Multimedia

§9 Video-Compression

Prof. Dr. Georg Umlauf

Content

§9.1 Basics of motion perception

§9.2 Code-Formats

§9.1 Basics of motion perception

Perception of motion

- Complex function of the human sensors.
- Dependent on **physiological factors**
  - Interdependence of adjacent light sensory cells in the retina.
  - Foveal object tracking (auto-tracking of the eye).
  - Vergence and accommodation.
- Dependent on **psychological factors**
  - Elimination of uniform motions by the brain.
  - Linked to motion and acceleration perception (vestibular system).

Content

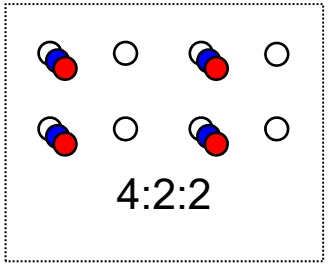
§9.1 Basics of motion perception

§9.2 Code-Formats

§9.2 Code-Formats

PAL (Phase Alternation Line)

- Analog video technique:
  - Row-wise sampling of single frames
  - Serializing of pixels using line interlacing
- Properties:
  - Video refresh rate 25 Hz
  - Image resolution 720×576 pixel
  - Color depth: 8 bit
  - Color space subsampling: 4:2:2
  - Interlacing: Two interlaced half-imaged every 1/50s.
  - Band width: $\underbrace{(720\times576\times25\times8)}_{\text{Luminance pixel}}+2\times\underbrace{(360\times576\times25\times8)}_{\text{Chrominance pixel}} = 166\text{Mbps}$



HDTV (High Definition Television)

- Properties:
  - Video refresh rate up to 60 Hz
  - Image resolution e.g. 1920×1080 pixel
  - Color depth: 8 bit
  - Color space subsampling: 4:2:2
  - Band width: $\underbrace{(1920\times1080\times60\times8)}_{\text{Luminance pixel}}+2\times\underbrace{(960\times1980\times60\times8)}_{\text{Chrominance pixel}} = 1{,}99\text{Gbps}$

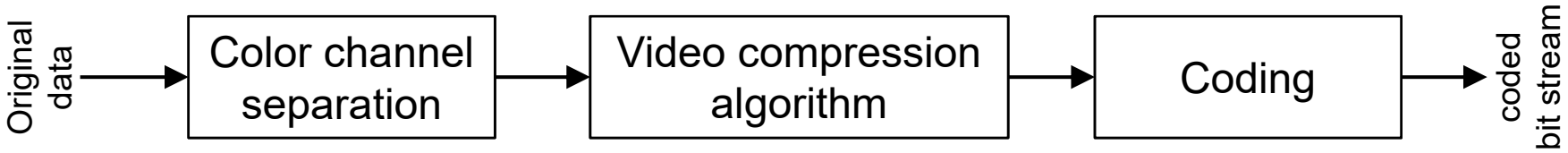
Compression of video signals

- Video sequences contain usually much redundant information
  - **Spatial redundancy:** Intra-frame-compression, i.e. within one frame
  - **Temporal redundancy:** Inter-frame-compression, i.e. between subsequent frames.

➡ High data reduction for video coding.

➡ Grouping: Consider groups of images (GOP), that are compressed together.

➡ General approach:

