

Video Compression: MPEG-4 and Beyond

[Ali Saman Tosun](#), tosun@cis.ohio-state.edu

Abstract

Technological developments in the networking technology and the computers make use of video possible. Storing and transmitting uncompressed raw video is not a good idea, it requires large storage space and bandwidth. Special algorithms which take the characteristics of the video into account can compress the video with high compression ratio. In this paper I will give a overview of the standardization efforts on video compression: MPEG-1, MPEG-2, MPEG-4, MPEG-7, and I will explain the current video compression trends briefly.

See also: [Multimedia Networking Products](#) | [Multimedia Over IP: RSVP, RTP, RTCP, RTSP](#) | [Multimedia Networking References](#) | [Books on Multimedia](#) | [Protocols for Multimedia on the Internet](#) (Class Lecture) | [Video over ATM networks](#) | [Multimedia networks: An introduction](#) (Talk by Prof. Jain)

| [Other Reports on Recent Advances in Networking](#)

[Back to Raj Jain's Home Page](#)

Table of Contents:

- [1. Introduction](#)
- [2. H.261](#)
- [3. H.263](#)
- [4. H.263+](#)
- [5. MPEG](#)
 - [5.1 MPEG-1](#)
 - [5.2 MPEG-2](#)
 - [5.3 MPEG-3](#)
 - [5.4 MPEG-4](#)
 - [5.5 MPEG-7](#)
- [6. J.81](#)
- [7. Fractal-Based Coding](#)
- [8. Model-based Video Coding](#)

- [9. Scalable Video Coding](#)
- [10. Wavelet-based Coding](#)
- [Summary](#)
- [References](#)
- [List of Acronyms](#)

1. Introduction

Over the last couple of years there has been a great increase in the use of video in digital form due to the popularity of the Internet. We can see video segments in web pages, we have DVDs to store video and HDTV will use a video format for broadcast. To understand the video formats, we need to understand the characteristics of the video and how they are used in defining the format.

Video is a sequence of images which are displayed in order. Each of these images is called a frame. We cannot notice small changes in the frames like a slight difference of color so video compression standards do not encode all the details in the video, some of the details are lost. This is called lossy compression. It is possible to get very high compression ratios when lossy compression is used. Typically 30 frames are displayed on the screen every second. There will be lots of information repeated in the consecutive frames. If a tree is displayed for one second then 30 frames contain that tree. This information can be used in the compression and the frames can be defined based on previous frames. So consecutive frames can have information like "move this part of the tree to this place". Frames can be compressed using only the information in that frame (intraframe) or using information in other frames as well (interframe). Intraframe coding allows random access operations like fast forward and provides fault tolerance. If a part of a frame is lost, the next intraframe and the frames after that can be displayed because they only depend on the intraframe.

Every color can be represented as a combination of red, green and blue. Images can also be represented using this color space. However this color space called RGB is not suitable for compression since it does not consider the perception of humans. YUV color space where only Y gives the grayscale image. Human eye is more sensitive to changes in Y and this is used in compression. YUV is also used by the NTSC, PAL, SECAM composite color TV standards.

Compression ratio is the ratio of the size of the original video to the size of the compressed video. To get better compression ratios pixels are predicted based on other pixels. In spatial prediction of a pixel can be obtained from pixels of the same image, in temporal prediction, the prediction of a pixel is obtained from a previously transmitted image. Hybrid coding consists of a prediction in the temporal dimension with a suitable decorrelation technique in the spatial domain. Motion compensation establishes a correspondence between elements of nearby images in the video sequence. The main application of motion compensation is providing a useful prediction for a given image from a reference image.

