Homework

Prof. Dr. Bernd Steinbrink

Multimedia Applications

MPEG Standard for Web-Based-Training?

by

Markus Alber

Content

1. Introduction

The Internet has spawned a revolution in higher education. New learning environments are emerging rapidly to meet the needs and expectations of students. So, what comprises e-learning infrastructures or e-infrastructures?

As reported in EduCause Quarterly, Number 2, 2000, EduCause Current Issues Committee Chair Paul B. Gandel lists some of the following challenges:

- Distance Education
- E-Learning Environments

Working on this field of **Web-Based-Training** there is a great challenge on using Audio and Video embedded in the application you want to offer or distribute. The questions are:

- How to use video, audio or graphics as teaching and learning tools?
- How get students access to these media themselves, during class or homework time or WBT-session?
- Is there precise control over playback of video or audio?
- How can the students benefit from audiovisual presentation of material?
- How to implement and organize streaming media in the internet in order to cover dynamic and synchronous or isochronous media use?

Digital media can help accomplishing these delivery tasks more efficiently, and the investment in time and money becomes surprisingly modest, if we choose the right standard.

1.1 Reason for Focus on MPEG

MPEG is a family of ISO/IEC standards for digital video and audio compression which optimize the match between quality and storage requirements. These standards are established and maintained by the Moving Pictures Experts Group.

MPEG-1 is the best established of these standards. MPEG-2, targeted for broadcast-quality television, requires significantly higher data transmission rates and is not practical for most instructional use. MPEG-4, a broader standard for interactive multimedia including graphics and text in addition to video and audio and encompassing a greater range of data rates. Since 2001 MPEG-7 is the content representation standard for multimedia information search, filtering, management and processing (to be approved July 2001).

MPEG is the work of the „**M**oving **P**ictures **E**xperts **G**roup“ , which was meant for developing high-quality video compression standards. So the MPEG specification does not define a special protocol, but a special type of compressing data. The following pages should only be a short introduction and discussion to MPEG.

The **M**oving **P**ictures **E**xperts **G**roup is a working group of ISO/IEC in charge of the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio and their combination. So far **MPEG** has *produced:*

MPEG-1 - standard for storage and retrieval of moving pictures and audio on storage media up to about 1,5 Mbit/s. 1993

MPEG-2 - coding of moving pictures and associated audio information and standard for digital television, 1994

and is now *developing*:

MPEG-4 - standard for multimedia applications; Version 1; building on digital television, interactive graphics applications (synthetic content) and the World Wide Web (distribution of and access to content) and will provide the standardized technological elements. Formal ISO/IEC designation is released in November 1998.

MPEG-7 - the content representation standard for multimedia information search, filtering, management and processing, 2001.

2. MPEG-*Video* - defining video compression

- MPEG-Audio - defining audio compression
- MPEG-Video - defining video compression
- MPEG-System - defines the interaction between Video, Audio and private streams. There can be up to 32 audio, 16 video and 2 private streams. Every stream is divided into packets. Timestamps are defining, when it has to be shown (further information can be found on www.mpeg.org)

