4 bit, mono	14	Talk Radio and meetings
MP3 11.5kHz,	16 Kbps, mono	4.5	Low-quality Internet music
PCM 44.1kHz,	16 bit, stereo	600	High-quality recordings

* Note that dlls (programs) are available that will create high-quality MP3 files. Refer to the MP3 page on the High Criteria web site for more information.

DSP Group Truespeech

The DSP Group Truespeech CODEC was written by DSP Group of Santa Clara California. This compression method was written specifically to address the requirement of recording human speech. This method removes a considerable portion of the potential sound spectrum. However, the data removed has little impact on a listener's ability to understand what was said. The algorithm "rounds off" many of the highs and lows associated with the tones found in the original spoken words. Listeners may no longer detect some of the emotions that may have been inferred by these tones, but the actual words will remain quite clear and distinguishable.

This CODEC is an excellent choice for recording dictation. The CODEC supports a sample rate of 8.0 kHz with a sample size of 8 bits and mono recording. One hour of Truespeech recording can fit into about 4.5 MB of disk space. We suggest the use of Truespeech for dictation-type applications where the speaker's emotions are not relevant.

4

Lernout & Hauspie

There are several Lernout & Hauspie CODECs included with Windows. All of these CODECs have a small sample rate and were written specifically to address the requirement of recording human speech.

The CODECs have a sample rate of 8.0 kHz with a sample size of 16 bits. All of the CODECs are mono. The higher sample size provides additional tone to these recordings but will increase the file size. A file recorded with the Lernout & Hauspie SBC 16-Kbps CODEC requires less than 9 MB of disk space for a one-hour recording. We suggest the use of a Lernout and Hauspie CODEC for recording applications that require a level of quality beyond that provided by Truespeech. Recording a telephone conversation is a good example for using this CODEC.

Microsoft ADPCM

The Microsoft ADPCM CODEC provides a middle ground for audio recording. This CODEC retains much of the audio spectrum while at the same time providing a good level of compression. The algorithm reproduces most of the tones that were present in the original spoken words. We suggest the Microsoft ADPCM CODEC be used for high-quality telephone recording and for recording meetings or conversations where capturing the emotions behind the words is also important. This CODEC is ideal for talk show Internet broadcasts.

The Microsoft ADPCM CODEC offers sample rates from 8.0 kHz to 44.1 kHz. All sample sizes are 8 bits, and the sounds can be captured in either mono or stereo.

MPEG Layer - 3

On some versions of Windows, Microsoft provides an MPEG Layer - 3 CODEC to create files compressed into MP3 format. This CODEC supports sample rates from 8.0 kHz to 24.0 kHz with a sample size of 16 bits in both mono and stereo. This CODEC is applicable to low-quality music recordings. Refer to the MP3 page on the High Criteria web site for more information.

PCM

PCM is a completely uncompressed sound format. Because it is uncompressed, there is no loss of quality due to the deletion of data. The most widely distributed PCM CODEC has sample rates from 8.0 kHz up to 48.0 kHz, a sample size of both 8 and 16 bits, and supports both mono and stereo.

This format is best used when the size of the file is not an issue. This would be the case if the file were to be quickly moved to a CD, for example.

MP3, WMA, and Ogg Vorbis Files

For detailed information on these types of files and how High Criteria's software works with such files, refer to the following URLs:

Type	URL
MP3	http://www.highcriteria.com/tr_ext_mp3.htm
WMA	http://www.highcriteria.com/tr_ext_wma.htm
Ogg Vorbis	http://www.highcriteria.com/tr_ext_ogg.htm

Audio CDs

The audio found on CDs played in home and car stereos is not in a file format that is recognized by Windows. Windows Explorer can detect .cda files on this type of CD. However, these .cda files are not audio files, and you cannot use Windows Explorer to copy the audio portions. These types of CDs are sometimes called CDA CDs, CD Audio CDs, or Red Book Audio CDs (since the original specification was published in a red book).

Users wanting to create their own CDA CDs must have a CD burner to write the CDs, software to control the CD burner, and a source for the audio they want to place onto the CD. Usually, the software to control the CD burner comes packaged with the burner. The audio source is usually a wav file found on the PC's hard drive.

When writing to a CD, the software program that controls the burner asks the user if he wants a music CD or a data CD. If the user chooses a music CD, then the software converts the source files to CDA format, and the resulting CD can be played in a home or car stereo. If the user specifies that he wants a data CD, then the program will not convert the files, but will instead copy them as a .wav file, .mp3 file, or whatever file type was specified. Generally, the data type of CD cannot be played on a standard home stereo system. One exception to this is if an MP3 file is copied to CD and played on a player that supports the MP3 format.

High Criteria develops a number of software programs that are designed to record audio into wav or mp3 files. These files can then be moved onto a CDA CD using the software that comes with the CD burner. High Criteria does not supply software to write wav files onto a CD. If you have problems moving your audio files onto a CD, please consult the documentation associated with your CD burner and the software that came with the burner.

Transfer of Audio from LPs and Cassettes to CDs

The transfer of recordings from tapes, records, or other media to CD is a two-step process. The first step is to get the recording to a PC file in a format the PC understands. The second step is to copy the PC file to a CD. Our software, either Total Recorder or Dictation Buddy, addresses the first step. You will also need a cable with audio plugs at both ends. This cable will go from the headphone/earphone jack of your stereo or cassette player to the mic or line-in jack of your computer's sound card. You should be able to get such a cable for a few dollars from Radio Shack or other similar store. Make sure you get plugs that fit the jacks at both ends!

Once you have the cable, use Total Recorder or Dictation Buddy to record from the stereo/cassette player to the PC. If you have Total Recorder, go to "Recording Source and Parameters" and change the setting from "Software" to "Sound Board". You will then be able to record whatever is played into the mic or line-in jack from the output of the stereo/cassette player. Our programs record into a WAV file—the industry standard for PC-based audio recordings.

If you don't already have a CD burner, then you must get one. There are many different burners on the market from many manufacturers. It is best to source this from your local computer supplier.

You will get software with the burner that allows you to copy .wav files over to a CD. The important issue in this step is to decide upon the CD's format. The burner software will ask if you want a "data"-format CD or a "music"-format CD. If you specify data, then you will get a .wav file on the CD that can only be played by a PC. If you specify "music", then the .wav file will be converted to a format that can be played by a standard home/car CD player. This will be a CDA (CD Audio) CD. See the section above on Audio CDs for more information about this type of CD.

Generally, when a "music" CDA CD is created, the software that comes with the burner creates separate "tracks" for each .wav file that is copied onto the CD. So, if you want your home stereo to recognize the different songs on a CD as separate tracks, each of the songs must be separated into different .wav files before they are copied to the CD.

Note: Almost all commercial recordings are protected by copyright. In the past, most jurisdictions have allowed a user to make a small number of copies of copyrighted material, provided that the copies are created solely for use by the original purchaser. If you have any questions about the legality of creating a copy, we suggest you contact an authority on the copyright regulations in your jurisdiction.