DCT (Discrete Cosine Transform) is used in almost all of the standardized video coding algorithms. DCT is typically done on each 8x8 block. 1-D DCT requires 64 multiplications and for an 8x8 block 8 1-D DCTs are needed. 2-D DCT requires 54 multiplications and 468 additions and shifts. 2-D DCT is used in

MPEG, there is also hardware available to do DCT. When DCT is performed, the top left corner has the highest coefficients and bottom right has the lowest, this makes compression easier. The coefficients are numbered in a zig-zag order from the top left to bottom right so that there will be many small coefficients at the end. The DCT coefficients are then divided by the integer quantization value to reduce precision. After this division it is possible to lose the lower coefficients if they are much smaller than the quantization. The coefficients are multiplied by the quantization value before IDCT (inverse DCT).

[Back to Table of Contents](#)

2. H.261

H.261 [[Clarke95](#)] was developed for transmission of video at a rate of multiples of 64Kbps. Videophone and videoconferencing are some applications. H.261 standard is similar to JPEG still image compression standard. H.261 uses motion-compensated temporal prediction. H.261 coder has a layered structure with 4 layers. 4 layers are picture layer, group of block (GOB) layer, macroblock (MB) layer and block layer. Each block is 8x8. The layers are multiplexed for transmission in series. Each layer has a header. Frame format of H.261 is called common intermediate format (CIF).

[Back to Table of Contents](#)

3. H.263

H.263 [[Rao96](#)] was designed for very low bit rate coding applications. H.263 uses block motion-compensated DCT structure for encoding. H.263 encoding has higher efficiency than H.261 encoding. An encoding specification called test model (TMN) was used for optimization in H.263. There are different versions of test models, the latest version is called TMN5. H.263 is based on H.261 but it is significantly optimized for coding at low bitrates. Video coding is performed by partitioning each picture into macroblocks. Each macroblock consists of 16x16 luminance block and 8x8 chrominance blocks of Cb and Cr. Each macroblock can be coded as intra or as inter. Spatial redundancy is exploited by DCT coding, temporal redundancy is exploited by motion compensation. H.263 includes motion compensation with half-pixel accuracy and bidirectionally coded macroblocks. 8x8 overlapped block motion compensation, unrestricted motion vector range at picture boundary, and arithmetic coding are also used in H.263. These features are not included in MPEG-1 and MPEG-2 since they are useful for low bitrate applications. H.263 decoding is based on H.261 with enhancements to support coding efficiency. Four negotiable options are supported to improve performance. These are unrestricted motion vector mode, syntax-based arithmetic coding mode, advanced prediction mode and PB-frames mode. Unrestricted motion vector mode allows motion vectors to point outside a picture. Syntax-based arithmetic coding mode allows using arithmetic coding instead of Huffman coding. Advanced prediction mode uses overlapped block motion compensation with four 8x8 block vectors instead of a single 16x16 macroblock motion vector. PH-frames mode allows a P-frame and a B-frame to be coded together as a

single PB-frame.

[Back to Table of Contents](#)

4. H.263+

H.263+ [[Villasenor](#)] is an extension of H.263. It has several additional features and negotiable additional modes. It provides SNR scalability as well as spatial and temporal scalability. It has custom source formats. Advanced intra coding is used to improve the compression efficiency for intra macroblock encoding by using spatial prediction of DCT coefficient values. Deblocking filter mode reduces the amount of block artifacts in the final image by filtering across the block boundaries using an adaptive filter. Slice structure allows a functional grouping of a number of macroblocks in the picture, enabling improved error resilience, improved transport over packet networks and reduced delay. Reference picture resampling mode allows a resampling of a temporally previous reference picture prior to its use as a reference for encoding, enabling global motion compensation, predictive dynamic resolution conversion, predictive picture area alteration and registration and special-effect warping. Reduced resolution update mode allows an encoder to maintain a high frame rate during heavy motion by encoding a low-resolution update to a higher resolution picture while maintaining high resolution in stationary areas. Independent segment decoding mode enhances error resilience by ensuring that corrupted data from some region of the picture cannot cause propagation of error into other regions. Alternate inter VLC mode reduces the number of bits needed for encoding predictively-coded blocks when there are many large coefficients in the block. Modified quantization mode improves the bitrate control by changing the method for controlling the quantizer step size on a macroblock basis, reduces the prevalence of chrominance artifacts by reducing the step size for chrominance quantization, increases the range of representable coefficient values for use with small quantizer step sizes, and increases error detection performance and reduces decoding complexity by prohibiting certain unreasonable coefficient representations.

