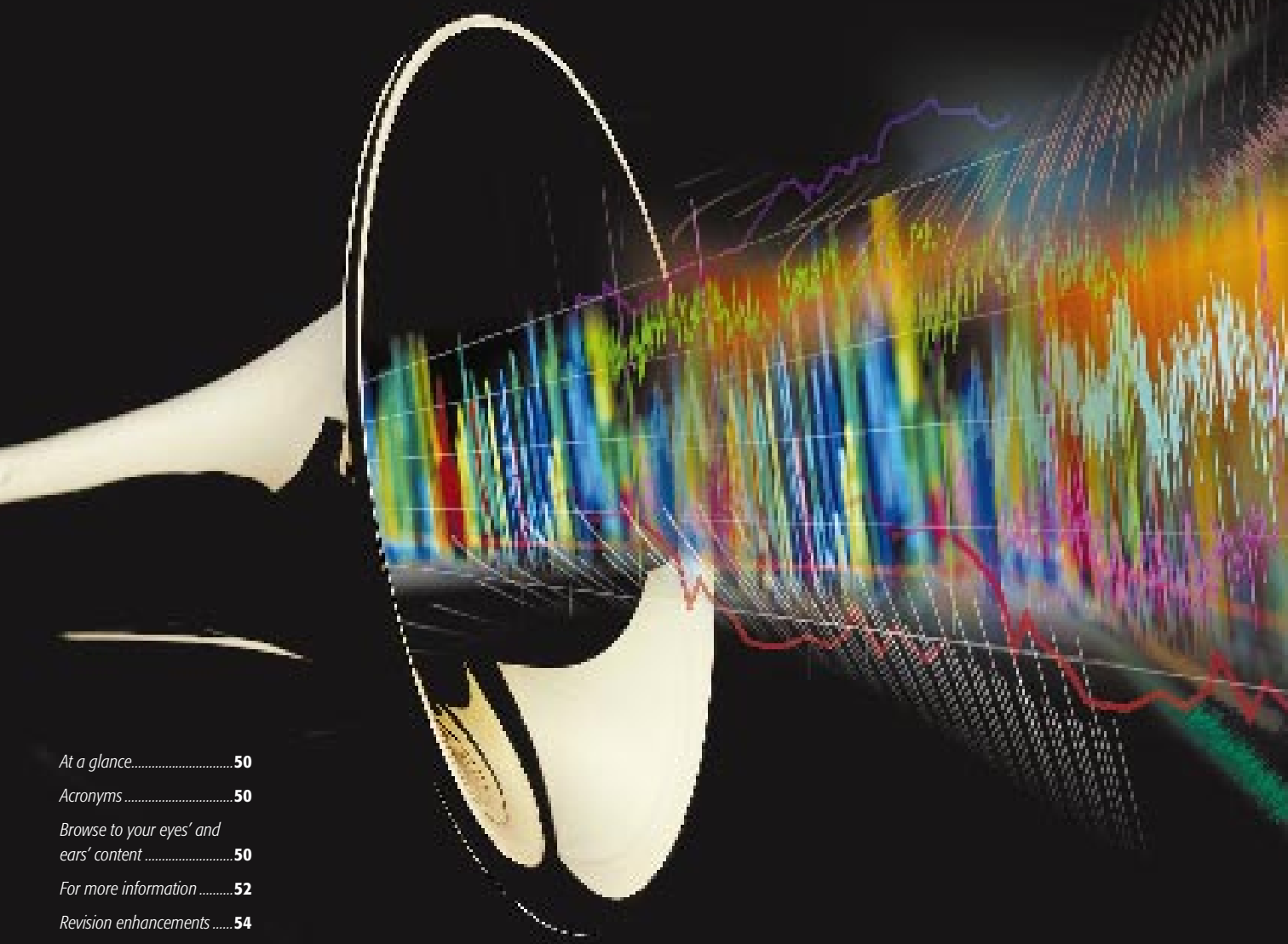
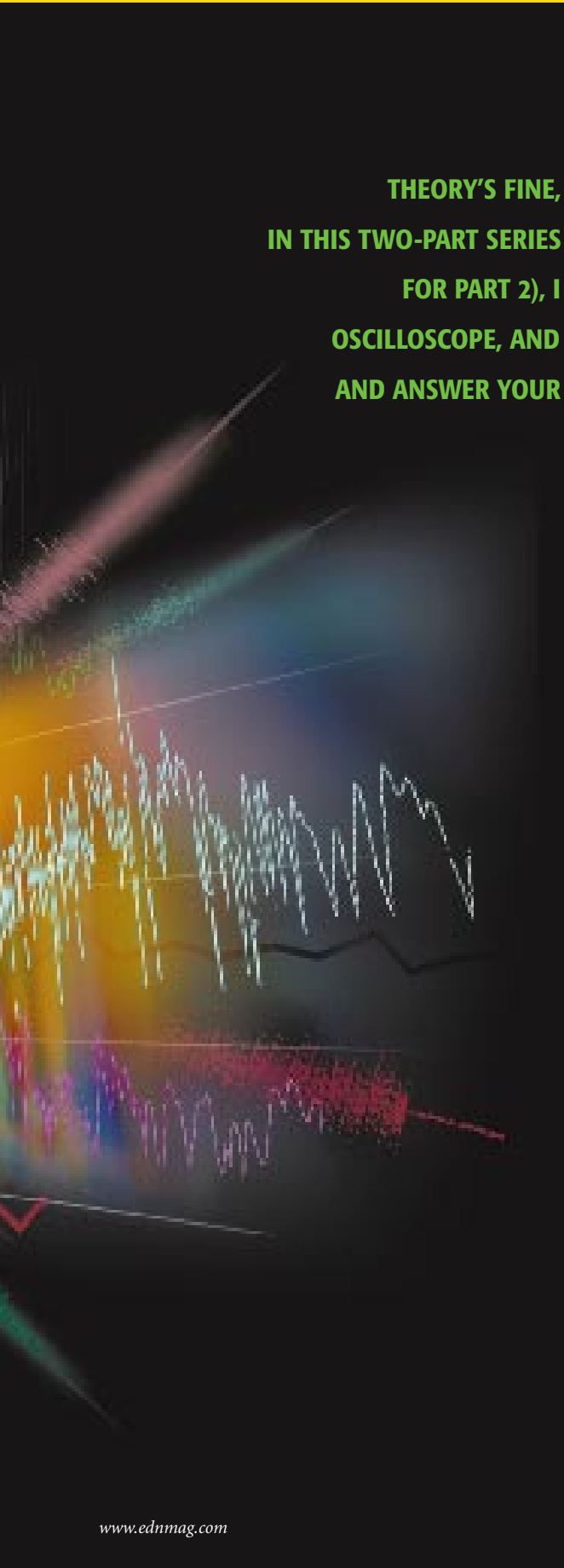


