



# Mathematically Lossless Video — Better Than Uncompressed

Digital video technology is in widespread use today in many professional video application areas including nonlinear editing systems (NLE), digital disk recorders (DDR), and media servers. All of these systems need to record and playback one or more digital video streams in real time. Such digital video systems typically perform the following:

- Conversion of an analog video signal to the digital domain using analog-to-digital converters.
- Transfer of the digitized video stream to a storage device such as a hard disk.
- Retrieval of the video stream for manipulation/transformation/compositing.
- Encoding of the video stream back into the analog domain using digital-to-analog converters.

Most professional systems employ the ITU-R 601 video standard. The digitized 601 video signal contains 20 megabytes (MB) of data per second of video (see Table 1). For example, one hour of digitized video requires 70.4 gigabytes (GB) of storage space. Since some systems need to playback two streams simultaneously, such as an NLE system performing A/B roll mixing, the throughput demand can be as high as 40 MB/sec of sustained data. Storing or retrieving that much data to/from a hard disk in real time demands a very high-performance system with high bandwidth characteristics throughout the entire system and massive storage devices, leading to a very expensive implementation. Until now, only high-end post-production houses have been able to afford such systems.

<b>Spatial resolution</b>	720 × 486
<b>Frame rate</b>	30 fps
<b>Bits /component sample</b>	8 bits (1 byte)
<b>Sampling rate/component</b>	13.5 MHz
<b>Color space</b>	YUV 4:2:2
<b>Luminance samples /line</b>	720
<b>Chrominance samples/line</b>	2 × 360

**Data-rate calculation:**

720 × 486 pixels × 30 fps × 2 components/pixel × 1 byte/component = 20.02 MB/sec.

**Table 1** ITU-R 601 Digital Video Characteristics

In order to reduce system cost, compression techniques such as Motion-JPEG are used. M-JPEG reduces the data-rate and storage requirements of digital video streams, but the compression process also introduces loss of information in the video signal. Video quality is inversely proportional to the compression factor — a higher compression factor leads to lower video quality with more compression artifacts. Until now, users have had to make the trade-off of either paying a very high price for uncompressed video systems or accepting the video-quality compromise inherent in affordable systems.

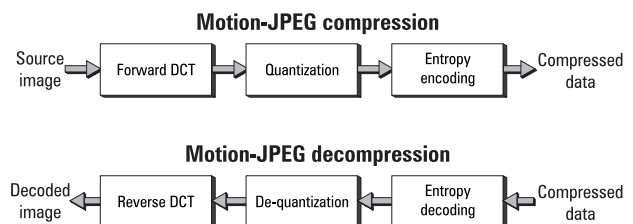
Fortunately, technology advances. Early in 1996, Matrox introduced a digital video platform called DigiSuite that offers the best of both worlds — the quality of uncompressed video at a reasonable cost. Using mathematically lossless encoding, DigiSuite performs data reduction on the video image using the entropy encoding part of the M-JPEG standard. Entropy encoding introduces no artifacts or loss of information — the process is **mathematically** lossless; that is, the digital binary code of the decoded image is identical to the source image. Users benefit from the data rate and storage space reduction of M-JPEG without the trade-off of introducing artifacts into the video signal.

In order to better understand how this is possible, a review of the M-JPEG algorithm is in order.

### M-JPEG compression without compromise

M-JPEG is an extended version of the still-image JPEG algorithm that's applied to motion video. In the M-JPEG algorithm, the still-image JPEG compression process is applied to each video field, one after the other, in real time.

JPEG is a symmetrical algorithm (see Figure 1). The complexity of the compression and the decompression methods are the same. The decompression process is the reverse operation of the compression process.



**Figure 1** Motion-JPEG algorithm

The operation of the JPEG algorithm can be divided into three basic steps:

- Reduction of the data redundancy by means of a discrete cosine transform (DCT).
- Quantization of the DCT coefficients using weighting functions optimized for the human visual system.
- Entropy encoding of the quantized DCT coefficients.

### Discrete Cosine Transform

DCT is a specific implementation of a Fourier transform whereby the image is transformed from its spatial representation to its frequential equivalent. Each element in the picture is represented by certain frequency coefficients. Most of the energy in an image is contained in the low-frequency coefficients (objects with large surfaces, such as sky, walls). Details in an image are expressed by the high-frequency coefficients (objects with fine details, such as hair, grass, tree branches). The human eye focuses primarily on the low frequency part of an image, since most of the energy is concentrated in this area. Interestingly, the high-frequency coefficients contain more information (data) than the low-frequency ones. DCT re-orders the information in preparation for quantization.