[Back to Table of Contents](#)

5. MPEG

MPEG (Moving Picture Experts Group) [[MPEG](#)] is a ISO/IEC working group developing international standards for compression, decompression, and representation of moving pictures and audio.

5.1 MPEG-1

MPEG video compression standard [[MPEGa, LeGall91](#)] is a layered, DCT-based video compression standard that results in VHS quality compressed video stream that has a bit rate of approximately

1.5Mbps at a resolution of approximately 352x240. At a high level, MPEG video sequences consist of several different layers that provide the ability to randomly access a video sequence as well as provide a barrier against corrupted information. Size layers within MPEG are given in Table 1.

Sequence layer	Random access unit: content
Group of pictures layer	Random access unit: video
Picture layer	Primary coding unit
Slice layer	Resynchronization unit within picture
Macroblock layer	Motion compensation unit within slice
Block layer	DCT unit within macroblock

Table 1: Size layers within MPEG

All MPEG frames are encoded in one of three different ways: Intra-coded (I-frames), Predictive-coded (P-frames), or Bidirectionally-predictive-coded (B-frames). I-frames are encoded as discrete frames, independent of adjacent frames. Thus, they provide randomly accessible points within the video stream. Because of this, I-frames have the worst compression ratio of the three other frames. P-frames are coded with respect to a past I-frame or P-frame, resulting in a smaller encoded frame size than the I-frames. The B-frames require a preceding and a future frame, which may be either I-frames or P-frames, in order to be decoded, but they offer the highest degree of compression.