DIGITAL AUDIO GETS AN AUDITION

PART ONE: LOSSLESS COMPRESSION



At a glance.....	50
Acronyms	50
Browse to your eyes' and ears' content	50
For more information	52
Revision enhancements	54



**THEORY'S FINE, BUT REALITY'S BETTER.
IN THIS TWO-PART SERIES (SEE OUR JAN 18 ISSUE
FOR PART 2), I DUST OFF MY PC, STOPWATCH,
OSCILLOSCOPE, AND SPECTRUM ANALYZER
AND ANSWER YOUR AUDIO-COMPRESSION QUESTIONS.**

LAST YEAR, I wrote about the theory behind digital-audio compression, basing my research mostly on second-hand anecdotes, technical papers, and marketing collateral (**references 1 to 4**). Theory is all well and good, but as I'm sure you've already discovered in your careers, even the best academic analysis doesn't always match real-life results. I prefer to come to my own conclusions, after running my own experiments, so I jumped at the chance to work on this article series. I'll be putting a number of lossless and lossy codecs under the magnifying glass (as well as my through headphones and speakers) to see how they perform.

AT A GLANCE

- ▶ Real-life data confirms (or contradicts) theoretically derived predictions.
- ▶ In a world seemingly obsessed with lossy compression, lossless codecs also have a place.
- ▶ CPUs' multimedia-tuned features can be as important as clock speed in determining performance, but only if the software exploits them.
- ▶ Generic file-compression software is practical for use with only the most random of sound clips.
- ▶ Differential predictive algorithms deliver greater compression and higher performance with most multimedia content.
- ▶ Supposedly significant differences in algorithm complexity don't lead to corresponding variances in speed or compression ratio.

Part 2 of this article series will appear in *EDN's* Jan 18, 2001, issue. Additional subsequent work will appear on *EDN's* Web site (see **sidebar** "Browse to your eyes' and ears' content").

When people seem to want to talk only about MP3 and other lossy codecs, why should you care about lossless compression

(**Table 1**)? Well, people who tape live performances are going to want to archive their recordings in the best-quality format, while still shrinking the file sizes to as small as possible. Remember that an uncompressed dual-channel, 16-bit, 44.1-kHz-sampled audio stream gobbles up nearly 11 Mbytes of storage space per minute of music. Even if lossless compression shrinks the file by only one-third or one-half, that size reduction is better than nothing. For similar reasons, audiophiles prefer to listen to their multichannel, surround-sound, extended-precision, and high-sampling-rate music in a lossless format, such as the MLP used in DVD Audio, rather than a lossy-compression technique, such as DTS or Dolby Digital (**Reference 5**).

Lossless compression may also make sense for musicians and audio engineers who want to temporarily store tracks prior to editing and mixing and archive the final master tracks for future fame and fortune, or folks who decide to digitize those LP records collecting dust in their garages. Once you convert a sound clip to a lossy format, you cannot recover the information that the encoder discarded. And just as with repeated saving of JPEG still images, every time you modify and resave an MP3 or other lossy-compression audio file—or transcode it from one lossy format to another—the audio quality further degrades.

