2.3 Video Compression

2.3.1 MPEG

MPEG stands for Moving Picture Experts Group (a committee of ISO).

Goal of MPEG-1: compress a video signal (with audio) to a data stream of 1.5 Mbit/s, the data rate of a T1 link in the U.S. and the rate that can be streamed from a CD-ROM.

Goals of the MPEG-1 Compression Algorithm

- Random access within 0.5 s while maintaining a good image quality for the video
- Fast forward / fast rewind
- Possibility to play the video backwards
- Allow easy and precise editing

MPEG Frame Types

Four types of frames:

I-Frame (Intra Frame)

Intra-coded full image, very similar to the JPEG image, encoded with DCT, quantization, run-length coding and Huffman coding

P-Frame (Predicted Frame)

Uses delta encoding. The P frame refers to preceeding I- and P-frames. DPCM encoded macroblocks, motion vectors possible.

B-Frame (Interpolated Frame)

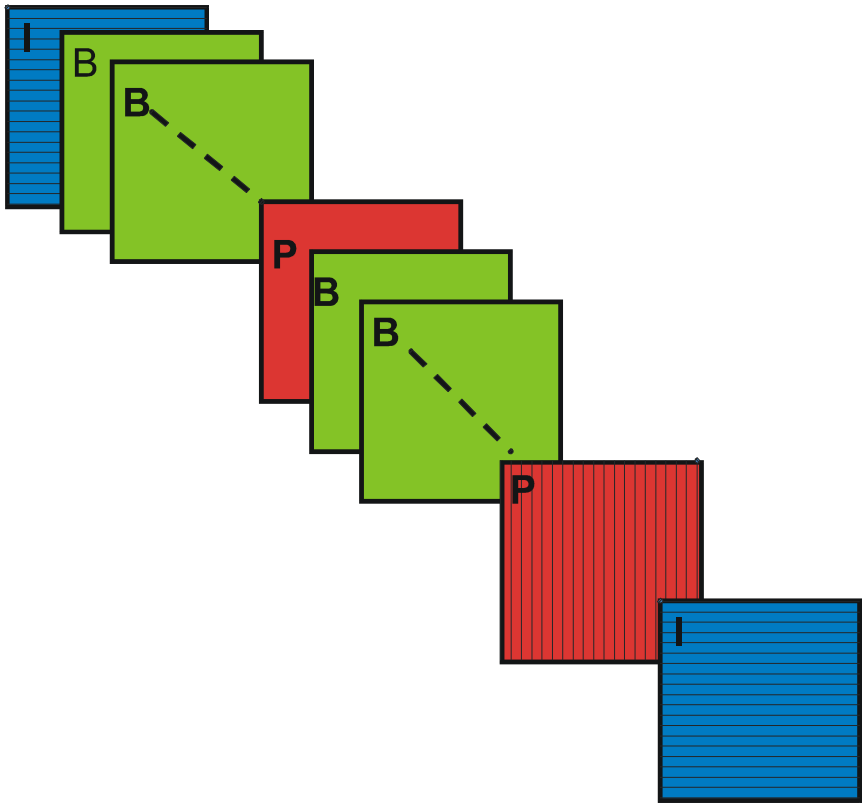
"bidirectionally predictive coded pictures„. The B frame refers to preceeding and succeeding frames, interpolated the data and encodes the differences.