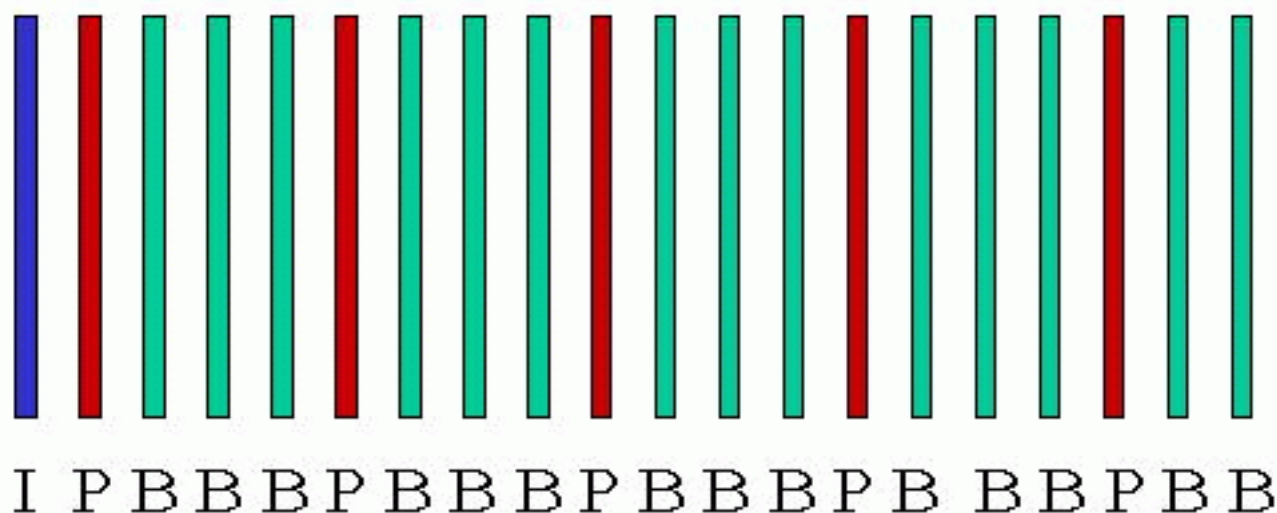


Figure 1: A Group of pictures in coding order

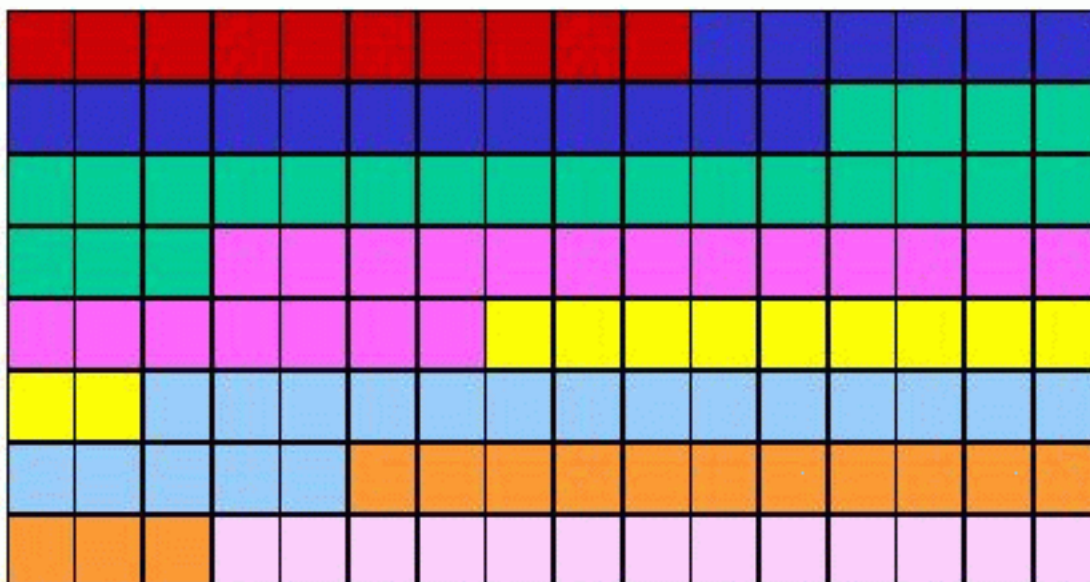


Figure 2: A possible slice structure for a frame

As can be seen in Table 1, an MPEG video consists of many video sequences, each of which has many group of pictures in it. Typically a group of pictures (GOP) has a single I-frame and many P and B-frames. Although there is no limit on the size of GOP, values around 15 is common. A typical GOP is given in Figure 1. The frames are divided in blocks of size 16x16 and these blocks are called macroblocks. A sequence of macroblocks is called a slice. A slice can be an entire row of macroblocks or it can start in one row and end in another. Each frame consists of many slices. A possible slice structure is given in Figure 2. Each color represents a different slice. Most of the parameters can be specified during the encoding stage. You can specify the GOP pattern, size of a slice, desired number of frames per second. Each of these layers has a header used for synchronization. If part of a slice is lost, the decoder will skip the rest of the slice when it detects the error and start decoding from the beginning of the next slice.

Video decoding is a computationally expensive operation, MPEG-1 decoding can be done in real time using a 350MHz Pentium processor.

5.2 MPEG-2

MPEG-1 has a bit rate of about 1.5Mbps, MPEG-2 [Watkinson99] is designed for diverse applications which require a bit rate of up to 100Mbps. Digital high-definition TV (HDTV), interactive storage media (ISM), cable TV (CATV) are sample applications. Multiple video formats can be used in MPEG-2 coding to support these diverse applications. MPEG-2 has bitstream scalability: it is possible to extract a lower bitstream to get lower resolution or frame rate. Decoding MPEG-2 is a costly process, bitstream scalability allows flexibility in the required processing power for decoding. MPEG-2 is upward, downward, forward, backward compatible. Upward compatibility means the decoder can decode the

pictures generated by a lower resolution encoder. Downward compatibility implies that a decoder can decode the pictures generated by a higher resolution encoder. In a forward compatible system, a new generation decoder can decode the pictures generated by an existing encoder and in a backward compatible system, existing decoders can decode the pictures generated by new encoders.