Storing digital audio clips as compressed data files instead of burning

ACRONYMS

AAC: Advanced Audio Coding

codec: compressor/decompressor, also sometimes used to define a single-chip A/D-plus-D/A converter.

DAE: digital audio extraction

DTS: Digital Theater Systems

EBU: European Broadcast Union

FDD: floppy disk drive

HDD: hard disk drive

LPAC: Lossless Predictive Audio Compression

MIPS: million instructions per second

MLP: Meridian Lossless Packing

RIMM: Rambus in-line memory module

SIMD: single instruction/multiple data

SPS: Sound Processing Software

SQAM: Sound Quality Assessment Material

BROWSE TO YOUR EYES' AND EARS' CONTENT

For additional information, visit the Web-site addendum to this article series at www.ednmag.com/ednmag/extras/01csaddendum.asp. Readers of this article's upcoming part two may notice that I haven't exactly been fair to MP3. Due to page-count limitations, I've shown and discussed only the 64-kbit "fastest encode" results, not the "highest quality" outputs, which theoretically might not exhibit the same type or extent of artifacts. Also, for each codec, a whole range of higher bit rates exist for which I

have not provided frequency- and time-based snapshots. The Web site contains all of these graphics, including electronic versions of the figures from part two of the article.

The Web-site addendum also provides downloadable WAV source files for the first 10 test tones shown in **Table 3**. I'll also post some of the lossy-compressed file versions of the tones that I generated. Copyright restrictions preclude me from posting **Table 2's** song clips, as well as the EBU SQAM CD

tracks. But the Massachusetts Institute of Technology (Cambridge, MA) has obtained permission to post some of the EBU SQAM tracks, so the Web-site addendum will include a link to MIT's Web site. Also included are links to other sources of useful audio files, comprising both synthetic test patterns and real music and voice, that you might find useful in your own analyses, as well as other lossless and lossy compression studies that I came across in my research.

As time allows, I plan to work

on additional analyses. I'd like to retest MP3 with other encoder and decoder combinations. I'd also like to look more closely at the three lossy codecs that part two of this article describes for evidence of other compression artifacts and to compile additional quantifiable comparison data. And plenty of other codecs await inspection, chief among them is AAC. (See **sidebar** "The never-ending sonic story" in this article's part two in *EDN's* Jan 18, 2001, issue for details.)

them onto an audio CD has another key advantage: There's no need for a subsequent DAE process to copy the information from the audio CD. DAE is notorious for injecting clicks, pops, stutters, blank spots, and other irregularities into the WAV file created during the extraction process. These flaws result from less-than-perfect laser-head positioning, defects in the audio-CD media, and programs running in the background that periodically distract the PC from its primary extraction task. If the lossless codec you choose contains built-in playback capability, or if the vendor supplies a plug-in for Nullsoft's WinAmp or other popular audio-playback software, your customers can enjoy tunes without needing to first decompress the files.

For the lossless-compression analysis, I wanted to measure several key parameters. First, I wanted to determine how long each algorithm takes to both encode

(compress) WAV files and decode (decompress) back to WAV files. I also wanted to compare the compressed file sizes with the originals to determine each algorithm's compression efficiency. Certainly, I needed to compare each original WAV file to the one that the encode-and-decode process created to ensure that the

based on the percentage of CPU resources they consumed during encoding and decoding, but I found that they all took whatever spare MIPS capability was available during at least a portion of their compression and decompression routines.

Lossless compression results depend

TABLE 1—REPRESENTATIVE LOSSLESS COMPRESSION ALGORITHMS

Codec	Version	URL
DAKX		www.dakx.com
LPAC		www-ft.ee.tu-berlin.de/~liebchen/lpac.html
LTAC		www-ft.ee.tu-berlin.de/~liebchen/ltac.html
MKW		http://home.att.net/~mkw
Monkey's Audio		www.monkeysaudio.com
*MusiCompress (WaveZip)	2.01	http://members.aol.com/sndspace ; www.gadgetlabs.com/wavezip.htm
Perfect Clarity Audio		www.sonicfoundry.com
*RAR	2.71	www.rarsoft.com
RKAU		http://rksoft.virtualave.net
*Shorten	2.3a1	www.softsound.com/Shorten.html
SPS		www.krishnasoft.com/sps.htm ; www.pegasusimaging.com/sound.html
Waveform Archiver		www.simtel.net/pub/simtelnet/msdos/arcers/wavarc10.zip
*WavPack	3.2	www.wavpack.com
*WinZip	8.0	www.winzip.com
Zap		www.emagic.de

* Algorithms that I tested.

algorithm was indeed lossless. (Fortunately, in all cases, it was.) I also originally hoped to differentiate the algorithms

greatly on the characteristics of the source material: for example, total frequency range that the material spans; the

FOR MORE INFORMATION...