### Quantization

The quantization process is the lossy part of the JPEG algorithm. Quantization assigns a specific resolution (number of bits) to each frequency coefficient and then compresses the data by assigning fewer bits to the high-frequency coefficients, which matter less to the viewer. Quantization parameters are optimized for ITU-R 601 imagery; but still, the process compromises video quality. It's generally accepted that, at a 2:1 compression factor (approx. 10 MB/sec), the integrity of the video signal is **visually** lossless (Table 2). In applications such as digital video compositing and 3D animation, however, where video quality is of primary importance, no compromise can be accepted. In these applications, uncompressed video systems are currently being used.

	Data rate	Video quality
<b>Uncompressed</b>	20 MB/sec	D1
<b>Mathematically lossless M-JPEG</b>	<13 MB/sec	D1
<b>M-JPEG with quantization</b>	10 MB/sec	Visually lossless
	6 MB/sec	Betacam SP
	2 MB/sec	S-VHS
	1 MB/sec	VHS

**Table 2** M-JPEG Compression Factors

### Entropy encoding

The quantized coefficients contain runs of consecutive zeros. Therefore a data reduction advantage can be obtained using run-length encoding, where the data is represented by code symbols instead of the data itself. Further data reduction is also achieved by performing Huffman encoding. The Huffman algorithm assigns shorter symbols to code symbols that

statistically occur more often. Run-length and Huffman entropy encoding are steps that reduce the size of the data, but don't cause any data loss. Another example of entropy encoding is the zip file format in WinZip, which is widely used to compress computer files without loss of information.

### Lossless mode

Lossless M-JPEG encoding skips the DCT and quantization processes. Only entropy encoding is used. Since data is lost during quantization, the decoded image resulting from the application of the lossless mode is identical to the source image. Typically a compression factor of 1.6 to 1 or better is achieved.

This compression factor translates into a data rate of approximately 13 MB/sec but varies according to the frequential content of the source image. This method reduces system bandwidth requirements to a reasonable level and significantly increases the length of video that can be stored in a given storage space without any video quality compromise.

### Sounds easy — what took so long?

In theory, designing a mathematically lossless encoding system based on M-JPEG is simple. In practice, however, a number of issues must be taken into account. All the individual components such as the M-JPEG chip set, the disk controller and the hard disks must perform at a sufficiently high level. As well, system-level hardware and software architecture that is capable of performing data-stream management and data buffering at the necessary speed is required. All the pieces of the puzzle have finally come together to make a cost-effective, mathematically lossless A/B roll editing system a reality.

### M-JPEG chip set

To achieve mathematically lossless compression at ITU-R 601 resolution, a Motion-JPEG codec chip must have an internal pipeline architecture that performs at no less than 27 MHz (ITU-R 601 clock rate). The Zoran ZR36050 chip set achieves this level of performance.

### Storage device technology

In a cost-effective A/B roll editing system, two video streams must be read from a storage device simultaneously. This unified storage architecture eliminates the need to copy video segments from one device to another in order to perform A/B transitions. Two mathematically lossless video streams plus audio require a storage device that's capable of delivering an average of at least 27 MB/sec (two videos streams @ 13 MB/sec + audio). Ultra Wide SCSI supports a throughput of 40 MB/sec.

There are two different storage options available that can support the throughput possible with Ultra Wide SCSI — software stripe sets and hardware RAID's.

### Software stripe sets

Under the Windows NT file system (NTFS), multiple individual hard drives can be configured in parallel to provide overall throughput that is the sum of the throughput of the individual drives minus some overhead. The digital video data stream is split by the built-in NTFS software and chunks of data are written in a round-robin fashion to each disk in turn. Four individual Ultra Wide SCSI drives with individual throughputs in the neighborhood of 8-9 MB/sec could provide a software stripe set with performance in the desired range.

### Hardware RAIDs

Hardware RAIDs can also provide the needed throughput. They tend to be more expensive but provide the added features of data redundancy and hot-wap capability that may be useful in some applications. In addition, since the system perceives the RAID as a single drive, no extra software overhead is incurred.