D-Frame

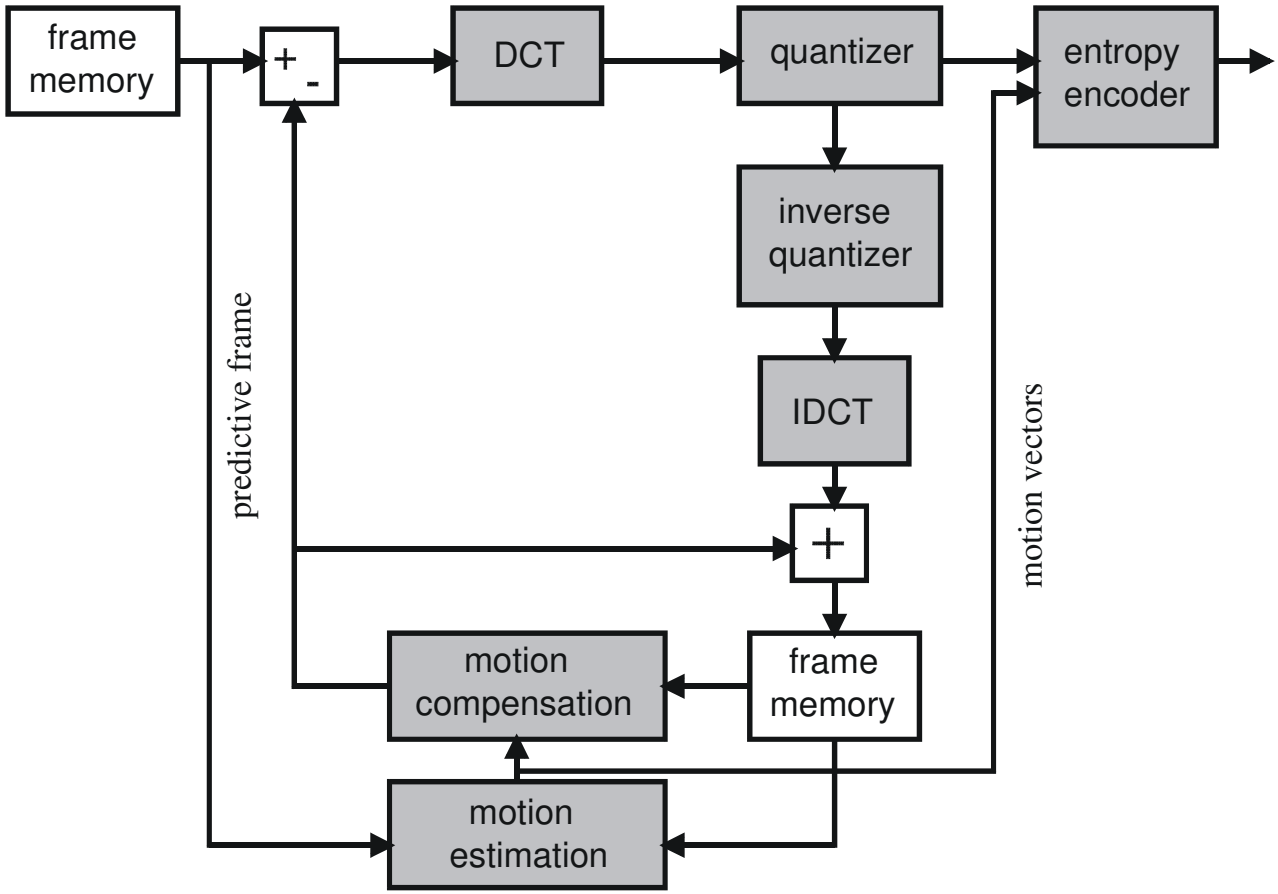
"DC coded picture", only the DC coefficient of each block is coded (upper left-hand corner of the matrix), e.g., for previews.