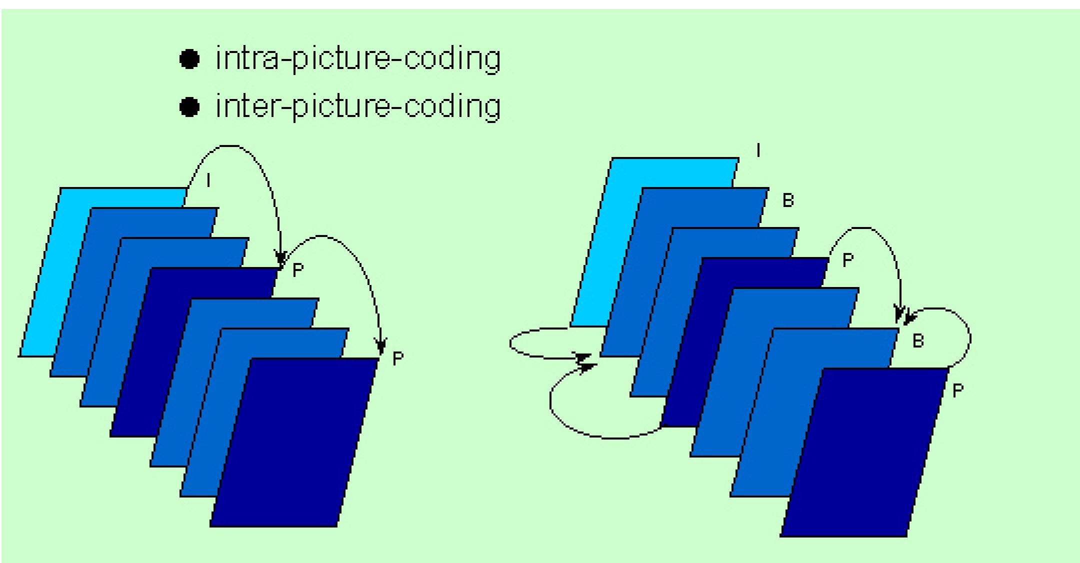
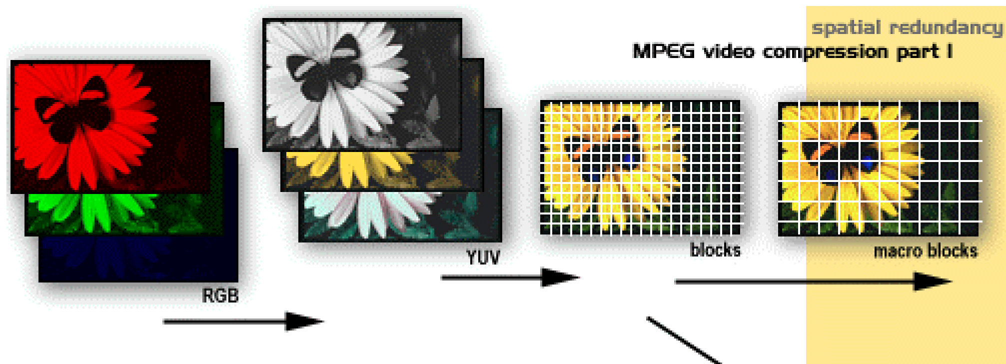


figure: Axel Maurer, Universität Karlsruhe, www.ubka.uni-karlsruhe.de/~axel/MM/sld017.htm

2.1 The Compression Technique

Video compression relies on the eye's inability to resolve **H**igh **F**requency color changes, and the fact that there's a lot of **redundancy** within each frame (spatial redundancy) and between frames (temporal redundancy).

By doing *spatial* redundancy we are creating an **I-frame** (*Intra-frame*). The frame includes all information to decode the picture. To reduce the bitrate, not every picture results in an I-frame. The main thing about doing this, is storing only the difference to a specific frame by *temporal* redundancy.



We know, that the eye is not very sensitive to high-frequency color changes. In order to compress the picture the RGB signal is converted into a **YUV** signal (luminance and two color difference signals).

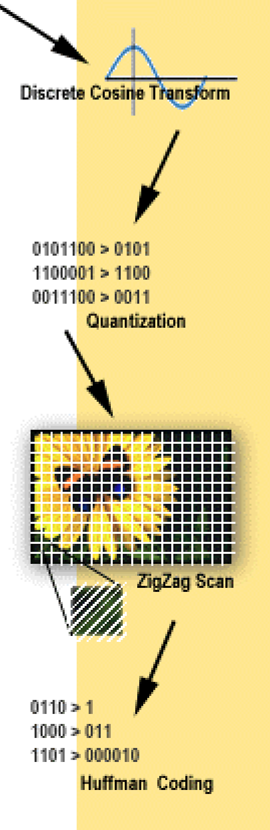
The **D**iscrete **C**osine **T**ransform (taking 8x8 pixels blocks to save processing time) is used, along with quantization and Huffmann coding; to predict a pixel value from all adjacent pixel values, and minimize the overall bit rate.

The DCT is lossless and reversible, but the next stage, the **quantization** causes the compression. The quantisized data (8x8 pixel blocks) is then ZigZag scanned to help the following **entropy coding** (Huffman). The more often a special value occurs, the smaller the binary value representing the pixels value.

This generates the **Intra-frames (I-frames)**.

<u>**Prediction & motion compensation,**</u> predicts the value of pixels in a frame, from the information in adjacent frames. Therefore the picture is divided into macro blocks (typically 16x16 blocks).