For more information on products such as those discussed in this article, go to our information-request page at www.rscanners.ims.ca/ednmag/. When you contact any of the following manufacturers directly, please let them know you read about their products in *EDN*.

RARSoft

www.rarsoft.com
Enter No. 301

SoftSound Ltd

+44 1223 421754
www.softsound.com
Enter No. 302

SoundSpace Audio

1-408-221-1191
<http://members.aol.com/sndspace>
Enter No. 303

WinZip Computing

1-612-253-8488
www.winzip.com
Enter No. 304

OTHER COMPANIES MENTIONED IN THIS ARTICLE

Adaptec

www.adaptec.com

Belkin

www.belkin.com

Cirque

www.cirque.com

Digigram

www.digigram.com

DTS (Digital Theater Systems)

www.dtsonline.com

Dolby Labs

www.dolby.com

Ego-Sys

www.egosys.net

High Criteria

www.highcriteria.com

Hitachi

www.hitachi.com

Intel

www.intel.com

Iomega

www.iomega.com

Kingston Technology

www.kingston.com

Linksys

www.linksys.com

M-Audio (Midiman)

www.m-audio.com

Maxtor

www.maxtor.com

Microsoft

www.microsoft.com

NEC

www.nec.com

Nullsoft

www.nullsoft.com

Nvidia

www.nvidia.com

Plextor

www.plextor.com

Sonic Foundry

www.sonicfoundry.com

Syntrillium Software

www.syntrillium.com

Zoltrix

www.zoltrix.com

SUPER INFO NUMBER

For more information on the products available from all of the vendors listed in this box, enter No. 305 at www.rscanners.ims.ca/ednmag/.

amount of sample-to-sample frequency variation; whether the material is monophonic or stereo; and, if the material is stereo, the difference in the channels' fre-

quency, amplitude, and phase. **Table 2** includes a variety of music genres to throw at the codecs. I chose not only modern, digitally captured music but also older,

analog-recorded, as well as live-performed, tracks, which potentially exhibit worst-case channel-to-channel phase differences.

(continued on pg 58)

TABLE 2—MUSIC SAMPLES

Genre	Artist	Song	Album	Duration (minutes: seconds. thousandths of a second)	Uncompressed size (bytes)
Adult contemporary	Fleetwood Mac	The Chain	Rumours	4:31.173	47,835,020
Blues	ZZ Top	Blue Jean Blues	Fandango!	4:45.493	50,361,068
Classical (instrumental)	Academy of St Martin-in-the-Fields, conducted by Neville Marriner	Mozart's Piano Concerto in D minor, 2nd movement	Amadeus soundtrack	9:52.467	104,511,164
Classical (with vocals)	Academy of St Martin-in-the-Fields, conducted by Neville Marriner	Mozart's Requiem (Lacrymosa)	Amadeus soundtrack	3:49.933	40,560,284
Country	Grateful Dead	Ripple	American Beauty	4:09.707	44,048,300
Folk	Cat Stevens	The Wind	Teaser and the Firecat	1:42.733	18,122,204
Hard rock (instrumental)	Joe Satriani	Surfing With the Alien	Surfing With the Alien	4:24.827	46,715,468
Hard rock (with vocals)	Motley Crue	Wild Side	Girls, Girls, Girls	4:40.000	49,392,044
Instrumental (piano)	George Winston	Carol of the Bells	December	4:00.240	42,378,380
Jazz (classic)	Miles Davis	So What	Kind of Blue	9:25.427	99,741,308
Jazz (contemporary)	Steely Dan	Bodhisattva	Countdown to Ecstasy	5:19.000	56,271,644
Live (classic rock)	Peter Dinklage	Do You Feel Like We Do?	Frampton Comes Alive!	13:45.000	145,530,044
Live (contemporary rock)	Phish	Bouncing Around the Room	A Live One	4:08.440	43,824,860
Modern rock	Nine Inch Nails	Into the Void	The Fragile	4:49.307	51,033,740
New age	Trillium Green	Liquid Moon	Metamorphoses	4:28.960	47,444,588
Rap	Beastie Boys	The Negotiation Limerick File	Hello Nasty	2:46.200	29,317,724
Techno	Propellerheads	Take California	Decksanddrumsand rockandroll	7:21.333	77,851,244
Vocals (choral)	Coro de monjes del Monasterio Benedictino de Santo Domingo de Silos	Puer Natus Est Nobis	Canto Gregoriano	3:41.000	38,984,444
Vocals (solo)	Suzanne Vega	Tom's Diner	Solitude Standing	2:09.107	22,774,460

Notes: All music clips are 44.1-kHz-sampled, 16-bit stereo.

REVISION ENHANCEMENTS

At this point, I've been working on this project for nearly six months. My lossless compression study took place back in the late summer of 2000, after which I shared my results with vendors. David Bryant, developer of the WavPack algorithm, cleared up some of my post-analysis confusion (and tempted me to do even more analysis) when he told me that the compressors I tested actually are similar, despite vendors' claims. However, he advised me that two of the compressors that I didn't test (LPAC and RKAU) are

different and produce better compression albeit at slower speeds. Unfortunately, "much" better compression is maybe only a 5% improvement or a little more.