„Group of Pictures“ in MPEG

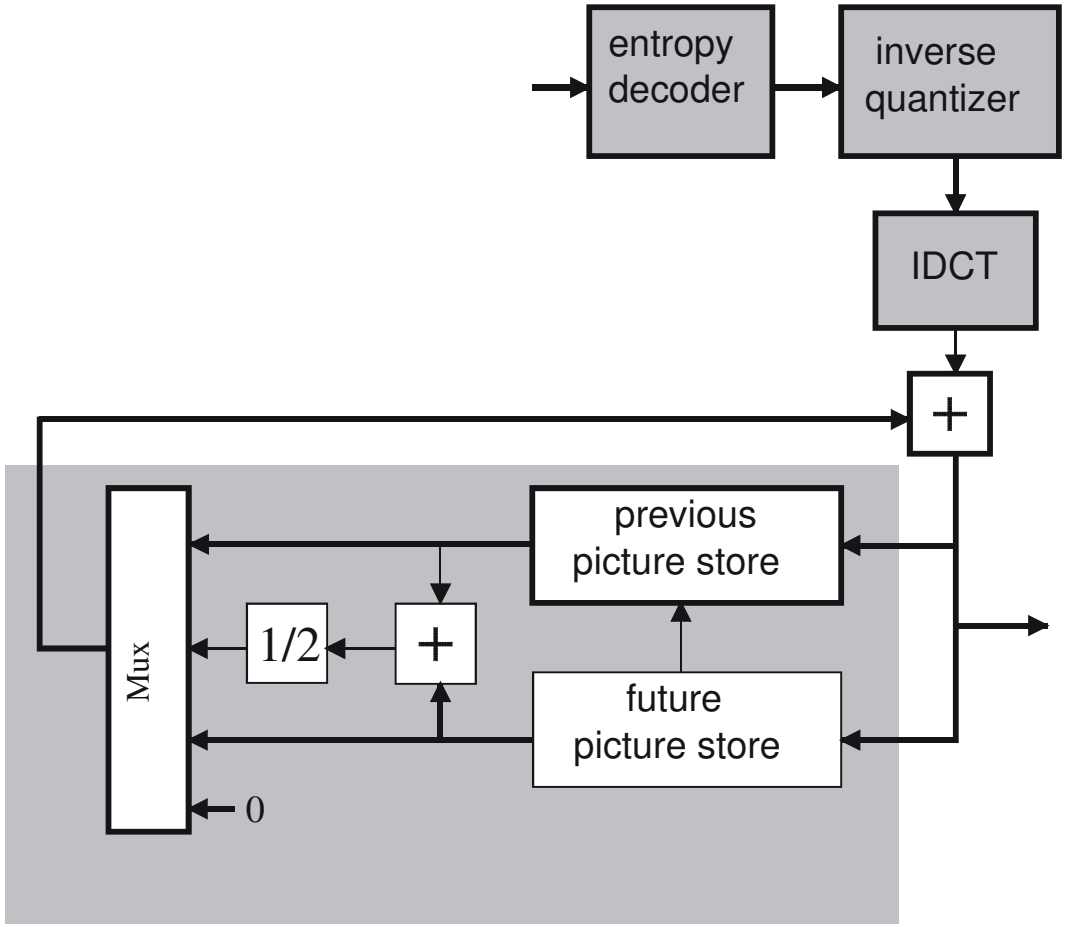
The sequence of I, P and B frames is not standardized but can be chosen according to the requirements of the application. This allows the user to chose his/her own compromise between video quality, compression rate, ease of editing, etc.



MPEG Encoder



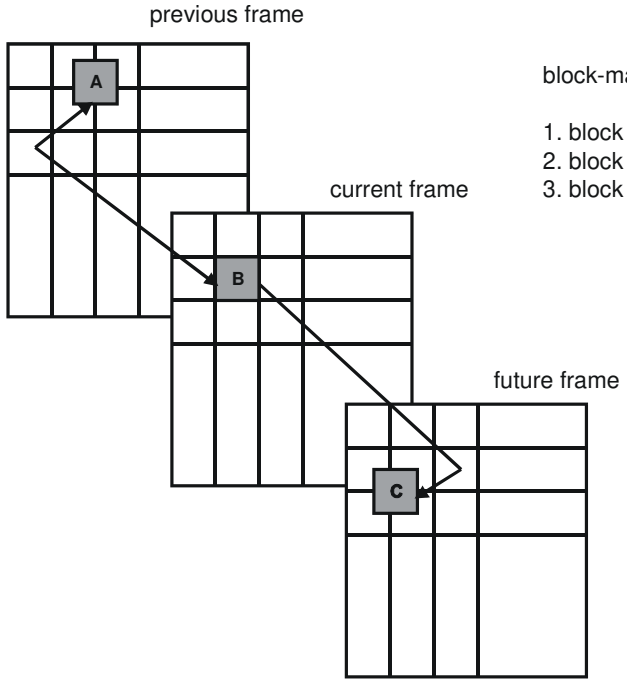
MPEG Decoder



motion compensation

Temporal Redundancy and Motion Vectors

"Motion Compensated Interpolation"



block-matching technique

1. block B = block A
2. block B = block C
3. block B = (block A + block C) / 2

On the encoder side the search range can be chosen as a parameter: the larger the search range, the higher the potential for compression, but the longer the run-time of the algorithm.

