

Putting The Squeeze On Video

An introduction to video compression.

Text:/Rod Sommerich & Andy Ciddor

THE PURPOSE OF COMPRESSION is to reduce the size of a message without losing its meaning. The idea has been around forever in such applications as smoke signals, Morse code and shorthand. The idea behind simple methods of compression is to make commonly used information shorter than the normal form. In Morse code, the most common letters use the shortest codes (E = one dot, T = one Dash, A=Dot Dash) while the less common characters use the long codes (Z=Dash Dash Dot Dot). While this technique is quite efficient for compressing textual material (in European languages) it is ineffective for more complex messages such as images and sounds. A more effective approach is to look for repeated patterns in the message (common phrases, repeated pixel patterns) and to replace the second and subsequent appearances of the pattern with a short code. This technique lies at the core of such familiar compression formats as, FLAC, TAR, WAV, ZIP, RAR and Stuffit, and is very effective with many types of data.

All of these techniques make the data smaller to reduce storage space and transmission bandwidth and most importantly, when the data is uncompressed what comes out is identical to what went into the compression process. Unfortunately these so-called 'lossless' compression methods are only capable of relatively minor levels of compression, and thus far have proved to be unsuitable for storing video and audio on available media or transmitting them over the relatively narrow bandwidth that has hitherto been available (i.e. before RuddNet/NBN).

DISCARD PILES

To achieve the required levels of compression for practical storage and transmission, the compression techniques developed have all involved discarding part of the original data. The trick to the various 'lossy' compression methods has been to work out just how much of what kind of information can be discarded before the result is either too noticeable (or more than the end user will put up with) in exchange for the convenience of carrying around their media library or receiving device.

There are two types of compression that

can be applied to a video stream. Intraframe compression attempts to minimise the amount of information required to represent the contents of a single frame of the video stream and is essentially the same process used to compress still images. Interframe compression attempts to minimise the amount of information required to represent the contents of series of frames, usually by discarding the information that remains unchanged between successive frames in the stream.

A 'codec' (a contraction of the term 'coder-decoder') is the software that takes a raw data file and encodes it into a compressed file, then attempts to recreate the original information from the compressed data. Because compressed files only contain some of the data found in the original file, the codec is the necessary 'processor' that decides what data makes it in to the compressed version, what data gets discarded and what gets transformed into a more compact format. Different codecs for the same compression standard may encode in slightly different ways, so a video file compressed using Intel's codec will be different from a file compressed using the Cinepak codec. Sometimes the difference is noticeable on replay, sometimes not, but it's good to be aware of a codec's strengths and matching them to what you're trying to do, in order to maintain the best balance between file size and quality.

STANDARDS - LEFT TO EXPERTS

The two most widely used video compression types are standardised and managed by organisations comprised of representatives from academic research organisations and industry. The Moving Picture Experts Group (MPEG) and the Joint Photographic Experts Group (JPEG) represent two different approaches to the video compression process, although some aspects of each set of standards benefit from the work of both groups.

The standards have continued to develop as computing power and transmission bandwidth has increased. As more local processing power became available through faster and more advanced processor chips, it became possible to require more complex calculations in the

decoding (replay) process. For example, MPEG4-encoded files tend to be smaller but require more processor activity to replay than an MPEG1 video file of the same quality. The MPEG1 file will be larger, requiring more storage space and bandwidth to move the file, but will replay on a device with less computational power.

JPEG - OLD KID ON THE BLOCK

JPEG is an intraframe compression standard that was developed for still images around 1992. When method began to be applied to video streams, a variation named Motion JPEG was developed. Both standards work by dividing the frame into squares or blocks of pixels (usually 16 per block) and calculating the average value of each block. The smaller the sample block, the better the quality of the image and the larger the size of the file. If the block size is too large, then the level of detail is poor and the image can begin to look like a group of squares. Everyone who uses a computer is familiar with the effect of over-enthusiastic use of JPEG compression.