Although I tested the latest available version of WavPack, Bryant has been hard at work improving his product. Several of the other algorithm developers also follow an aggressive updating schedule. This rapid-revision phenomenon was largely why, for example, I chose to not include in the study the popular Monkey's Audio algorithm.

When software goes through multiple revisions in a short time, analysis results become obsolete long before they can appear in print.

Bryant claims that if I had tested the latest version of WavPack (3.6 beta), the results would have had more variety. According to Bryant, the new WavPack is a 32-bit program, so it executes much faster than the version I tested, and it now contains a new "high"-compression option (-h) that gives compression ratios much closer to the best programs while still being

reasonably fast. Bryant says he achieves this high performance-with-high compression combination by including more samples in the predictor, but he still retains an all-integer approach.

For readers who would like to test new versions of WavPack, any of the other algorithms I evaluated, or codecs that I didn't have time to benchmark, most of the test files are available at *EDN's* Web site, and **Table 2** lists the compressed audio-CD song files. I welcome your feedback.

TABLE 3—TEST TONES

Tone	Description	Duration (minutes: seconds. thousandths of a second)	Uncompressed size (bytes)
1	Pink noise; both channels at approximately -12-dB average rms power; 44.1-kHz-sampled, 16-bit stereo	00:30.000	5,292,044
2	Pink noise; left channel at approximately -12-dB average rms power; right channel at approximately -32-dB average rms power 44.1-kHz-sampled, 16-bit stereo	00:30.000	5,292,044
3	White noise; both channels at approximately -9.6-dB average rms power; 44.1-kHz-sampled, 16-bit stereo	00:30.000	5,292,044
4	White noise; left channel at approximately -9.6-dB average rms power; right channel at approximately -29.6-dB average rms power; 44.1-kHz-sampled, 16-bit stereo	00:30.000	5,292,044
5	Left and right channels 180° out of phase; all frequencies present over entire time interval; 44.1-kHz-sampled, 16-bit stereo; the following half-band frequencies are at -10-dB combined peak amplitude (both channels): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz	00:30.000	5,292,044
6	Left and right channels 180° out of phase; all frequencies present over entire time interval; 44.1-kHz-sampled, 16-bit stereo; the following half-band frequencies are at -10-dB combined peak amplitude (left channel) and -30-dB combined peak amplitude (right channel): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz	00:30.000	5,292,044
7	Frequency masking test; left and right channels 180° out of phase; all frequencies present over entire time interval; 44.1-kHz-sampled, 16-bit stereo; the following half-band frequencies are at -10-dB combined peak amplitude (both channels): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz; the following quarter-band frequencies are at -30-dB combined peak amplitude (both channels): 25, 125, 225, 325, 427.5, 540, 665, 807.5, 960, 1127.5, 1322.5, 1540, 1790, 2080, 2415, 2812.5, 3287.5, 3875, 4625, 5575, 6725, 8150, 10125, 12875, 17137.5 Hz; the following three-quarter-band frequencies are at -30-dB combined peak amplitude (both channels): 75, 175, 275, 375, 482.5, 600, 735, 882.5, 1040, 1222.5, 1427.5, 1660, 1930, 2240, 2605, 3037.5, 3562.5, 4225, 5075, 6125, 7375, 9050, 11375, 14625, 20412.5 Hz	00:30.000	5,292,044
8	Frequency masking test; left and right channels 180° out of phase; all frequencies present over entire time interval; 44.1-kHz-sampled, 16-bit stereo; the following half-band frequencies are at -10-dB combined peak amplitude (left channel) and -30-dB combined peak amplitude (right channel): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz; the following quarter-band frequencies are at -30-dB combined peak amplitude (left channel) and -50-dB combined peak amplitude (right channel): 25, 125, 225, 325, 427.5, 540, 665, 807.5, 960, 1127.5, 1322.5, 1540, 1790, 2080, 2415, 2812.5, 3287.5, 3875, 4625, 5575, 6725, 8150, 10125, 12875, 17137.5 Hz; the following three-quarter-band frequencies are at -30-dB combined peak amplitude (left channel) and -50-dB combined peak amplitude (right channel): 75, 175, 275, 375, 482.5, 600, 735, 882.5, 1040, 1222.5, 1427.5, 1660, 1930, 2240, 2605, 3037.5, 3562.5, 4225, 5075, 6125, 7375, 9050, 11375, 14625, 20412.5 Hz	00:30.000	5,292,044
9	Temporal masking test; left and right channels 180° out of phase; 44.1-kHz-sampled, 16-bit stereo; over first 50-msec interval, the following half-band frequencies are at -30-dB combined peak amplitude (both channels): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz; over subsequent 30-sec interval, the following half-band frequencies are at -10-dB combined peak amplitude (both channels): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz; over final 200-msec interval, the following half-band frequencies are at -30-dB combined peak amplitude (both channels): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz	00:30.250	5,336,384
10	Temporal masking test; left and right channels 180° out of phase; 44.1 kHz sampled, 16-bit stereo; over first 50-msec interval, the following half-band frequencies are at -30-dB combined peak amplitude (left channel) and -50-dB combined peak amplitude (right channel): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz; over subsequent 30-sec interval, the following half-band frequencies are at -10-dB combined peak amplitude (left channel) and -30-dB combined peak amplitude (right channel): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz; over final 200-msec interval, the following half-band frequencies are at -30-dB combined peak amplitude (left channel) and -50-dB combined peak amplitude (right channel): 50, 150, 250, 350, 455, 570, 700, 845, 1000, 1175, 1375, 1600, 1860, 2160, 2510, 2925, 3425, 4050, 4850, 5850, 7050, 8600, 10750, 13750, 18775 Hz	00:30.250	5,336,384
11	EBU SQAM track 27 (Castanets) 44.1-kHz-sampled, 16-bit stereo	00:20.000	3,528,044
12	EBU SQAM track 32 (Triangle) 44.1-kHz-sampled, 16-bit stereo	01:14.000	13,053,644
13	EBU SQAM track 35 (Glockenspiel) 44.1-kHz-sampled, 16-bit stereo	00:59.000	10,407,644