MPEG-2

MPEG-2 extends MPEG-1 for higher bandwidths and better image qualities, up to HDTV. It was developed jointly by ISO and ITU-T (where the standard is called H.262).

MPEG-2 defines scalable data streams which allow receivers with different bandwidth and processing power to receive and decode only parts of the data stream.

Scalability in MPEG-2 (1)

- **"SNR scalability“:** Each frame is encoded in several layers. A receiver who only decodes the base layer will get a low image quality. A receiver decoding addtional (higher) layers gets a better image quality. An example is color subsampling: the base layer contains only one quarter of the values for the U and V components, compared to the Y components. The enhancement layer contains the U and V components in full resolution, for better color quality.

- **"Spatial scalability“:** The frames are encoded with different pixel resolutions (e.g., for a standard TV set and for an HDTV TV set. Both encodings are transmitted in the same data stream.

Scalability in MPEG-2 (2)

- **"Temporal scalability“:** The base layer contains only very few frames per second, the enhancement layers additional frames per second. Receivers decoding the higher layers will thus get a higher frame rate (i.e., a higher temporal resolution).

- **"Data partinioning“:** The data stream is decomposed into several streams with different amounts of redundany for error correction. The most important parts of the stream are encoded in the base layer, e.g., the low-frequency coefficients of the DCT and the motion vectors. This layer can then be enriched with an error correcting code for better error resilience than the enhancement layers where errors are not as harmful.

MPEG-2 Video Profiles

	Simple profile no B frames not scalable	Main profile B frames not scalable	SNR scalable profile B frames SNR scaling	Spatially scalable profile B frames spatial scaling	High profile B frames spatial or SNR scaling
High level 1920x1152x60		<=80 Mbits/s			<=100 Mbits/s
High-1440 level 1440x1152x60		<=60 Mbits/s		<=60 Mbits/s	<=80 Mbits/s
Main level 720x576x30	<=15 Mbits/s	<=15 Mbits/s	<=15 Mbits/s		<=20 Mbits/s
Low level 352x288x30		<=4 Mbits/s	<=4 Mbits/s		

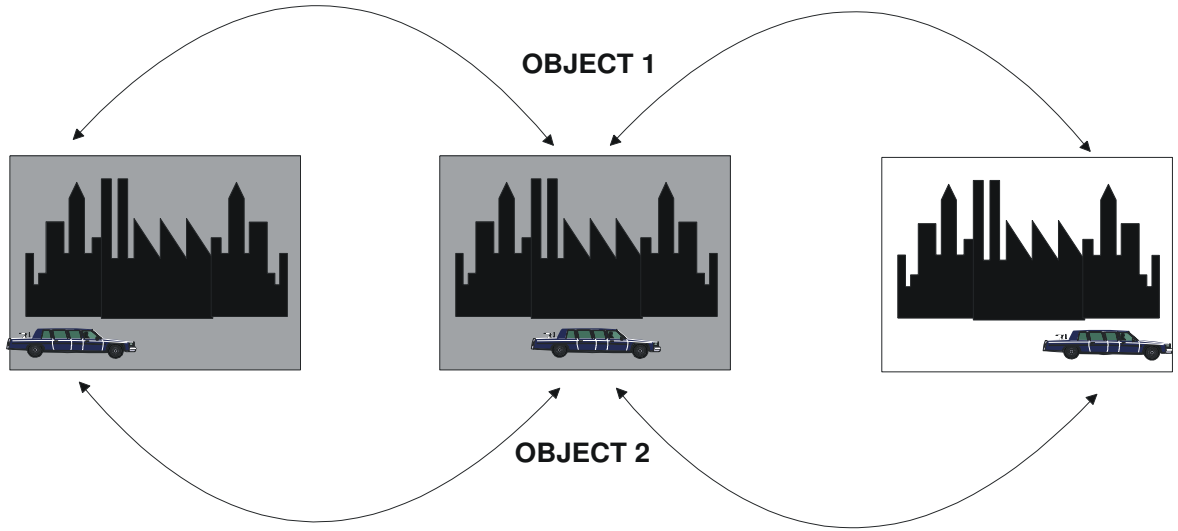
MPEG-4 (1)