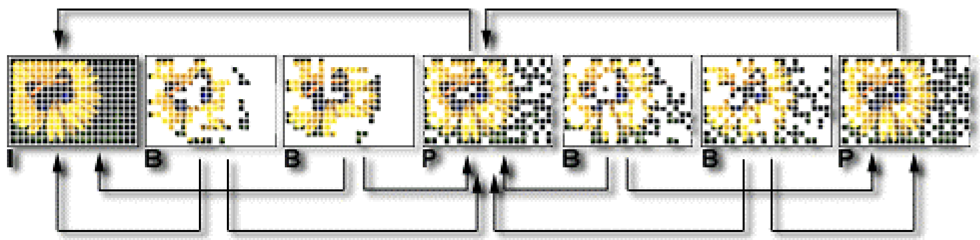
Interframe prediction and motion estimation: Frame similarity is used to reduce the overall bitrate. So predicted **B-** and **P-frames** are generated.



The MPEG- Standard defines **three** different type of frames:

- **I-Frames** (*intra-frames*) They don't depend on following or previous frames. This is the only type of frame to continue after an error has been detected. They are also used for fast-forward, etc.
- **P-Frames** (*predicted frames*) These frames are additionally predicted from a previous P- or I- Frame.
- **B-Frames** (*bi- directional predicted frames*) They offer the greatest compression and use past and future I- and/or P- frames for motion compensation. This frame type is the most error- sensitive frame

The following diagram shows the dependencies between the different type of frames. **For example:** to encode/decode the third picture in the below shown sequence, you need to know the first frame (I- frame) and the fourth one (P- frame).



The order of the frames will depend on the application. Usually every 12th or 15th frame is an **Intra-Frame.**

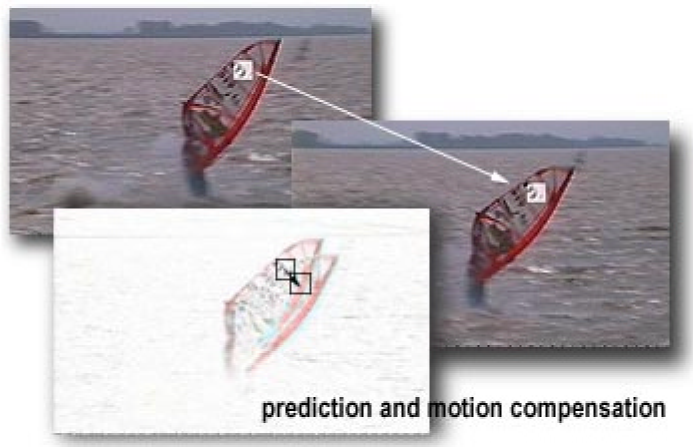
A video sequence coded using I-frames only **(I I I I I I ...)** allows very good random access, Fast Forward/Fast Rewind and editability, but achieves only low compression.

A sequence coded with a regular I-picture update and no B-frames **(i.e I P P P P P P I P P P P ...)** achieves moderate compression and using all three frame types, **(as i.e. I B B P B B P B B I B B P ...)**, may achieve high compression and reasonable random access and FF/FR functionality but also increases the coding delay significantly. This delay may not be tolerable for e.g. videotelephony or videoconferencing applications.

(Video Editing is currently dominated by Motion-JPEG, because of a disadvantage of the MPEG-Codec: most frames depend on other frames, so editing a specific frame may cause problems)

2.2 Motion Compensation

The following illustration shows the block matching approach for **motion compensation**: One motion vector is estimated for each block in the actual frame to be coded. The **motion vector** points to a reference block of same size in a previously coded frame.



In the MPEG compression algorithms the motion compensated **prediction** techniques are used for reducing temporal redundancies between frames and only the prediction error images - the **difference between original images and motion compensated prediction images** - are encoded.

2.3 Frame Dependencies

When you look at the frame dependencies, describing the frame order, you will recognize, that the displayed frames can't be transmitted in the same order, because of the dependencies between the specific frames.

example:

display order:	I_1	B_1	B_2	P_1	B_3	B_4	P_2	B_5	B_6	P_3	B_7	B_8	I_2
file order:	I_1	P_1	B_1	B_2	P_2	B_3	B_4	P_3	B_5	B_6	I_2	B_7	B_8

For the bidirectional frame $\mathbf{B}_7$ you need the frames $\mathbf{I}_2$ and $\mathbf{P}_3$, so these two frames have to be transmitted before, in order to decode the picture.

