

RE: compressing mp3 files

Source:

<http://www.tech-archive.net/Archive/WinXP/microsoft.public.windowsxp.music/2005-12/msg00096.html>

- *From:* marcinont <marcinont@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Wed, 21 Dec 2005 10:41:16 -0800
-

What?

"Bhavesh" wrote:

- > MPEG standards in general and MP3 in particular
- >
- > MPEG-1 Layer 3 (known as "MP3") is most widespread and popular today. It has
- > won its popularity quite deservedly – it is the first widespread lossy-codec
- > which reached such a high data compression factor, together with very good
- > sounding quality. A little bit of history. MPEG is an abbreviation of "Moving
- > Pictures Coding Experts Group". MPEG has been started at January, 1988. Since
- > the first assembly in May, 1988, the group began to grow, and has grown up to
- > unusual dense experts collectively. Usually, in MPEG assembly about 350
- > experts participate, from more than 200 companies. The largest part of
- > participants are the experts occupied in various scientific and academic
- > establishments. Today MPEG group has developed the following standards and
- > algorithms:
- >
- > MPEG-1 (November 1992) – the standard of coding, storage and decoding of
- > moving pictures and audio data;
- > MPEG-2 (November 1994) – the standard of data coding for digital TV;
- > MPEG-4 – the standard for multimedia applications;
- > MPEG-7 – universal standard for multimedia, intended for processing,
- > filtration and management of multimedia data.
- > Let us consider the set of standards MPEG-1. This set, according to ISO
- > standards (International Standards Organization), includes three algorithms
- > of different levels of complexity: Layer 1, Layer 2 and Layer 3. Our well
- > known friend MP3 in exact designation is "MPEG-1 Layer 3". The general
- > structure of encoding process is identical in all Layers. At the same time,
- > in spite of similarity of the Layers in the general approach to encoding, the
- > Layers differ on target use and internal mechanisms. By the way, this fact
- > determines the degree of similarity of the algorithms which have "grown" from
- > MPEG-1 (such as, Ogg Vorbis and MusePack). Each Layer has its own format of
- > data stream and decoding algorithm. MPEG-1 algorithms are mainly based on
- > known properties of perception of sound signals by a hearing aid of human (we
- > have mentioned above about these techniques).
- >
- > Briefly about encoding algorithm used in MPEG-1. At the beginning of

RE: compressing mp3 files

- > encoding, the source audio stream with the help of filters is divided on
- > bandwidth. The continuation of the encoding process depends on used Layer.
- >
- > In the case of Layer 3 (MP3) the signal in each obtained bandwidth is
- > decomposed on frequency components by applying MDCT (Modified Discrete Cosine
- > Transform – a special case of Fourier Transform) that gives a set of
- > coefficients. Further processing is focused on simplification of the signal
- > in order to perform re-quantization of its spectral coefficients. Obtained
- > spectrum is cleared (by filtering) of obviously inaudible components –
- > low-frequency noise and high imperceptible spectrum components. At the next
- > stage, considerably more complex psycho acoustic analysis is applied (as was
- > described earlier) on the audible part of spectrum. After all these
- > manipulations, the source signal is deprived of more than half of its
- > information. In completion of all, compression of obtained stream by the
- > simplified analogue of Huffman algorithm is performed (this is lossless
- > compression method), that allows to reduce noticeably the stream size.
- >
- > In the case of Layer 2 the simplification process is quite similar. The
- > difference consists in the object of re-quantization: re-quantization is
- > performed on amplitude signal in each sub-band and not on the spectrum
- > coefficients (some non-MP3 lossy encoders are based on the same technique).
- >
- > Complete set MPEG-1 is intended for coding signals with sample rates of 32,
- > 44.1 and 48 kHz. Three MPEG-1 Layers that were mentioned above have
- > distinctions in encoding mechanisms and, thus, they provide different
- > compression factors and sounding quality of resulting streams. Layer 1 allows
- > keeping signals in format 44.1 KHz / 16 bits without significant losses of
- > quality at bitrate of 384 Kbps that gives 4 times profit of data size. Layer
- > 2 provides, subjectively, the same quality at 192 – 224 Kbps, when Layer III
- > (MP3) gives the same results at 128–160 Kbps. It is impossible to speak about
- > advantages and disadvantages of one Layer compared to another, because each
- > Layer is developed to achieve its own aim. For example, the advantage of
- > Layer 3 actually consists in allowing of data compression 8–12 times
- > (depending on bitrate) without significant losses of original sound quality.
- > At the same time, speed of a compression provided by this Layer is the lowest
- > (it is necessary to note, that on modern CPU's this restriction is not
- > appreciable at all). Layer II is potentially capable to provide higher
- > quality of coding on account of "easier" internal signal processing during
- > transformation. However, Layer II does not allow to reach so high compression
- > factors, which may be reached by using Layer III.
- >
- > Nuances of coding
- >
- > The technique of audio coding is complex enough and has a set of nuances.
- > All of them cannot be explained within the framework of one article; however
- > all the most important should be considered, as almost every user meets with
- > them when encoding.
- >
- > Data encoding into MP3 (as well as into WMA and OGG) is performed by blocks:
- > the coded file is divided on so-called frames of a certain equal length and
- > each frame is encoded separately and is stored in a target stream. Thus, the