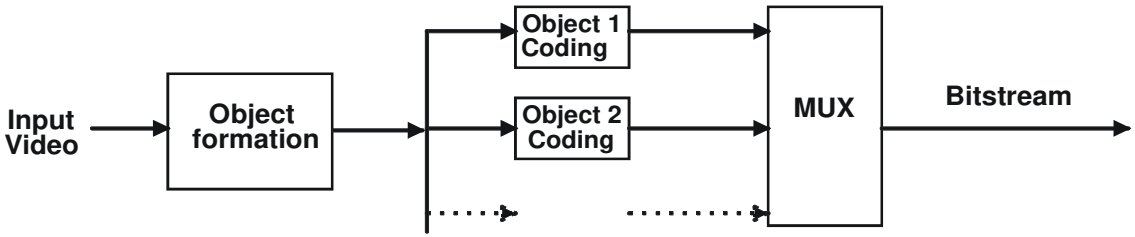
Originally ISO and ITU-T had planned a standard MPEG-3 for HDTV at very high data rates. This work was later integrated into MPEG-2. This explains why there is no MPEG-3 standard.

MPEG-4 was originally planned for video at very low bandwidths (e.g., for wireless PDAs). Later the ISO committee decided to concentrate on an entirely new technology, namely encoding in the form of **sets of objects overlaid to form an image**. The encoding technique can be chosen separately for each object. This object-oriented encoding also opens up much richer possibilities for processing on the receiver side.

MPEG-4 (2)

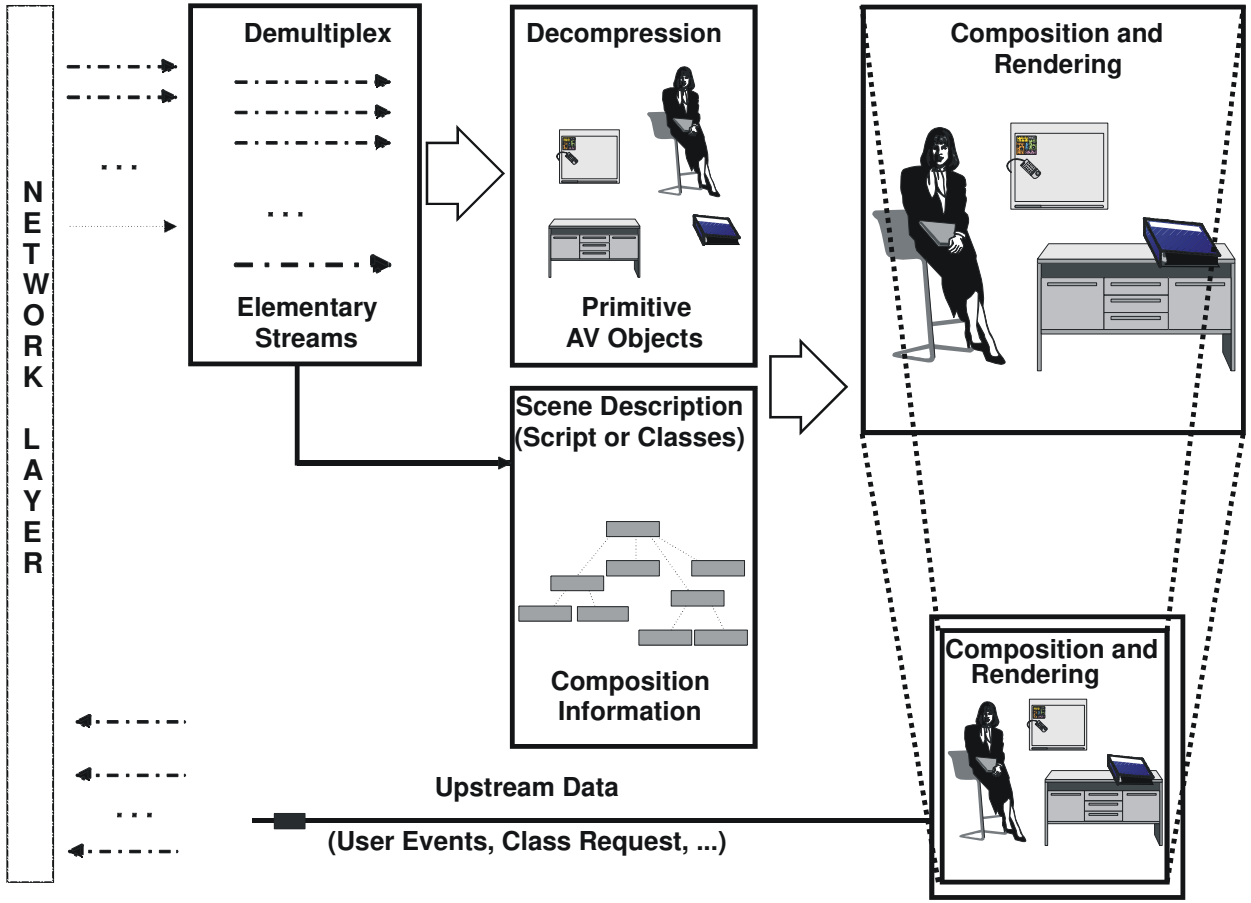
Separate encoding of background and foreground. The background is static.