### System-level design considerations

At the system level, a data-stream management and buffering architecture that can take full advantage of high-performance codec chips and storage technologies is needed. The DigiMotion codec card, part of the Matrox DigiSuite hardware platform, and Microsoft DirectShow (ActiveMovie) software architecture combine to provide the required system-level functionality.

### Hardware platform

The DigiMotion codec card provides the data-stream management and data buffering required to ensure mathematically lossless system performance (see Figure 2 below).

Data read from a storage device is always delivered in bursts for two reasons. 1. There's a time delay when the disk head seeks from one sector location to another. 2. The data streams are interleaved. On the other hand, video codecs require a constant flow of data to refresh video in real time. DigiMotion provides a memory buffer where data bursts written from the SCSI controller on one end are read out smoothly on the other end to supply the video codecs with the constant flow of data that they require. The memory access bandwidth has to deliver, on average, four times the data bandwidth since two streams are written and read simultaneously. The total required, on average, is four times 13 MB/sec = 52 MB/sec but instantaneous peak demand could reach 80 MB/sec. DigiMotion provides more than 100 MB/sec of bandwidth, thus leaving plenty of bandwidth for the audio channels.

### Software architecture

A software architecture that's capable of keeping up with the high-bandwidth hardware performance is required to complete the system. Microsoft DirectShow is a stream-oriented multimedia architecture that supports high data rates for realtime system operation.