To convert most of the audio tracks to WAV files, I used an enhanced version of Windows 98's CDFX.VXD driver, which automatically displays the tracks in Windows Explorer as stereo and mono, 11-, 22.05-, and 44.1-kHz-sampled data files (Reference 6). I had only to copy them to the hard drive. I could have also used specialized audio-extraction utilities, such as Exact Audio Copy or Plextor Manager (references 7 and 8). The Beastie Boys hybrid CD contained both audio tracks and PC software and was, therefore, unusable with CDFX.VXD; I extracted the song using the WAV conversion utility built into Music Match Jukebox. Note that all the music tracks are stereo; one common method of cheating that some codec developers use in benchmarking against their competition involves employing low-fidelity mono, less random versions of a sound clip for themselves and a "harder" high-fidelity stereo version for the alternatives.

To fully understand the codecs' capabilities and gain some insight into how they work, I used both "real" music clips and test tones. I generated the first 10 files that Table 3 shows using Syntrillium Software's Cool Edit Pro. White noise is a random audio pattern spanning the entire frequency spectrum, representing all frequencies in equal proportions. Pink noise's sample proportions, in contrast, follow a $1/\text{frequency}$ pattern that places equivalent audio energy in each frequency octave. In other words, low frequencies are logarithmically more likely to exist in each sample than high frequencies. The pink-noise

**WHITE NOISE IS A
RANDOM AUDIO PATTERN
SPANNING THE ENTIRE
FREQUENCY SPECTRUM,
REPRESENTING ALL
FREQUENCIES IN EQUAL
PROPORTIONS.**



pattern more closely matches the human-auditory-system response, giving pink noise a more natural sound, with less high-frequency "hiss" than white noise. In contrast, brown-noise proportions follow a $1/(\text{frequency})^2$ ratio, giving brown noise an audio presentation dominated by bass much more so than its colorful counterparts. Part two of this article discusses the remainder of Table 3's test tones.

Keep in mind when you analyze the re-

sults that the PC I used has a fast and full-featured microprocessor inside (Table 4). If a given format's encoders and decoders take advantage of the integer and floating-point SIMD instructions that the Pentium III supports, they might run faster on my PC than their nonenhanced counterparts; the differences would be less significant on a more conventional CPU. Other important CPU-related questions include whether the algorithms fit within the L1 or L2 caches and whether they use streaming data instructions to prevent cache corruption. With 256 Mbytes of Direct Rambus DRAM, the PC should have sufficient available memory for even the largest music clip. And, although some level of system overhead is inevitable when reading and writing large audio files, the PC's 7200 RPM, Ultra ATA/66 hard drive should keep this overhead to a minimum.

One other note about my PC processor: As a Katmai-generation CPU (the 533B), it's different from the latest-and-greatest Coppermine-generation Pentium IIIs (such as the 533EB) coming out of Intel's fabs. Both processors employ a 32-kbyte (16-kbyte code, 16-kbyte data) internal L1 cache, and both use a 133-MHz system bus. My 533B, however, includes a 512-kbyte half-speed external L2 cache; the 533EB and its Coppermine companions employ a smaller (256-kbyte) but faster (full-speed) integrated L2 cache, with more advanced features. Coppermine-era processors also beef up the system buffering to increase bus usage: They include four write-back buffers, six fill buffers, and eight bus-queue buffers. My 533B should still run rings around a 533 (Pentium II core) or 533A (Pentium III Katmai core) Celeron CPU, which have 66-MHz system buses and even smaller 128-kbyte integrated L2 caches.

