About MPEG Compression

HD video requires significantly more data than SD video. A single HD video frame can require up to six times more data than an SD frame. To record such large images with such a low data rate, HDV uses long-GOP MPEG compression. MPEG compression reduces the data rate by removing redundant visual information, both on a per-frame basis and also across multiple frames.

Note: HDV specifically employs MPEG-2 compression, but the concepts of long-GOP and I-frame-only compression discussed below apply to all versions of the MPEG standard: MPEG-1, MPEG-2, and MPEG-4 (including AVC/H.264). For the purposes of this general explanation, the term *MPEG* may refer to any of these formats.

Spatial Compression

Within a single frame, areas of similar color and texture can be coded with fewer bits than the original frame, thus reducing the data rate with a minimal loss in noticeable visual quality. JPEG compression works in a similar way to compress still images. Spatial, or *intraframe*, compression is used to create standalone video frames called *I-frames* (short for *intraframe*).

Temporal Compression

Instead of storing complete frames, temporal (*interframe*) compression stores only what has changed from one frame to the next, which dramatically reduces the amount of data that needs to be stored while still achieving high-quality images. Video is stored in three types of frames: a standalone I-frame that contains a complete image, and then predictive *P-frames* and bipredictive *B-frames* that store subsequent changes in the image. Every half second or so, a new I-frame is introduced to provide a complete image on which subsequent P- and B-frames are based. Together, a group of I-, P-, and B-frames is called a *group of pictures*, or *GOP*. HDV uses a long-GOP pattern, which means that there is at least one P- or B-frame for each I-frame.

For example, suppose you record some typical "talking head" footage, such as an interview in which a seated person moves very little throughout the shot. Most of the person's body stays still, so most of the visual information is stored in an I-frame; the subsequent P- and B-frames store only the changes from one frame to the next.

Because P- and B-frames depend on other frames to create a meaningful image, your computer spends more processing power decoding HDV frames for display than it does when displaying intraframe-only formats such as DV, uncompressed video, or the Apple Intermediate Codec.

More About Long-GOP Video

The term *long* refers to the fact that P- and B-frames are used between I-frame intervals. At the other end of the spectrum, the opposite of long-GOP MPEG is I-frame-only MPEG, in which only I-frames are used. Formats such as IMX use I-frame-only MPEG, which reduces temporal artifacts and improves editing performance. However, I-frame-only formats have a significantly higher data rate because each frame must store enough data to be completely self-contained. Therefore, although the decoding demands on your computer are decreased, there is a greater demand for scratch disk speed and capacity.

1080-line HDV media uses an *open GOP* structure, which means that B-frames in the MPEG stream can be reliant on frames in adjacent GOPs. 720-line HDV media uses a *closed GOP* structure, which means that each GOP is self-contained and does not rely on frames outside the GOP.

Transcoding HDV to Other Apple Codecs

Instead of working with native MPEG-2 HDV video, you can transcode your HDV video to an Apple ProRes codec or the Apple Intermediate Codec during capture. For more information about these codecs, see Transcoded HDV Editing Workflow, Working with Apple ProRes, and About the Apple Intermediate Codec.

Unlike MPEG-2 HDV, these Apple codecs do not use temporal compression, so every frame can be decoded and displayed immediately, without first decoding other frames.

You can also capture and edit native HDV but render your footage using an Apple ProRes codec. For more information, see Stage 4: Choosing a Render File Format for HDV Sequences in Native HDV Editing Workflow.

About MPEG Compression

MPEG encoding is based on eliminating redundant video information, not only within a frame but over a period of time. In a shot where there is little motion, such as an interview, most of the video content does not change from frame to frame, and MPEG encoding can compress the video by a huge ratio with little or no perceptible quality loss.

MPEG compression reduces video data rates in two ways:

- *Spatial (intraframe) compression:* Compresses individual frames.
- *Temporal (interframe) compression:* Compresses groups of frames together by eliminating redundant visual data across multiple frames.

Intraframe Compression

Within a single frame, areas of similar color and texture can be coded with fewer bits than the original, thus reducing the data rate with minimal loss in noticeable visual quality. JPEG compression works in a similar way to compress still images. Intraframe compression is used to create standalone video frames called *I-frames* (short for *intraframe*).

Interframe Compression

Instead of storing complete frames, temporal compression stores only what has changed from one frame to the next, which dramatically reduces the amount of data that needs to be stored while still achieving high-quality images.