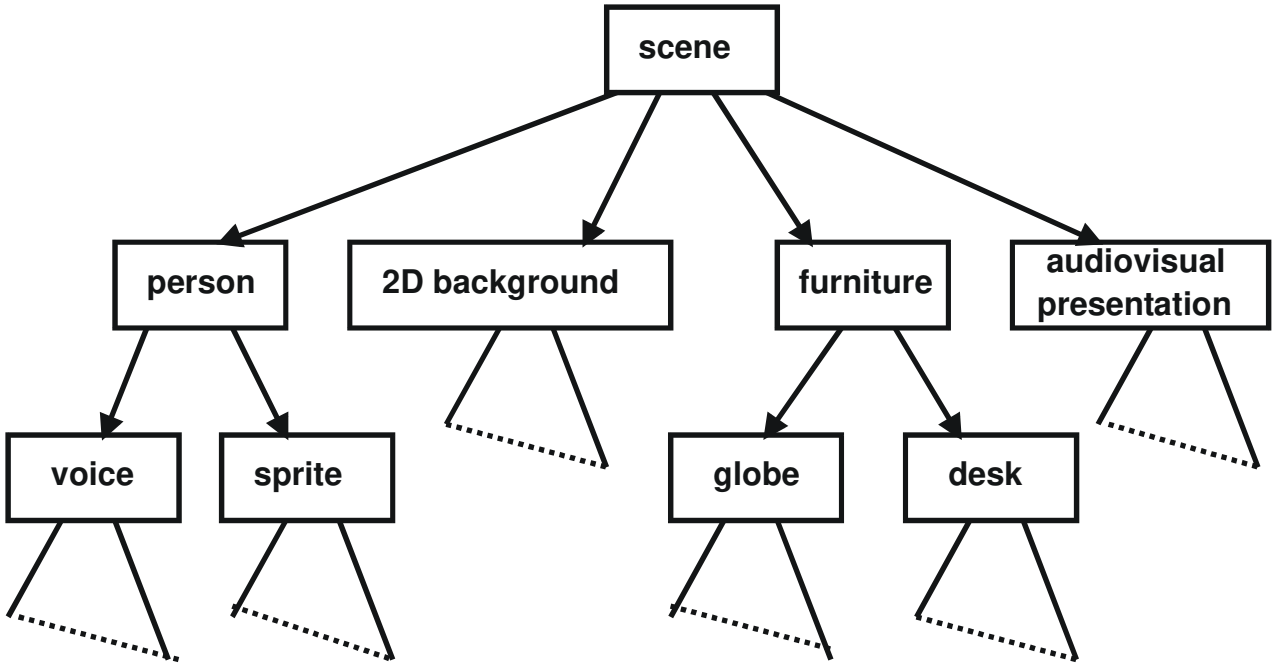
MPEG-4 (3)