In MPEG-2 the input data is interlaced since it is more oriented towards television applications. Video sequence layers are similar to MPEG-1 the only improvements are field/frame motion compensation and DCT processing, scalability. Macroblocks in MPEG-2 has 2 additional chrominance blocks when 4:2:2 input format is used. 8x8 block size is retained in MPEG-2, in scaled format blocks can be 1x1, 2x2, 4x4 for resolution enhancement. P and B frames have frame and field motion vectors.

MPEG-2 decoding can not be done in real time with current processors.

5.3 MPEG-3

MPEG-3 was intended for EDTV and HDTV for higher bit rates and later merged with MPEG-2.

5.4 MPEG-4

Success of digital television, interactive graphics applications and interactive multimedia encouraged MPEG group to design MPEG-4 [[Wen98](#), [Villasenor](#), [MPEG4](#), [MPEG4f](#), [German](#), [Avaro97](#)] which allows the user to interact with the objects in the scene within the limits set by the author. It also brings multimedia to low bitrate networks.

MPEG-4 uses media objects to represent aural, visual or audiovisual content. Media objects can be synthetic like in interactive graphics applications or natural like in digital television. These media objects can be combined to form compound media objects. MPEG-4 multiplexes and synchronizes the media objects before transmission to provide QoS and it allows interaction with the constructed scene at receiver's machine.

MPEG-4 organizes the media objects in a hierarchical fashion where the lowest level has primitive media objects like still images, video objects, audio objects. MPEG-4 has a number of primitive media objects which can be used to represent 2 or 3-dimensional media objects. MPEG-4 also defines a coded representation of objects for text, graphics, synthetic sound, talking synthetic heads.

MPEG-4 provides a standardized way to describe a scene. Media objects can be placed anywhere in the coordinate system. Transformations can be used to change the geometrical or acoustical appearance of a media object. Primitive media objects can be grouped to form compound media objects. Streamed data can be applied to media objects to modify their attributes and the user's viewing and listening points can be changed to anywhere in the scene.

Visual part of the MPEG-4 standard describes methods for compression of images and video, compression of textures for texture mapping of 2-D and 3-D meshes, compression of implicit 2-D meshes, compression of time-varying geometry streams that animate meshes. It also provides algorithms

for random access to all types of visual objects as well as algorithms for spatial, temporal and quality scalability, content-based scalability of textures, images and video. Algorithms for error robustness and resilience in error prone environments are also part of the standard.

For synthetic objects MPEG-4 has parametric descriptions of human face and body, parametric descriptions for animation streams of the face and body. MPEG-4 also describes static and dynamic mesh coding with texture mapping, texture coding with view dependent applications.

MPEG-4 supports coding of video objects with spatial and temporal scalability. Scalability allows decoding a part of a stream and construct images with reduced decoder complexity (reduced quality), reduced spatial resolution, reduced temporal resolution., or with equal temporal and spatial resolution but reduced quality. Scalability is desired when video is sent over heterogeneous networks, or receiver can not display at full resolution (limited power)

Robustness in error prone environments is an important issue for mobile communications. MPEG-4 has 3 groups of tools for this. Resynchronization tools enables the resynchronization of the bitstream and the decoder when an error has been detected. After synchronization data recovery tools are used to recover the lost data. These tools are techniques that encode the data in an error resilient way. Error concealment tools are used to conceal the lost data. Efficient resynchronization is key to good data recovery and error concealment.

5.5 MPEG-7

Lots of audio-visual information is available on the web, but there are not adequate search tools to locate this information. The aim of MPEG-7 [[MPEG7](#)] is to specify a set of descriptors to describe various forms of multimedia. It will also standardize ways to define other descriptors as well as structures for the descriptors and their relationship. This information will be associated with the content to allow fast and efficient search. MPEG-7 will also standardize a language to specify description schemes.