Groups of Pictures

MPEG formats use three types of compressed frames, organized in a *group of pictures*, or *GOP*, to achieve interframe compression:

- *I-frames:* Intra (I) frames, also known as *reference* or *key frames*, contain all the necessary data to re-create a complete image. An I-frame stands by itself without requiring data from other frames in the GOP. Every GOP contains one I-frame, although it does not have to be the first frame of the GOP. I-frames are the largest type of MPEG frame, but they are faster to decompress than other kinds of MPEG frames.
- *P-frames:* Predicted (P) frames are encoded from a "predicted" picture based on the closest preceding I- or P-frame. P-frames are also known as *reference frames*, because neighboring B- and P-frames can refer to them. P-frames are typically much smaller than I-frames.
- *B-frames:* Bi-directional (B) frames are encoded based on an interpolation from I- and P-frames that come before and after them. B-frames require very little space, but they can take longer to decompress because they are reliant on frames that may be reliant on other frames. A GOP can begin with a B-frame, but it cannot end with one.

GOPs are defined by three factors: their pattern of I-, P-, and B-frames, their length, and whether the GOP is "open" or "closed."

GOP Pattern

A GOP pattern is defined by the ratio of P- to B-frames within a GOP. Common patterns used for DVD are IBP and IBBP. All three frame types do not have to be used in a pattern. For example, an IP pattern can be used. IBP and IBBP GOP patterns, in conjunction with longer GOP lengths, encode video very efficiently. Smaller GOP patterns with shorter GOP lengths work better with video that has quick movements, but they don't compress the data rate as much.

Some encoders can force I-frames to be added sporadically throughout a stream's GOPs. These I-frames can be placed manually during editing or automatically by an encoder detecting abrupt visual changes such as cuts, transitions, and fast camera movements.

GOP Length

Longer GOP lengths encode video more efficiently by reducing the number of I-frames but are less desirable during short-duration effects such as fast transitions or quick camera pans. MPEG video may be classified as *long-GOP* or *short-GOP*. The term *long-GOP* refers to the fact that several P- and B-frames are used between I-frame intervals. At the other end of the spectrum, short-GOP MPEG is synonymous with I-frame-only MPEG. Formats such as IMX use I-frame-only MPEG-2, which reduces temporal artifacts and improves editing performance. However, I-frame-only formats have a significantly higher data rate because each frame must store enough data to be completely self-contained. Therefore, although the decoding demands on your computer are decreased, there is a greater demand for scratch disk speed and capacity.

Maximum GOP length depends on the specifications of the playback device. The minimum GOP length depends on the GOP pattern. For example, an IP pattern can have a length as short as two frames.

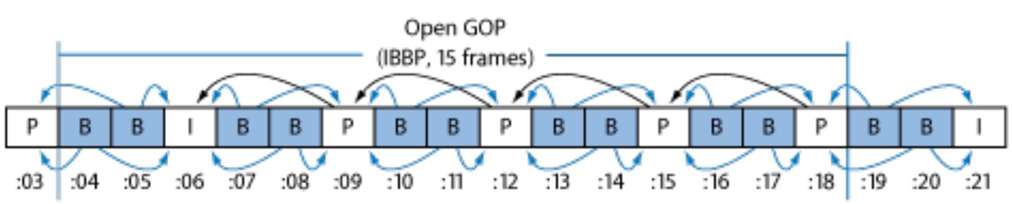
Here are several examples of GOP length used in common MPEG formats:

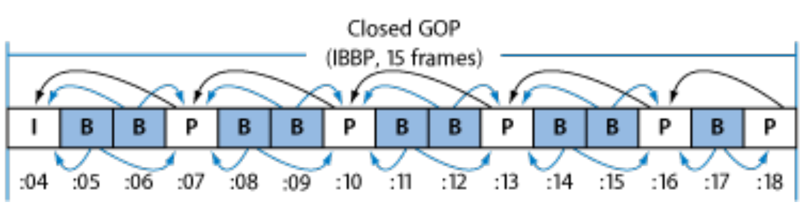
- *MPEG-2 for DVD:* Maximum GOP length is 18 frames for NTSC or 15 frames for PAL. These GOP lengths can be doubled for progressive footage.
- *1080-line HDV:* Uses a long-GOP structure that is 15 frames in length.
- *720-line HDV:* Uses a six-frame GOP structure.
- *IMX:* Uses only I-frames.

Open and Closed GOPs

An open GOP allows the B-frames from one GOP to refer to an I- or P-frame in an adjacent GOP. Open GOPs are very efficient but cannot be used for features such as multiplexed multi-angle DVD video. A closed GOP format uses only self-contained GOPs that do not rely on frames outside the GOP.

The same GOP pattern can produce different results when used with an open or closed GOP. For example, a closed GOP would start an IBBP pattern with an I-frame, whereas an open GOP with the same pattern might start with a B-frame. In this example, starting with a B-frame is a little more efficient because starting with an I-frame means that an extra P-frame must be added to the end (a GOP cannot end with a B-frame).





MPEG Containers and Streams

MPEG video and audio data are packaged into discrete data containers known as *streams.* Keeping video and audio streams discrete makes it possible for playback applications to easily switch between streams on the fly. For example, DVDs that use MPEG-2 video can switch between multiple audio tracks and video angles as the DVD plays.