My PC ran Microsoft's Windows 98 Second Edition. Before encoding or decoding, I used the task list (accessed via a CNTL-ALT-DEL keystroke combination) to terminate all running applications except Explorer and Systray. Eliminating all unnecessary background programs pushed available system resources to greater than 95% and ensured that random interruptions wouldn't un-

TABLE 4—SYSTEM SPECIFICATIONS

Hardware:

Intel VC820 motherboard with built-in Creative Labs Soundblaster audio; Intel 533-MHz (Katmai, 133-MHz front-side bus) CPU; two 128-Mbyte (Kingston Technology) PC800 RIMMs; Maxtor Fireball Plus KX 3.5-in., 27.3-Gbyte, 7200 RPM Ultra ATA/66 HDD; Nvidia GeForce 256 reference AGP graphics card; Linksys EtherFast 10/100 Ethernet PCI card; Adaptec 2906 SCSI PCI card; Hitachi GD-2500 ATAPI DVD-ROM drive; Plextor 12/10/32 ATAPI CD-RW drive; Iomega Zip 250 SCSI drive; 3.5-in. FDD; NEC MultiSync E1100 21-in. CRT; Cirque ergonomic keyboard and Logitech First Mouse Wheel; generic ATX midtower enclosure with 250W power supply; Belkin OmniView Pro keyboard, video, and mouse switch.

Software:

Microsoft Windows 98 Second Edition; Sonic Foundry Sound Forge 4.5h; Syntrillium Software Cool Edit Pro 1.2a

Future hardware and software (digital-I/O audio-subsystem options):

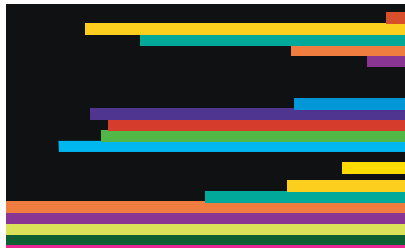
Digigram VXpocket; Ego-Sys Waveterminal U2A; High Criteria Total Recorder 2.2; M-Audio Delta DiO 2496; Zoltrix Nightingale.

fairly handicap one codec and not another.

RESULTS

I first decided to run each of the WAV files through the ubiquitous (at least for Windows-based PCs) WinZip program. I did two sets of tests; the first at the program’s “fastest speed” setting, then at its “highest compression” configuration. I’d previously heard that the compression algorithms that generic file-compression utilities, such as PKZip, WinRAR, and WinZip use, are not optimal for low-entropy multimedia files, and the results showed that this opinion is correct (Table 5). Table 5 in the version of this article on EDN’s Web site contains results for all 19 music genres shown in Table 2, as well as for pink and white noise test tones 1 through 4 in Table 3. Even with advanced warning, I was shocked by how much poorer the WinZip compression ratios were than the ratios of the other programs and how long the compression

I WAS SHOCKED BY HOW MUCH POORER THE WINZIP COMPRESSION RATIOS WERE THAN THE RATIOS OF THE OTHER PROGRAMS.



and decompression routines took to complete.

I also found it ironic that WinZip’s “highest compression” option—although it took noticeably longer to

execute than its “fastest speed” counterpart—didn’t achieve proportionally better compression success. WinRAR, a peer of WinZip, is unique among generic file-compression programs in offering a multimedia-optimized compression option. And, as the name promises, WinRAR’s multimedia mode delivered significantly better compression than WinZip. However, on average, WinRAR took roughly twice as long to compress a WAV file than WinZip at its slowest setting.

Now for the audio-exclusive compression algorithms. In comparing the data for MuziCompress, Shorten, and WavPack, you might be struck by the similarity of the results. This seeming near-parity is ironic because, if you believe the programs’ documentation, the algorithms vary widely in their complexity. Literature from Soundspace Audio, for example, touts MuziCompress as a lean, mean, integer-only routine employing a relatively simple predictive algorithm. Shorten, at the opposite end of the spec-

TABLE 5—LOSSLESS COMPRESSION RESULTS

Codec	Classical (instrumental)				Hard rock (with vocals)				Compressed file size (bytes)
	Compressed file size (bytes)	Compressed file size (% of original)	Compression time (minutes: seconds. thousandths of a second)	Decompression time (minutes: seconds. thousandths of a second)	Compressed file size (bytes)	Compressed file size (% of original)	Compression time (minutes: seconds. thousandths of a second)	Decompression time (minutes: seconds. thousandths of a second)	
MuziCompress (WaveZip)	41,208,646	39	00:34	00:28	38,355,946	78	00:19	00:16	5,222,716
Shorten	38,266,837	37	00:33	00:18	37,950,193	77	00:18	00:16	4,990,237
WavPack (fast option)	39,919,927	38	00:30.38	00:26.20	38,175,231	77	00:16.75	00:16.32	5,431,833
WavPack	39,821,655	38	00:33.56	00:32.09	37,645,751	76	00:19.78	00:18.84	5,159,333
WinRAR (multimedia option)	46,929,337	45	02:03	01:02	36,905,064	75	01:00	00:30	5,175,091
WinZip (fastest speed)	86,462,730	83	00:50	00:31	46,601,198	94	00:24	00:15	5,164,004
WinZip (highest compression)	85,698,366	82	01:10	00:31	46,504,885	94	00:27	00:16	5,159,247

Notes: File sizes determined via Windows Explorer; compression percentages manually calculated (data reported by software is sometimes incorrect).

trum, is a more elaborate and highly configurable floating-point-based scheme. In fairness to Shorten, note that I did not experiment with altering its default settings. As they say in college textbooks, “This exercise is left to the interested student” (see sidebar “Revision enhancements”).