MPEG standardization efforts and approval dates are given in Table 2

MPEG-1	November 1992
MPEG-2	November 1994
MPEG-4 version 1	October 1998
MPEG-4 version 2	December 1999
MPEG-7	July 2001

Table 2: MPEG approval dates

[Back to Table of Contents](#)

6. J.81

This standard [[Torress96](#)] specifies the coding and transmission of digital television signals at vbit rates of 34-45Mbps in the format specified by recommendation ITU-R 601. Net video capacity is between 26 and 31Mbps for Europe and depends on the number of optional channels used. J.81 provides very high quality which is suitable for transparent compression necessary for contribution applications.

ITU-T J.11 operates on a field basis, with three different processing modes for each macroblock: intra, interfield predicted and interframe predicted. An extension of this algorithm was used to in the first all-digital transmission of HSTV through satellite and fiber optic links in 1980 in Europe.

[Back to Table of Contents](#)

7. Fractal-Based coding

Fractal coding [[Torress96](#)] is a new and promising technique . In an image values of pixels that are close are correlated. Transform coding takes advantage of this observation. Fractal compression takes advantage of the observation that some image features like straight edges and constant regions are invariant when rescaled. Representing straight edges and constant regions efficiently using fractal coding is important because transform coders cannot take advantage of these types of spatial structures. Fractal coding tries to reconstruct the image by representing the regions as geometrically transformed versions of other regions in the same image. There is not much material on fractal-based video coding. Most of the research currenty focuses on fractal based image coding.

[Back to Table of Contents](#)

8. Model-based Video Coding

Model based schemes [[Torress96](#), [Lopez97](#)] define three dimensional space structural models of the scene. Coder and decoder use an object model. The same model is used by coder to analyse the image, and by decoder to generate the image. Traditionally research in model-based video coding (MBVC) focuses on head modeling, head tracking, local motion tracking, and expression analysis, synthesis. MBSV have bean mainly used for videoconferencing and videotelephony since mostly human head is modeled. MBVC has concentrated in modeling of image like head and shoulders because it is impossible to model every object that may be in the scene. There is lots of interest in applications such as speech driven image animation of talking heads and virtual space teleconferencing.

In model-based approaches a parametrized model is used for each object in the scene. Coding and transmission is done using the parameters associated with the objects. Tools from image analysis and computer vision is used to analyze the images and find the parameters. This analysis provides information on several parameters like size, location, and motion of the objects in the scene. Results have shown that it is possible to get good visual quality at rates as low as 16kbps.

[Back to Table of Contents](#)

9. Scalable Video Coding

Multimedia communication systems may have nodes with limited computation power to be used for decoding and heterogeneous networks such as combination of wired and wireless networks. In these cases we need the ability to decode at a variety of bit rates. Scalable coders [Girod96] has this property. Layered multicast has been proposed as a way to provide scalability in video communication systems.

MPEG-2 has basic mechanisms to achieve scalability but it is limited. Spatiotemporal resolution pyramids is a promising approach to provide scalable video coding. Open-loop and closed-loop pyramid coders both provide efficient video coding and inclusion of multiscale motion compensation. Simple filters can be used for spatial downsampling and interpolation operations and fast and efficient codecs can be implemented. Morphological filters can also be used to improve image quality.

Pyramid coders have multistage quantization scheme. Bit allocation to the various quantizers depending on the image is important to get efficient compression. Optimal bit allocation is optimally computationally infeasible when pyramids with more than two layers are used. Closed-loop pyramid coders are better suited for practical applications than open-loop pyramid coders since they are less sensitive to suboptimal bit allocations and simple heuristics can be used.

There are several ways to utilize multistage motion compensation. Efficiently computing motion vectors and then encoding them by hierarchical group estimation is one way. When video is sent over heterogeneous networks scalability is utilized by offering a way to reduce the bit rate of video data in case of congestion. By using priorities the network layer can reduce bitrate without knowing the content of the packet or informing the sender.