Decoding of an MPEG-4 system stream



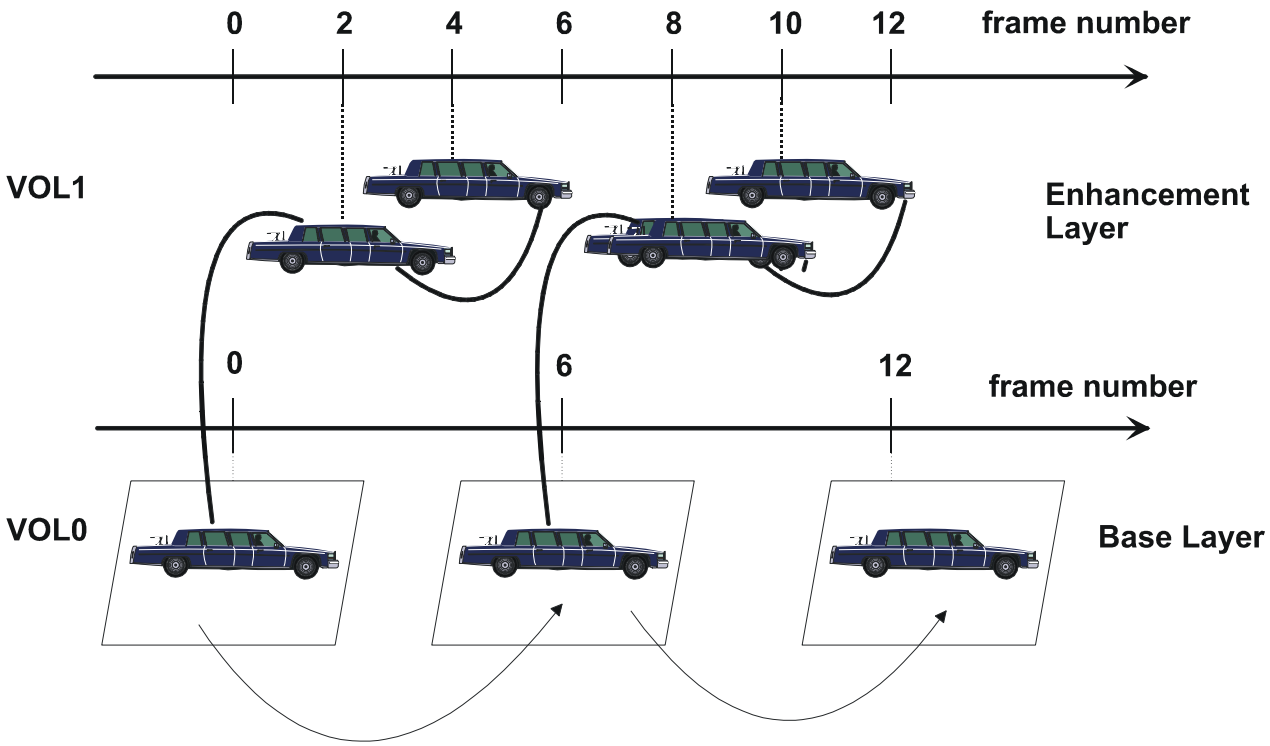
MPEG-4 (4)

Object hierarchy for our decoding example



MPEG-4 (5)

Scalability by „layered encoding“ in MPEG-4



2.3.2 ITU Recommendation H.261

Also known as „p*64 kbit/s“

- A video coding technique for video data at p x 64 kbit/s.
- Originally developed for ISDN
- Parameter p is chosen as follows: p in [1,30]
- p small implies low image quality at low data rates. An example is video telephony with p=1 or p=2.
- p larger implies better video quality at higher data rates. Typical is p=6 for company video conferencing over 6 parallel ISDN B-channels.
- **Intraframe-Coding:** based on the DCT. Very similar to JPEG but there is only one quantization factor for all values of the block (no quantization table).
- **Interframe-Coding:** very similar to the P frames in MPEG-1.
- There are no B frames in H.261.