3. MPEG – CODECs - Overview

3.1 MPEG-1

See the document by International Organisation for Standardisation

Overview

- specified 1993
- *non-interlaced*
- *motion compensation & prediction*

Some Technical Facts

resolutions

- Europe Y: 352x288; C_b, C_v: 176x144; 50 fields/sec (25 frames/sec)
- US: Y: 352x240; C_b, C_y: 176x120; 60 fields/sec (30 frames/sec)
- theoretical resolution: 4095x4095x60Hz
- bitrate: 4 Mbit/s for 352x240x30 Hz and 352x288x25Hz (Video)

data-rates for CD bit-rate 1,5MBit/sec - MPEG-1:

- CD-audio can be compressed down to 0,25MBit/sec (6:1)
- system data stream (synchronization, ...) 0,1MBit/sec
- 1,15MBit/sec left for Video, (26:1)

The video compression technique developed by MPEG-1 covers many applications from interactive systems on CD-ROM to the delivery of video over telecommunications networks. The MPEG-1 video should support a wide range of applications, so the input parameters including flexible picture size and frame rate can be specified by the user. MPEG has recommended a constraint parameter set: every MPEG-1 compatible decoder must be able to support at least video source parameters up to TV size: including a minimum number of 720 pixels per line, a minimum number of 576 lines per picture, a minimum frame rate of 30 frames (NTSC) per second and a minimum bit rate of 1.86 Mbits/s. The standard video input consists of a *non-interlaced* video picture format.

However, MPEG-1 was primarily targeted for multimedia CD-ROM applications, requiring additional functionality supported by both encoder and decoder. Important features provided by MPEG-1 include frame based random access of video, fast forward/fast reverse (FF/FR) searches through compressed bit streams, reverse playback of video and editability of the compressed bit stream.

interlacing: the standard video input format for MPEG-1 is **non-interlaced**. However, coding of interlaced colour television with both 525 and 625 lines at 29.97 (NTSC) and 25 (PAL) frames per second respectively is/was an important application for the MPEG-1 standard, based on the conversion of the interlaced source to a progressive intermediate format. Note that MPEG-2 supports interlacing.

3.2 MPEG-2

See the document by International Organisation for Standardisation

Overview

- finally introduced 1994
- supports *interlacing*
- new *motion compensation* modes
- different *chrominance formats* (Y:U:V)
- multiple bitstreams
- *DVD* and *digital TV* standard
- *zig-zag scan* before entropy coding (Huffman)
- optimized Huffman table

Some Technical Facts

examples for resolution and bitrates for NTSC:

- 4 Mbit/s: 352x240x30 Hz (Video)
- 15 Mbit/s: 720x480x30 Hz (SDTV, Standard Definition TV)
- 60 Mbit/s: 1440x1150x30 Hz (HDTV)
- 80 Mbit/s: 1920x1080x30 Hz

Basically MPEG-2 can be seen as a superset of the MPEG-1 coding standard and was designed to be **backward compatible** to MPEG-1 - every MPEG-2 compatible decoder can decode a valid MPEG-1 bit stream.

MPEG-2 is also a digital standard for video at TV resolution (i.e. CCIR-Norm 720 × 576 pixels). In MPEG-1 there's no special concept on how to cope with the to fields in an interlaced TV picture. MPEG-2 is able to encode interlaced video and is today the standard for **DVD** and **digital TV.**.

However, implementation of the full variety may not be practical for most applications. MPEG-2 has introduced the concept of "Profiles" and "Levels" handling equipment not supporting the full implementation.

Profiles and **Levels** provide means for defining subsets of the syntax and thus the decoder capabilities required to decode a particular bit stream, i.e. higher profiles support several video streams (parallel, for example at two different resolutions). MPEG-2 video is also divided into **levels**. A level specifies *image size, frame rate* and *bitrates* of an MPEG-2 video. Main Profile at Main level (MP@ML) supports non-scaleable coding of digital video with approximately digital TVparameters with 720 x 576. The bitrates at MP@ML are around 8 MBit/s. DVD and digital-TV (using MPEG-2) are additionally encrypted.