It’s possible that the hard-drive read-and-write delays normalized the results, although my gut feeling is that HDD overhead was only a secondary factor. Also, remember that in this study I’m using a high-speed, robust-featured microprocessor. A simpler, or slower, CPU, such as those in many embedded applications, might produce greater discrepancies among the competing alternatives. For similar reasons, I decided *not* to use my other PC, which has an even faster PIII 800 CPU and a dual-Rambus-channel i840 chip set. A 533-MHz CPU-based system represents a more mainstream configuration, particularly when you consider not only the PCs that are sell-

ing today but also those PCs that are already in users’ hands. In this era of 1.5-GHz Pentium 4’s, it’s sometimes easy to forget that not too long ago, 533 MHz represented the state of the art.

Not surprisingly, the compression routines targeting audio struggled the most with the random-white-noise and pseudorandom-pink-noise sound files, achieving limited compression success comparable with that of WinZip, which targets more random, generic data. When you examine Table 5 on EDN’s Web site, compare the compression-ratio differences that exist when right and left channels have equivalent intensity with the ratio when one channel is “louder” than the other. Also, note that the supposedly compressed WavPack “fast-option” white-noise equivalent-channel file is actually larger than the original WAV clip. The fact that a “compressed” version of a highly random data pattern is larger than the pattern itself isn’t a surprise; this phenomenon results

from the control bit and other overhead of the compression scheme. But most compression algorithms, in such a case, automatically use the original file instead.

Different types of music, even with similar time duration, achieve widely differing compression ratios and speeds. Instrumental-only classical data compresses to less than half the original file size, but more modern music types, such as hard rock, rap, and techno, are nearly as random and difficult to compress as pink and white noise. Perhaps our parents were right, after all, when they claimed that rock and roll was nothing but noise. MusiCompress was unable to handle the Beastie Boys rap WAV file I generated using MusicMatch Jukebox. I’ve heard that some “ripping” software packages produce nonstandard WAV files that other programs sometimes find difficult to read and play. I believe that this scenario occurred here.

Stay tuned for part 2 of my hands-on project in the Jan 18 issue of EDN. □

You can reach
Technical Editor
Brian Dipert at
1-916-454-5242,
fax 1-916-454-5101,
e-mail bdipert@pacbell.net



White noise (equivalent channels)

Compressed file size (% of original)	Compression time (minutes:seconds.thousandths of a second)	Decompression time (minutes:seconds.thousandths of a second)	*Notes
99	00:01	00:01	Compression and decompression times obtained from stopwatch
94	00:01	00:01	Block size (256), mean estimation (4) and predictor order (0) left at defaults; compression and decompression times obtained from stopwatch
103	00:01.37	00:01.43	Compression and decompression times reported by software
97	00:01.81	00:01.81	Compression and decompression times reported by software
98	00:16	00:02	Use multimedia compression archiving option enabled, nonmultimedia material compression method left at default normal; compression and decompression times reported by software
98	00:02	00:02	Super-fast compression setting; compression and decompression times obtained from stopwatch
97	00:03	00:02	Maximum (slowest) compression setting; compression and decompression times obtained from stopwatch

* See expanded version of this table on EDN’s Web site at www.ednmag.com/ednmag/reg/2001/01042001/01cs.htm.

REFERENCES

- Dipert, Brian, “Now hear this,” *EDN*, Feb 3, 2000, pg 50.
- Dipert, Brian, “Digital audio breaks the sound barrier,” *EDN*, July 20, 2000, pg 71.
- Dipert, Brian, “Security scheme doesn’t hold water (marking),” *EDN*, December 21, 2000, pg 35.
- Dipert, Brian, “Listen up,” *CommVerge*, January 2000, pg 46.
- Partkya, Jeff, “Stop, hey, what’s that sound,” *EMedia Magazine*, January 1999, pg 42.
- Starrett, Robert A, “Ripping Off Recordings: Digital Audio Extraction Do’s, Don’ts, and Do-ers,” *EMedia Magazine*, July 1999, pg 34.
- Dipert, Brian, “Software delivers the sound of savings,” *EDN*, Sept 1, 2000, pg 24.
- Dipert, Brian, “CD-RW drive, tutorial prevent a ‘slow burn,’” *EDN*, Oct 26, 2000, pg 34.