Each MPEG standard has variations, but in general, MPEG formats support two basic kinds of streams:

- *Elementary streams:* These are individual video and audio data streams.
- *System streams:* These streams combine, or multiplex, video and audio elementary streams together. They are also known as *multiplexed streams*. To play back these streams, applications must be able to demultiplex the streams back into their elementary streams. Some applications only have the ability to play elementary streams.

MPEG-1

MPEG-1 is the earliest format specification in the family of MPEG formats. Because of its low bit rate, MPEG-1 has been popular for online distribution and in formats such as Video CD (VCD). DVDs can also store MPEG-1 video, though MPEG-2 is more commonly used. Although the MPEG-1 standard actually allows high resolutions, almost all applications use NTSC- or PAL-compatible image dimensions at quarter resolution or lower.

Common MPEG-1 formats include 320 x 240, 352 x 240 at 29.97 fps (NTSC), and 352 x 288 at 25 fps (PAL). Maximum data rates are often limited to around 1.5 Mbps. MPEG-1 only supports progressive-scan video.

MPEG-1 supports three layers of audio compression, called *MPEG-1 Layers 1, 2, and 3*. MPEG-1 Layer 2 audio is used in some formats such as HDV and DVD, but MPEG-1 Layer 3 (also known as *MP3*) is by far the most common. In fact, MP3 audio compression has become so popular that it is usually used independently of video.

MPEG-1 elementary stream files often have extensions such as .m1v and .m1a, for video and audio, respectively.

MPEG-2

The MPEG-2 standard made many improvements to the MPEG-1 standard, including:

- Support for interlaced video
- Higher data rates and larger frame sizes, including internationally accepted standard definition and high definition profiles
- Two kinds of multiplexed system streams—Transport Streams (TS) for unreliable network transmission such as broadcast digital television, and Program Streams (PS) for local, reliable media access (such as DVD playback)

MPEG-2 categorizes video standards into MPEG-2 Profiles and MPEG-2 Levels. Profiles define the type of MPEG encoding supported (I-, P-, and B-frames) and the color sampling method used (4:2:0 or 4:2:2 $Y'C_BC_R$). For example, the MPEG-2 Simple Profile (SP) supports only I and P progressive frames using 4:2:0 color sampling, whereas the High Profile (HP) supports I, P, and B interlaced frames with 4:2:2 color sampling.

Levels define the resolution, frame rate, and bit rate of MPEG-2 video. For example, MPEG-2 Low Level (LL) is limited to MPEG-1 resolution, whereas High Level (HL) supports 1920 x 1080 HD video.

MPEG-2 formats are often described as a combination of Profiles and Levels. For example, DVD video uses Main Profile at Main Level (MP @ ML), which defines SD NTSC and PAL video at a maximum bit rate of 15 Mbps (though DVD limits this to 9.8 Mbps).

MPEG-2 supports the same audio layers as MPEG-1 but also includes support for multichannel audio. MPEG-2 Part 7 also supports a more efficient audio compression algorithm called *Advanced Audio Coding*, or *AAC*.

MPEG-2 elementary stream files often have extensions such as .m2v and .m2a, for video and audio, respectively.

MPEG-4

MPEG-4 inherited many of the features in MPEG-1 and MPEG-2 and then added a rich set of multimedia features such as discrete object encoding, scene description, rich metadata, and digital rights management (DRM). Most applications support only a subset of all the features available in MPEG-4.

Compared to MPEG-1 and MPEG-2, MPEG-4 video compression (known as *MPEG-4 Part 2*) provides superior quality at low bit rates. However, MPEG-4 supports high-resolution video as well. For example, Sony HDCAM SR uses a form of MPEG-4 compression.

MPEG-4 Part 3 defines and enhances AAC audio originally defined in MPEG-2 Part 7. Most applications today use the terms *AAC audio* and *MPEG-4 audio* interchangeably.

MPEG-4 Part 10, or H.264

MPEG-4 Part 10 defines a high-quality video compression algorithm called *Advanced Video Coding* (AVC). This is more commonly referred to as *H.264*. H.264 video compression works similarly to MPEG-1 and MPEG-2 encoding but adds many additional features to decrease data rate while maintaining quality. Compared to MPEG-1 and MPEG-2, H.264 compression and decompression require significant processing overhead, so this format may tax older computer systems.

About H.264 Video

You can use the H.264 video format, also known as MPEG-4 part 10 and the Advanced Video Codec (AVC), in your HD projects. The H.264 encoder is twice as efficient as the standard MPEG-2 encoder. When compared to encoding with MPEG-2, this means that with the H.264 encoder:

- You can use a lower bit rate to get the same quality, resulting in smaller files.
- You can use the same bit rate and get better quality with the same file size.

While DVD Studio Pro does not include an embedded H.264 encoder, Compressor includes presets specifically for use in HD projects. You are able to modify these presets to suit your specific needs.

Important: You must use H.264 for HD DVD presets in Compressor to encode H.264 video assets for native use in DVD Studio Pro HD projects.