[Back to Table of Contents](#)

10. Wavelet-based Coding

Wavelet transform [Chen99, Shen99, Spiht, Video, Asbun99] techniques have been investigated for low bit-rate coding. Wavelet-based coding has better performance than traditional DCT-based coding. Much lower bit-rate and reasonable performance are reported based on the application of these techniques to

still images. A combination of wavelet transform and vector quantization gives better performance. Wavelet transform decomposes the image into a multifrequency channel representation, each component of which has its own frequency characteristics and spatial orientation features that can be efficiently used for coding. Wavelet-based coding has two main advantages: it is highly scalable and a fully-embedded bit-stream may be easily generated. The main advantage over standard techniques such as MPEG is that video construction is achieved in a fully embedded fashion. Encoding and decoding process can stop at a pre-determined bit rate. The encoded stream can be scaled to produce the desired spatial resolution and frame rate as well as the required bit rate. Vector quantization makes use of the correlation and the redundancy between nearby pixels or between frequency bands. Wavelet transform with vector quantization exploits the residual correlation among different layers if the wavelet transform domain using block rearrangement to improve the coding efficiency. Further improvements can also be made by developing the adaptive threshold techniques for classification based on the contrast sensitivity characteristics of the human visual system. Joint coding of the WT with trellis coded quantization as a joint source/channel coding is an area to be considered.

Additional video coding research applying the wavelet transform on a very low bit-rate communication channel is performed. The efficiency of motion compensated prediction can be improved by overlapped motion compensation in which the candidate regions from the previous frame are windowed to obtain a pixel value in the predicted frame. Since the wavelet transform generates multiple frequency bands, multifrequency motion estimation is available for the transformed frame. It also provides a representation of the global motion structure. Also, the motion vectors in lower-frequency bands are predicted with the more specific details of higher-frequency bands. This hierarchical motion estimation can also be implemented with the segmentation technique that utilizes edge boundaries from the zero-crossing points in the wavelet transform domain. Each frequency band can be classified as temporal-activity macroblocks or no-temporal-activity macroblocks. The lowest band may be coded using an H.261 like coder which uses DCT, and the other bands may be coded using vector quantization or trellis coded quantization.

[Back to Table of Contents](#)

Summary

Video compression is gaining popularity since storage requirements and bandwidth requirements are reduced with compression. There are many algorithms for video compression each aimed at a different target. In this paper I explained the standardization efforts for video compression such as H.261, H.263, H.263+, MPEG-1, MPEG-2, MPEG-4, MPEG-7. There are many different criteria for video coding, scalable video coding is good for heterogeneous systems, wavelet coding is good for low bit-rate systems and there are some video coding algorithms which need to be researched further. Fractal-based video coding, model-based video coding and segmentation-based video coding are some of these.

[Back to Table of Contents](#)

List of Acronyms

MPEG: Moving Picture Experts Group
GOP: Group of Pictures
MBVC: Model Based Video Coding
DCT: Discrete Cosine Transform
JPEG: Joint Photographic Experts Group

References

[Chen99] Chang Wen Chen, Ya-Qin Zhang, Visual Information Representation, communication, and image processing, 1999 Marcel Dekker, 551 pages

This book is a collection of papers in image processing and video coding

[Clarke95] Roger J. Clarke, Digital Compression of Still image and video, 1995 Academic Press, 453 pages

This book describes different methods of intraframe coding and standards and alternative schemes for video coding

[Torres96] Lois Torres, Murat Kunt, Video Coding, 1996 Kluwer Academic Publishers, 433 pages

This book describes second generation video coding techniques

[Rao96] K. R. Rao, J. J. Hwang, Techniques and standards for image, video and audio coding, 1996 Prentice Hall, 563 pages

This book describes the techniques for image and video coding

[Watkinson99] John Watkinson, MPEG-2, 1999 Focal Press, 244 pages

This book describes MPEG-2 from an electrical engineering perspective

[MPEGa] Joan Mitchell, William B. Pennebaker, Chad, E. Fogg and Didier J. Legall, MPEG video compression standard, International thompson publishing, 470 pages,

This book explains the MPEG standard with fragments of code

[LeGall91] Didier LeGall, MPEG: a video compression standard for multimedia applications.