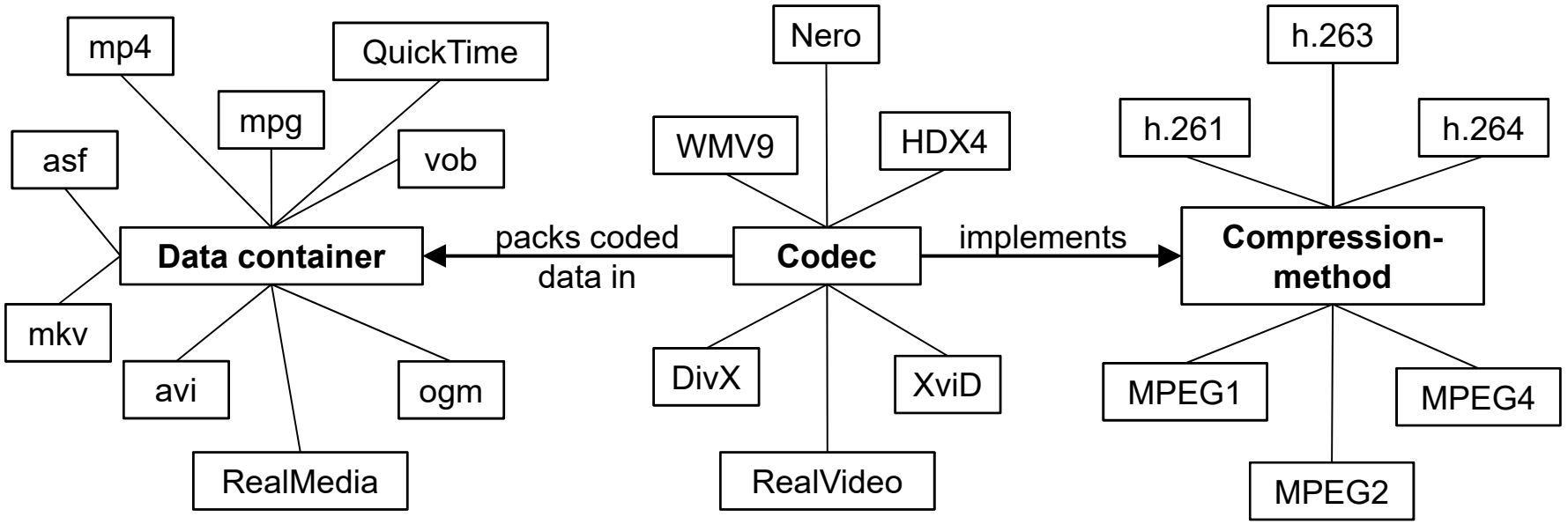


Video compression algorithm

- simple: limited color table, RLE, etc.
- **interpolating:**
  - Store only a subset of the pixel and compute the missing pixels to restore the complete image via interpolation.
  - Storage of differences of the original image to the interpolated image.
  - Application: many image areas change over time only marginally.
- **predictive:**
  - Background is static, only one object in the foreground moves.
  - ➡ **Objects and object motions** must be detected.
  - Cause for a change in the image of a video sequence is often a motion of the camera (translation, rotation, scaling).
  - ➡ **Motion compensation** eliminates motions of the camera.
- Transformation: e.g. DCT
- static: e.g. Huffman-coding

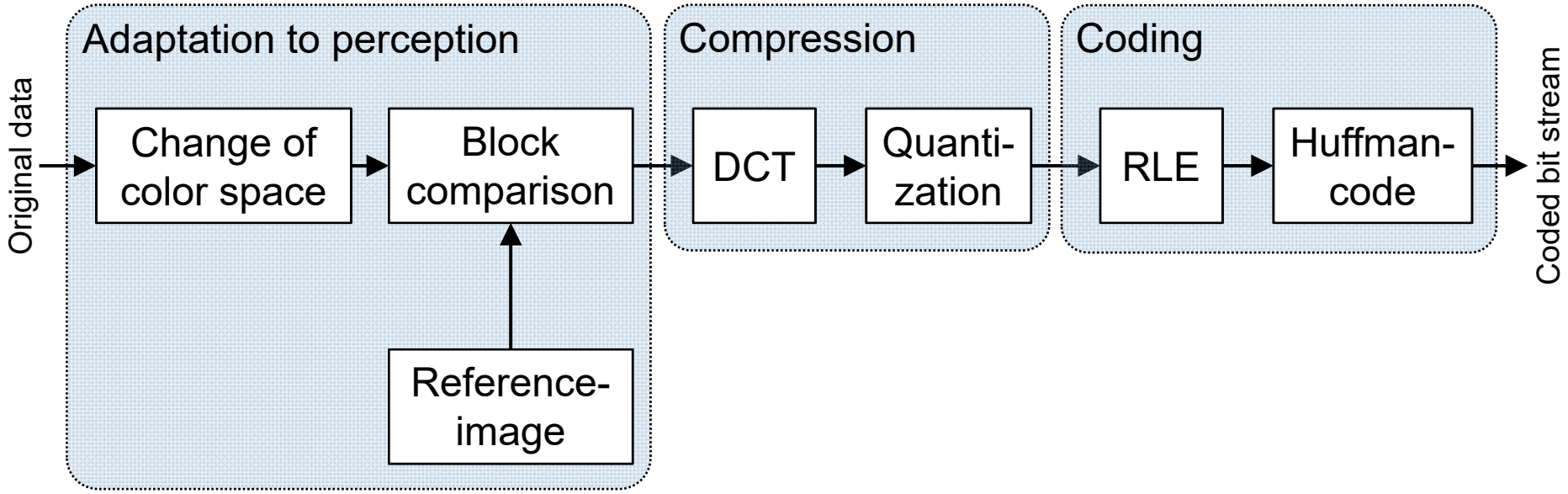
Video compression as defined by MPEG

- Codecs and compression methods



MPEG-1 (1)