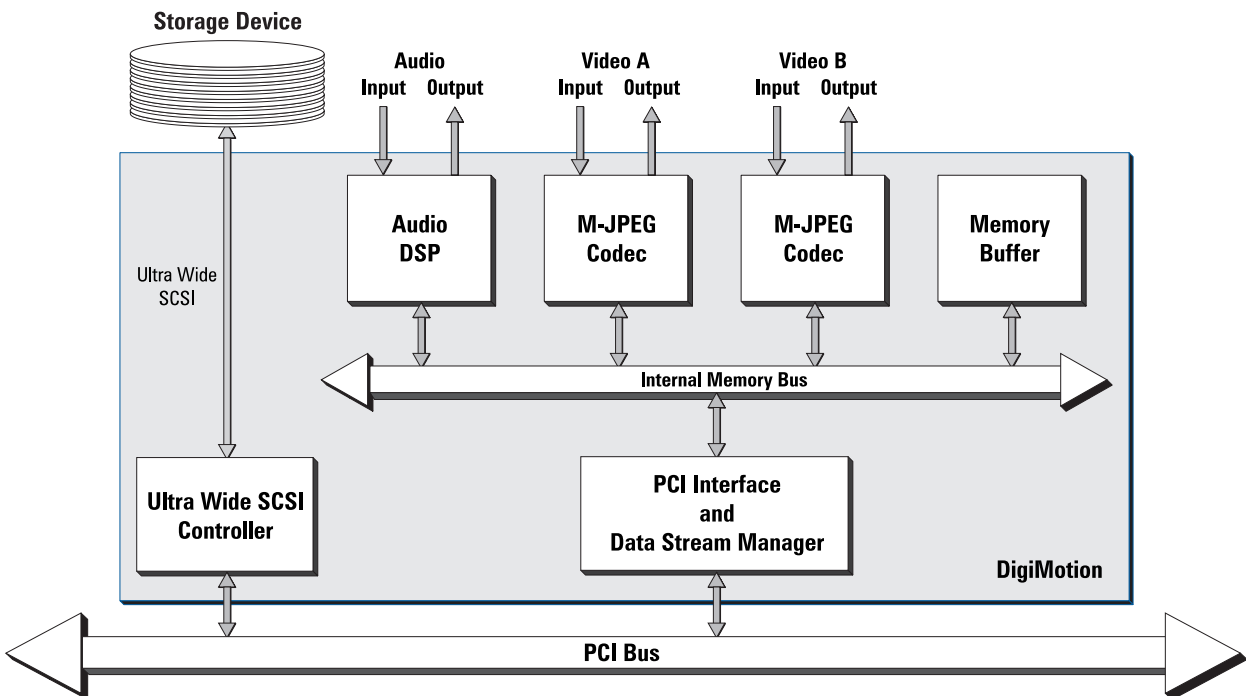


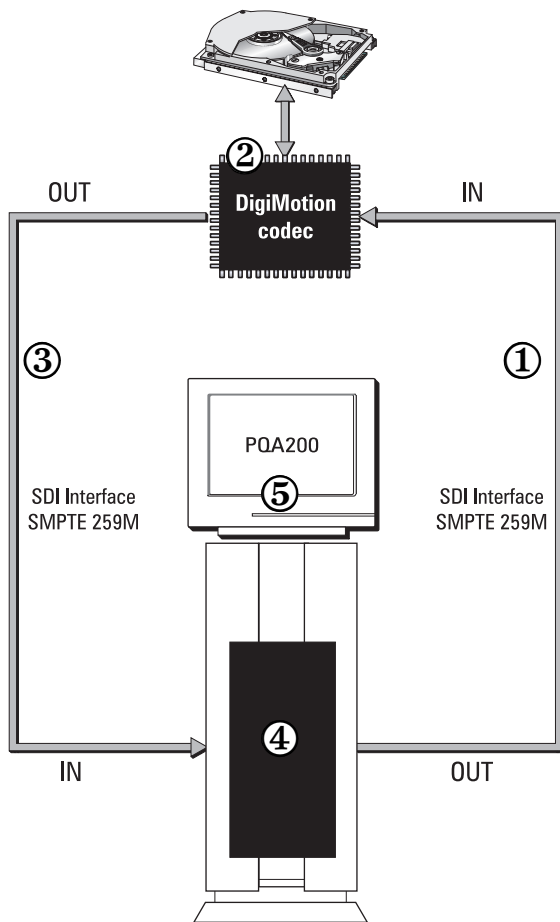
Figure 2 - Matrox DigiMotion data flow diagram

**High performance, low cost**

Mathematically lossless M-JPEG encoding is the basis for very cost-effective digital video systems. It offers video quality identical to that of uncompressed systems but reduces the throughput burden on the system, thereby allowing lower cost components to be used throughout. It also offers the important benefit of increased video storage capacity.

**Test method**

In order to demonstrate that the system is mathematically lossless, a test using the Tektronix PQA200 Picture Quality Analysis System is used.



**Figure 3** Test Setup

The Tektronix PQA200 performs a measurement resulting in a single numeric value of picture quality, called the Picture Quality Rating (PQR). Using a human vision system model

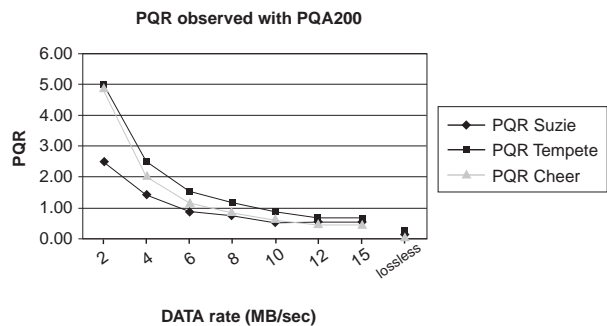
based on years of research by the Sarnoff Corporation, the PQA200 contains the three necessary dimensions for evaluation of dynamic and complex motion test sequences; spatial analysis, temporal analysis, and full color analysis. PQR values usually range between 5 and 0, with 5 being low quality (severe artifacts) and 0 indicating no loss. It's generally accepted that the human eye cannot detect artifacts on images with a PQR below 1.0. Images in this range are commonly described as "visually lossless." A PQR of 0 indicates no mathematical degradation at all.

**Test Procedure**

The test procedure involves the following steps:

- 1 Test video sequences output from the PQA200 are captured and compressed by the DigiMotion codec and written to disk.
- 2 The compressed streams are read from disk and decompressed by the DigiMotion codec.
- 3 The PQA200 captures back the decompressed video streams.
- 4 The PQA200 performs a comparison analysis between the video test sequences submitted to the codec prior to compression and the decompressed video streams captured back.
- 5 Numerical results are provided to the user on the PQA200 console.

Test sequences are performed in M-JPEG lossy at various data rates between 2 MB/sec and 15 MB/sec, and also in mathematically lossless mode. The video sequences, called Suzie, Tempete, and Cheer, are approximately 5 seconds long.



As can be observed, every test sequence run through the Matrox DigiMotion codec in mathematically lossless mode produces a PQR of zero (0), proving the concept.

*THIS DOCUMENT WAS ORIGINALLY PUBLISHED IN 1997*

Matrox Video Products Group  
 1055 St. Regis Blvd.  
 Dorval, Quebec, Canada H9P 2T4  
 1-800-361-4903 (U.S. and Canada)  
 (514) 685-2630