Communications of the ACM, April 1991, Vol 34, No 4, 13 pages

This paper gives an overview of MPEG standard

[Girod96] Bernd Girod, Khaled Ben Younes, Reinhard Bernstein, Peter Eisert, Niko Farber, Frank Hartung, Uwe Horn, Eckehard Steinbach, Klaus Stuhlmuller, Thomas Wiegand, Recent advances in video compression, Proceedings of ISCAS-96

<http://www-nt.e-technik.uni-erlangen.de/~bernstei/publika/mypubli.html>

[Shen99] Ke Shen, Edward Delp, Wavelet based rate scalable video compression. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No. 1, February 1999, pp. 109-122. .

<http://dynamo.ecn.purdue.edu/~ace/delp-pub.html>

[Villasenor] John Villasenor, Ya-Qin Zhang, Jiangtao Wen, Robust video coding algorithms and

systems. Special Issue of the Proceedings of the IEEE on Wireless Video. 20 pages

<http://www.icsl.ucla.edu/~ipl/papers/RobustVideo.html>

[MPEG4] Overview of MPEG-4 standard, 52pages

http://drogo.cselt.stet.it/mpeg/#The_MPEG_standards

[Spiht] Beong-Jo Kim, William A. Pearlman, An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT) 10 pages

[Video] Andreas Polzer, Hansjorg Klock, Joachim Buhmann, Video coding by region-based motion compensation and spatio-temporal wavelet transform, 4 pages

[Wen98] Jiangtao Wen, John D. Villasenor, Reversible Variable Length codes for efficient and robust image and video coding. Proceedings of the 1998 IEEE Data Compression Conference pp 471-480, 10 pages <http://www.icsl.ucla.edu/~ipl/publications.html>

[Asbun99] Eduardo Asbun, Paul Salama, Ke Shen, Edward J. Delp., Very low bit rate wavelet-based scalable video compression. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No. 1, February 1999, pp. 109-122. <http://dynamo.ecn.purdue.edu/~ace/delp-pub.html>

[MPEG7] MPEG-7: Context and Objectives at

<http://drogo.cselt.stet.it/mpeg/standards/mpeg-7/mpeg-7.htm>

[German] A. Mufit German, Bilge Gonsel, A. Murat Tekalp, Object based indexing of MPEG-4 compressed video. Proceedings of IST/SPIE Symposium on Electronic Imaging vol 3024 11 pages

<http://www.ee.rochester.edu:8080/users/ferman/publications.html>

[MPEG4f] H. Kalva, A. Eleftheriadis, A. Puri, and R. Schmidt, Stored file formats for MPEG-4 Contribution ISO-IEC JTC1/SC29/WG11 MPEG97/2062, April 1997, Bristol, UK (39th MPEG meeting) <http://www.ee.columbia.edu/~elef/publications.html>

[Avaro97] Olivier Avaro, Philip A. Chou, Alexandros Eleftheriadis, Casten Herpel, Cliff Reader, Julien Signes, The MPEG-4 systems and description languages: a way ahead in audio visual representation, Signal Processing Image Communication, Special Issue on MPEG-4, 1997, 43 pages

<http://www.ee.columbia.edu/~elef/publications.html>

[Lopez97] R. Lopez, A. Colmenarez, T. S. Huang, Head and Feature Tracking for Model-based Video Coding, International Workshop on Synthetic-Natural Hybrid Coding and 3-D Imaging, Greece, 1997, 4 pages <http://troi.ifp.uiuc.edu/~antonio/publications.html>

[MPEG] Official MPEG webpage at <http://drogo.cselt.stet.it/mpeg/>

[Back to Table of Contents](#)

Last Modified: November 23,1999

Note: This paper is available on-line at

<http://www.cis.ohio-state.edu/~jain/cis788-99/compression/index.html>