- Principal structure



MPEG-1 (2)

1. **Change of color space:** Conversation to YCrCb (4:2:2, 4:2:0)
2. **Block comparison:**
   a) **Blocking** of single images in slices of macro-blocks
      - 16×16 pixel luminance (macro-block)
      - 8×8 pixel chrominance (block)
   b) **Prediction of motion**
      - Compare luminance-macro-blocks of subsequent frames.
      - Change of position of luminance-macro-blocks are coded as vectors.
      - Quality depends on size of search space.

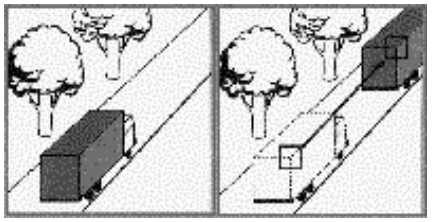
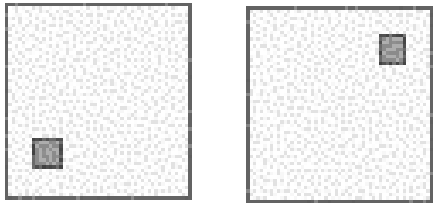
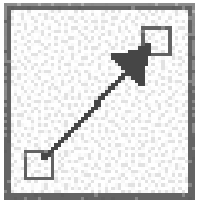