MPEG - IN THE FRAME

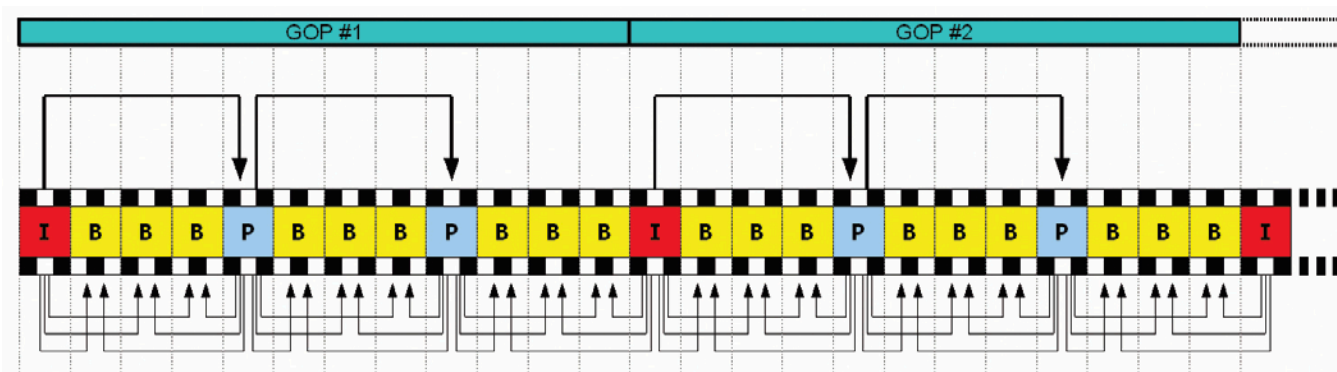
The various versions of MPEG employ both interframe and intraframe compression techniques. Its interframe process uses three types of frames.

Intra Frames (I-Frames): The I-Frame is the reference point for the frames that follow. It's a complete independent frame of information, compressed using only intraframe methods.

Predicted Frames (P-Frames): A P-Frame is consists only of the differences between the current frame and its reference I-Frame or the previous P-Frame, a process known as forward prediction. P-Frames can also serve as a prediction reference for B-Frames and future P-Frames. A P-Frame is much more compressed than an I-Frame.

Bidirectional Frames (B-Frames): Sometimes called Back Frames, B-Frames are derived using bi-directional prediction, which uses both a past and a future frame as references, only storing the differences between them. B-Frames show the highest level of compression.

The I, B and P Frames from a single shot form a block of information known as a Group



The I-Frame is the complete image that begins a Group Of Pictures. A Predictive P-Frame is derived from the difference between itself and the I-Frame. Back/Bidirectional B-Frames are derived from both the I-Frame and the P-Frame.



Acronym	Description	Application
M-JPEG	Intra Frame Compression	Video Editing
MPEG-1	up to about 1.5Mbit/s released circa 1993	VCD, Streaming
MP3	Common name for MPEG1-Layer3 audio compression	Audio Players
MPEG-2	Generic coding of moving pictures and audio	DVD, Satellite, HDTV
MPEG-3	Not completed and was incorporated in to MP2	Not Used
MPEG-4	Video Audio & DRM (Incl H.264)	Streaming, Video Conf, HD-DVD, BluRay

of Pictures(GOP). A GOP can vary in length from three frames to several seconds.

MPEG has many different variations, each one evolved to facilitate a specific application such as security, HD video, content protection, or multi-channel audio and video. The one which has made the biggest impact recently has been the MPEG4 Advanced Video Codec (AVC) implementation (also know as H.264), which, because of the highly flexible and extensible nature of the scheme, can be used on everything from mobile devices through to very high-resolution video imaging. Its scalability in compression and processing power has revolutionised the delivery of video in media as diverse as the Blu-Ray disc, YouTube, Digital Video Broadcasting and streaming video delivery over the 3G mobile network.

TURN UP THE SOUND

In addition to compressing the video stream we need similarly efficient systems to process the audio component of the program. The very different process of compressing audio will be the subject of other articles from other authors. To keep the audio and video components together requires encapsulating them in a container system that ensures that the components arrive together and remain synchronised when viewed.

Some compression systems such as MPEG have preferred containers such as MPG or MP4 for MPEG or Digital Video standard (often called DV), used in

camcorders. Other non-proprietary envelopes include AVI, MKV, TS or VOB, and 3GP. Other envelopes are created by special interest groups who wish to emphasise specific features of the compression to make the media more suited to their requirements. Some of these include FLV (Flash Video), Windows Media, RealMedia, DivX and XVID which have specific features that the creators want to access. These formats can change without notice and without explanation as the owners of the standards develop their requirements. This can be frustrating for users.

In the mid-to-late 1990s watching videos and listening to music online was like dealing with peak hour traffic, a little bit at a time.... If you had a slow internet connection, you could spend more time staring at the hourglass or a status bar than watching the video or listening to a song. Everything was choppy, pixilated and hard to see. Streaming video and audio have come a long way since then. Tens of millions of people watch and listen to Internet video and radio every day. In addition to this, YouTube is seeing 20 hours of video uploaded per minute... this is up from six hours per minute in 2007. Compression is making all this possible and accessible to anyone with a computer. 🖱

References:
The wondrous Wikipedia.com.
bbc.co.uk