RE: compressing mp3 files

- > target stream also has frame structure. Each frame can be encoded not on any
- > bitrate, but only on one of those included in the standard table for MPEG1
- > Layer 3 (Kbps): 32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256, 320
- > (coding on intermediate bitrates is not stipulated by the standard, though it
- > is possible). Because each frame is processed individually, it is possible to
- > speak about data compression with constant (CBR) and variable (VBR) bitrate.
- >
- > CBR (Constant Bitrate) is a way of encoding when all frames are encoded on
- > identical bitrate. In other words, bitrate of the whole encoded stream
- > remains constant all along the stream.
- >
- > VBR (Variable Bit Rate) is a way of encoding when each separate frame is
- > encoded with its own bitrate, calculated by encoder. The choice of bitrate
- > for each frame is performed by the encoder according to performed psycho
- > acoustic analysis.
- >
- > There is also one more encoding mode – ABR (Average bitrate). Encoding in
- > this mode (it is true, at least, for MP3 coders) is similar to CBR encoding.
- > However this encoding is performed on variable bitrate keeping the same
- > average. Not going into technical details, we shall note that VBR and ABR
- > encoding is much more flexible and, often, more favorable and qualitative,
- > rather than in CBR mode.
- >
- > It is important to note, that ABR, VBR and CBR modes are used also in many
- > coders rather than MP3.
- >
- > We shall consider now existing encoding techniques of stereo data stipulated
- > in MPEG–1 Layer 1, 2, 3 standards. These methods, probably, with some
- > different interpretations, are valid not only in MPEG, but also in other
- > codecs.
- >
- > Dual Channel. This mode is intended for encoding of audio information in two
- > channels as absolutely independent. In other words, encoding of audio occurs
- > separately in each channel without tracking dependence of a signal in
- > channels. As is implied from the name, this mode is mainly intended for
- > coding of data with two parallel independent channels (for example, speech in
- > English and German languages), and NOT with two channels carrying stereo
- > information of sounding. In general, this mode is not recommended to be used
- > for coding of stereo signal.
- > Stereo. This mode differs from the Dual Stereo mode in reservoir usage.
- > Reservoir – is a mechanism that is responsible for assignment of bits for
- > encoded frames in the target stream. During encoding in stereo mode both
- > channels are processed using the same reservoir, when in Dual Stereo mode,
- > the signal is encoded, using independent reservoir for each channel. There
- > are no other differences between the modes.
- > Joint Stereo is common definition of the encoding methods of stereo
- > information, which are based on the use of its redundancy. There are two
- > versions of this method described in MPEG–1.
- > MS Stereo. In this mode the encoded signal is re–divided on a middle channel
- > (common constituent for both right and left channels) and a side channel
- > (differentiated constituent of the channels) and processed as in Stereo mode,