MPEG-2 has introduced **new motion compensation modes** to efficiently cope with the temporal redundancies between fields, namely the "Dual Prime" prediction and the motion compensation based on 16x8 blocks. A discussion of these methods is beyond the scope of this introduction.

MPEG-2 has specified additional Y:U:V luminance and chrominance **subsampling ratio formats** for applications with highest video quality requirements. Next to the 4:2:0 format already supported by MPEG-1 the specification of MPEG-2 is extended to 4:2:2 formats suitable for high quality studio video coding applications.

Flexibly supporting **multiple resolutions** is of particular interest for interworking between HDTV and Standard Definition Television (SDTV), in which case it is important for the HDTV receiver tobe compatible with the SDTV product. Compatibility can be achieved by means of scalable coding of the HDTV source. Transmitting **two independent bit streams** to the HDTV and SDTV receivers is very wasteful and can be avoided by MPEG-2. Other important applications for scalable coding include video database browsing and multiresolution playback of video in multimedia environments.

3.3 MPEG-4 Version 1

See the document by International Organisation for Standardisation

Some Technical Facts on MPEG-4 Version 1 (may change):

features:

- progressive and interlaced *Scanning* Methods.
- Luminance Spatial *Resolutions* : sizes from 8x8 to 2048x2048, e.g. SQSIF/SQCIF and CCIR 601.
- *Color* Spaces: monochrome, Y, C_r, C_b,combined with an *Alpha* Channel
- *Chrominance Spatial Resolutions* : 4:0:0, 4:2:0, 4:2:2.
- T*emporal Resolutions* : various resolutions with as maximum the capture rate. The frame rate shall be continuously variable, on a frame-by-frame basis.
- *Pixel Color Depths* : up to 8 bits per component

bitrates supported:

MPEG-4 Video is optimized for :

- <64 kBit/sec (**low**)
- 64-384kBit/sec (**intermediate**) and
- 384-4MBit/sec (**high**) bitrates. It shall support both *constant* bitrate (CBR) and *variable* bitrate (VBR).

Some of the new MPEG-4 functionalities require **higher computational power**, but provide higher compression efficiency or new functionalities. Techniques which are considered to add computational complexity compared to the previous MPEG video standards include: shape coding of arbitrarily shaped objects, sprite generation, macroblock padding for arbitrarily shaped objects and rendering system at the decoder. Experience with previous video standards has shown that for computational intensive algorithms fast HW and SW implementations were found soon (compare e.g. DCT/IDCT or Motion Estimation).

MPEG-4 is a concept for broadcasting-, movie- and multimedia applications. Because of the scaleable bitrating it should be perfect for use on the net. It handles small bitrates (4-64 Kbit/s) for instance: 176 x 144 x 10 Hz for ISDN- Videophone. Take a look at invited papers at http://leonardo.telecomitalialab.com/icjfiles/mpeg-4_si/index.htm

Hint:
(MPEG-4 Version 2 and MPEG-7 are not explicitely discussed in this paper- please have a closer look to International Organisation for Standardisation and http://arge.tuwien.ac.at/text/AG1/MPEG7/sld001.htm)

4. MPEG-4 Standard for use in WBT-Applications

Streaming video over the Internet is becoming very popular, using viewing tools as software plug-ins for a Web browser. Wbt-applications is just an examples of many possible video streaming applications.

Here, bandwidth is limited due to the use of modems, and transmission reliability is an issue, as packet loss may occur. Increased error resilience and improved coding efficiency will improve the experience of streaming video. In addition, scalability of the bitstream, in terms of temporal and spatial resolution, but also in terms of video objects, under the control of the viewer, will further enhance the experience, and also the use of streaming video.