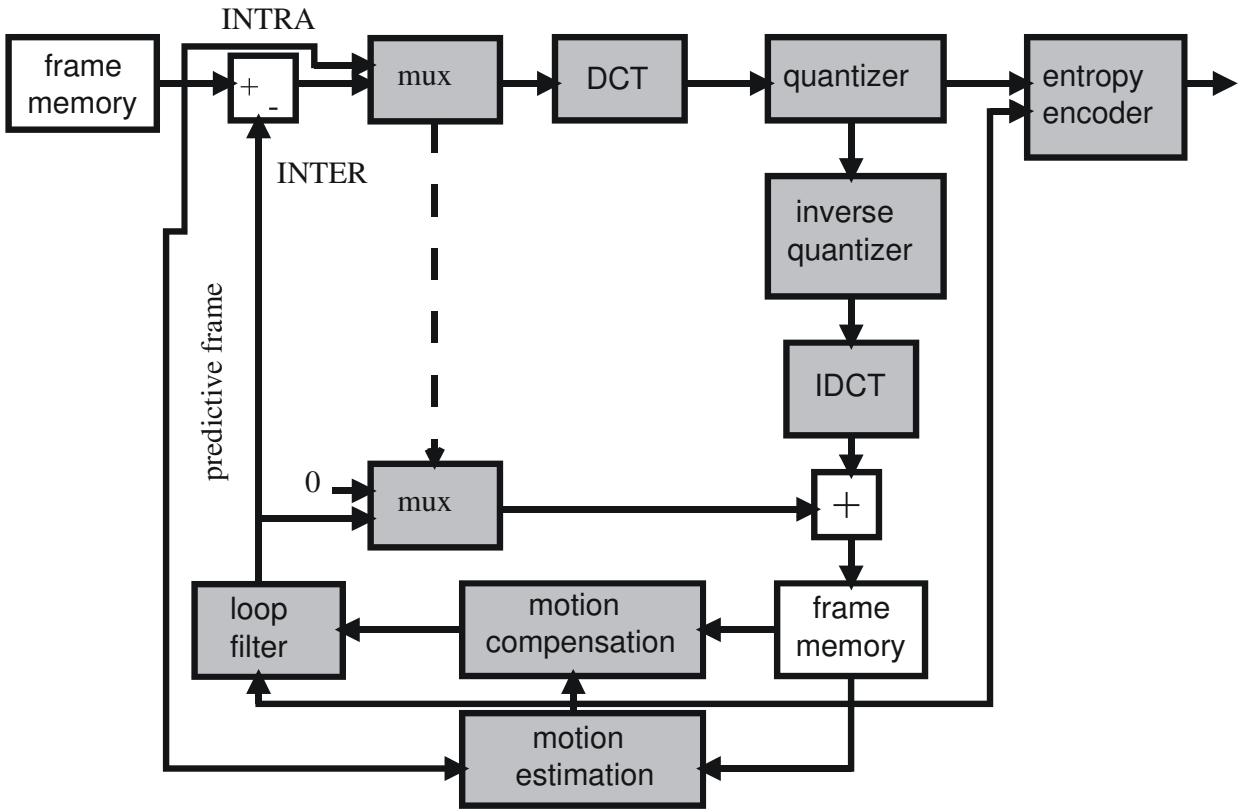
Important Parameters of H.261

	CIF (width x height)	QCIF (width x height)
Y	352 x 288	176 x 144
C_b	176 x 144	88 x 72
C_r	176 x 144	88 x 72

Hierarchy of the structuring elements of the data stream:

Structure Element	Description
picture	a full frame
group of blocks	33 macroblocks
macro block	16 x 16 Y, 8 x 8 C_b, C_r
block	8 x 8 pixels (unit for the DCT)

The H.261 Encoder



Status of H.261

Very widely used in practice, many products available in the market from many manufacturers. Has replaced earlier proprietary standards for video telephony.

Pure software implementations are available as well as stand-alone hardware solutions (“black boxes“) and combined solutions, mainly for the PC.

H.263

H.263 is the successor standard of H.261 at ITU-T, incorporating much of the experience gained with MPEG-1.

Some differences between H.263 and H.261 are:

- There are five image sizes instead of two.
- There is a bi-directional interpolation where exactly one B frame follows each P frame.
- There are negotiable options that allow to tailor the algorithm for specific applications. For example, arithmetic coding can be chosen instead of run-length/Huffman coding in the entropy encoding step.