4.1 Features and functionalities - Result

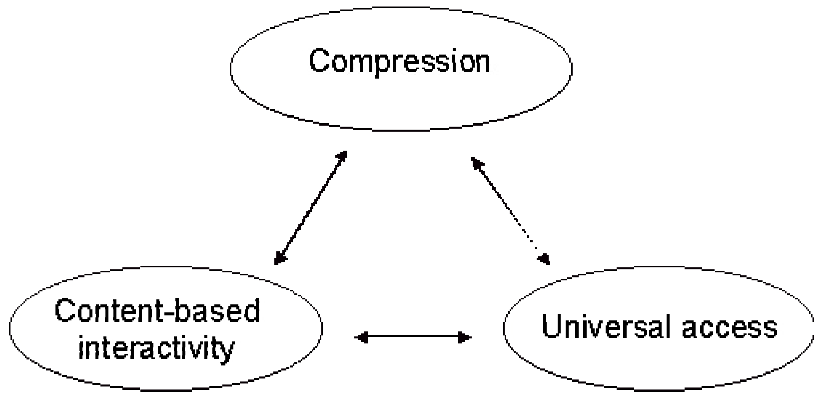
The MPEG-4 visual standard consists of a set of tools that enable applications by supporting several classes of functionalities. The most important features covered by MPEG-4 standard can be clustered in three categories (see Fig. below) and summarized as follows:

1) Compression efficiency: Compression efficiency (compare Steinbrink, 2002, Lecture-Slides “Selection of Multimedia Platforms”, No. 15) has been the leading principle for MPEG-1 and MPEG-2, and in itself has enabled applications such as Digital TV and DVD. Improved coding efficiency and coding of multiple concurrent data streams will increase acceptance of applications based on the MPEG-4 standard.

2) Content-based interactivity: Coding and representing video objects rather than video frames enables content-based applications. It is one of the most important qualities offered by MPEG-4. Based on efficient representation of objects, object manipulation, bitstream editing, and object-based scalability allow new levels of content interactivity

3) Universal access: Robustness in error-prone environments allows MPEG-4 encoded content to be accessible over a wide range of media, such as mobile networks as well as wired connections. In addition, object-based temporal and spatial scalability allow the user to decide where to use sparse resources, which can be the available bandwidth, but also the computing capacity or power consumption.

Functionalities offered by the MPEG-4 visual standard in a figure:



To support some of these functionalities, MPEG-4 should provide the capability to represent arbitrarily shaped video objects. Each object can be encoded with different parameters, and at different qualities. The shape of a video object can be represented in MPEG-4 by a binary or a gray-level (alpha) plane. The texture is coded separately from its shape. For low-bitrate applications, frame based coding of texture can be used, similar to MPEG-1 and MPEG-2. To increase robustness to errors, special provisions are taken into account at the bitstream level to allow fast resynchronization, and efficient error recovery.

The MPEG-4 visual standard has been explicitly optimized for three bitrate ranges:

1. Below 64 kbit/sec

2. 64 - 384 kbit/sec

3. 384- 4 Mbit/sec

For high quality applications, higher bitrates are also supported while using the same set of tools and the same bitstream syntax for those available in the lower bitrates. MPEG-4 provides support for both interlaced and progressive material.

The chrominance format that is supported is 4:2:0. In this format the number of Cb and Cr samples are half the number of samples of the luminance samples in both horizontal and vertical directions. Each component can be represented by a number of bits ranging from 4 to 12 bits.

Summary:

MPEG 4 is the actually data format for stored and streamed interactive multimedia content design and web-based-training applications (see also: Steinbrink 2002, Lecture-Slides "Multimedia Applications", No. 5)

literature:

Avaro, Olivier; Herpel, Carsten; Signes, Julien:
http://woody.imag.fr/MPEG4/syssite/syspub/docs/tutorial/sld007.htm

Koehnen, Rob:
http://leonardo.telecomitalialab.com/icjfiles/mpeg-4_si/11-Profiles_paper/11-Profiles_paper.htm

Kossmeier, Andreas:
http://wbt-3.iicm.edu/kossmeier.andreas/mpeg/mpeg4.htm

Steinbrink, Prof. Dr. Bernd:
Lecture-Slides (Quarter I, II, III)

Schinagel, Wolfgang:
http://arge.tuwien.ac.at/text/AG1/MPEG7/sld001.htm

Chiariglione, Leonardo:
http://imsc.usc.edu/Events/seminars/980128/sld003.html

Rehm, Eric: http://www.acm.org/sigs/sigmm/MM2000/ep/rehm/

Ebrahimi , Touradj; Horne, Caspar:
http://leonardo.telecomitalialab.com/icjfiles/mpeg-4_si/7-natural_video_paper/7-natural_video_paper.htm