Image n and image $n+1$



Moved object



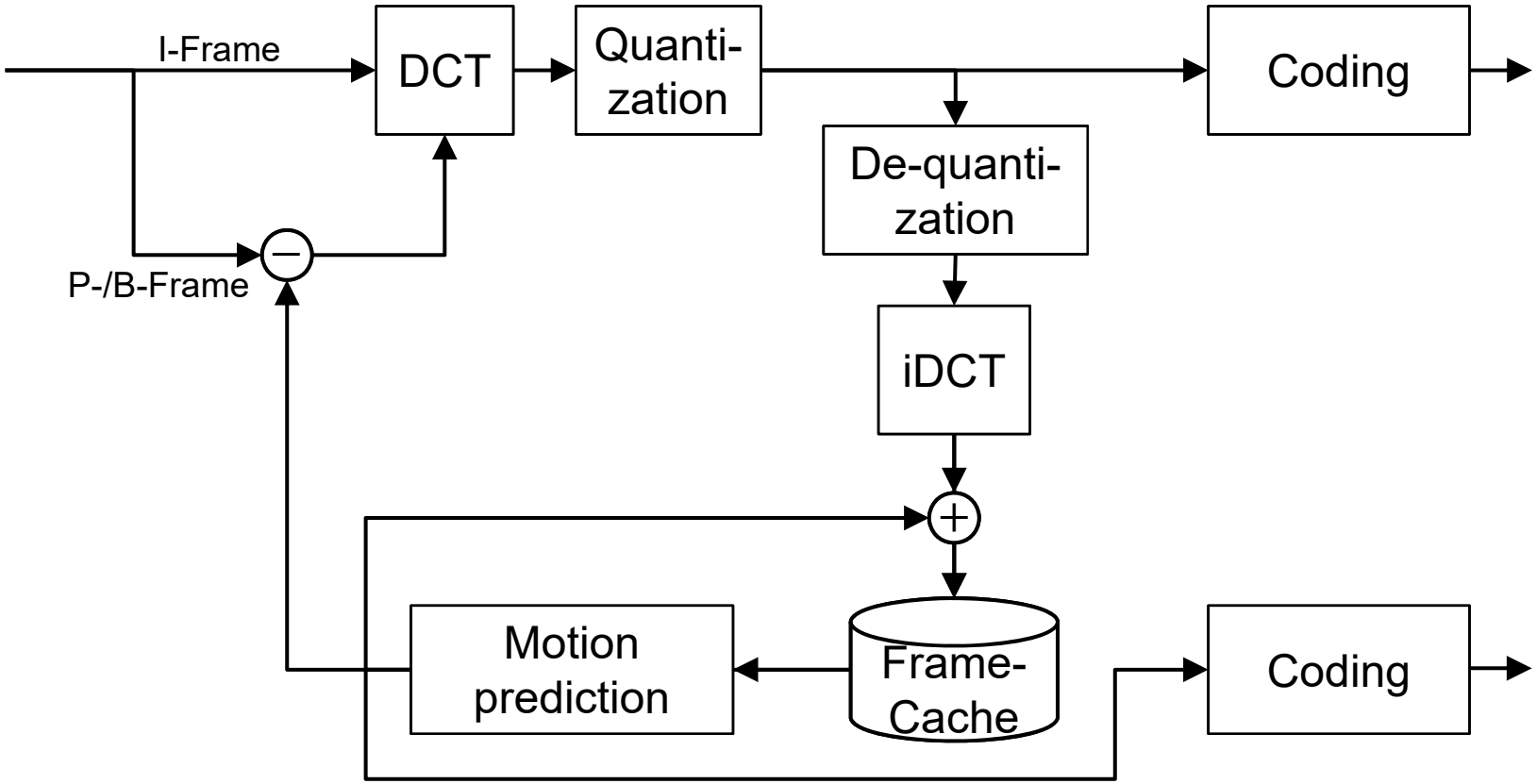
Vector of motion

MPEG-1 (3)

3. **DCT** of different image types
   - **I-Frame:**
     - complete single image (**intra-picture**),
     - compressed as a still image,
     - no motion prediction,
     - first image of a sequence of images (at least every 12th frame),
     - reference image for subsequent images.
   - **P-Frame:**
     - Uses motion prediction,
     - relative to the previous I-Frame (**predicted picture**).
   - **B-Frame:**
     - Uses motion prediction,
     - relative to the previous and subsequent P- or I-Frame (**bi-directional picture**).
   - Order and repetition of I-/P-/B-Frames is arbitrary.

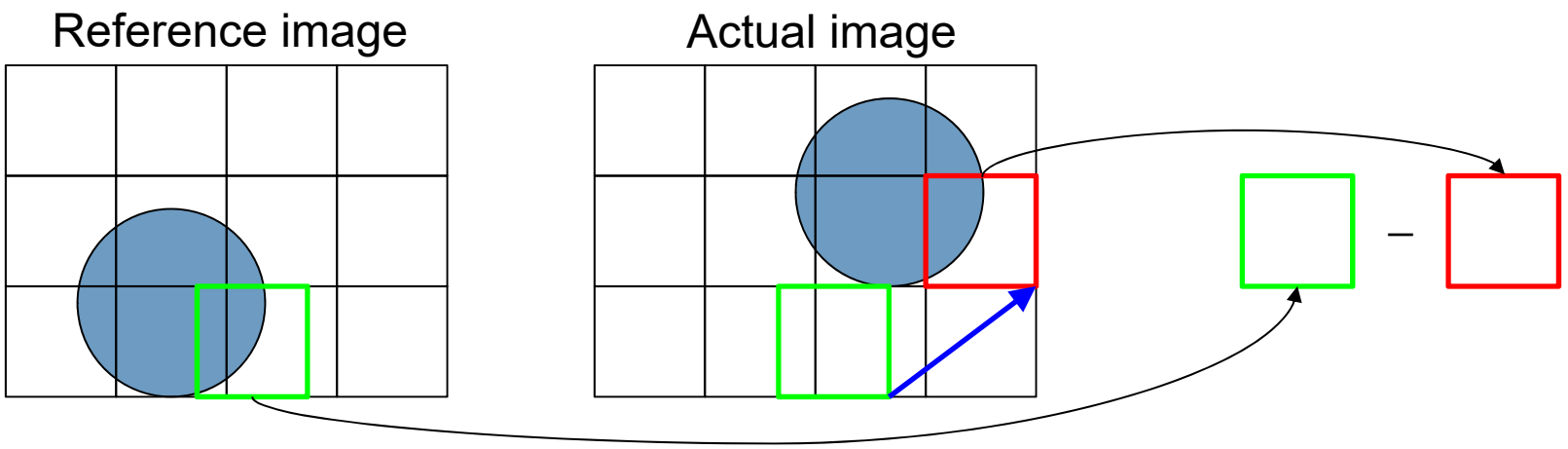
MPEG-1 (4)

I-/P-/B-Frame-Coding



MPEG-1 (5)

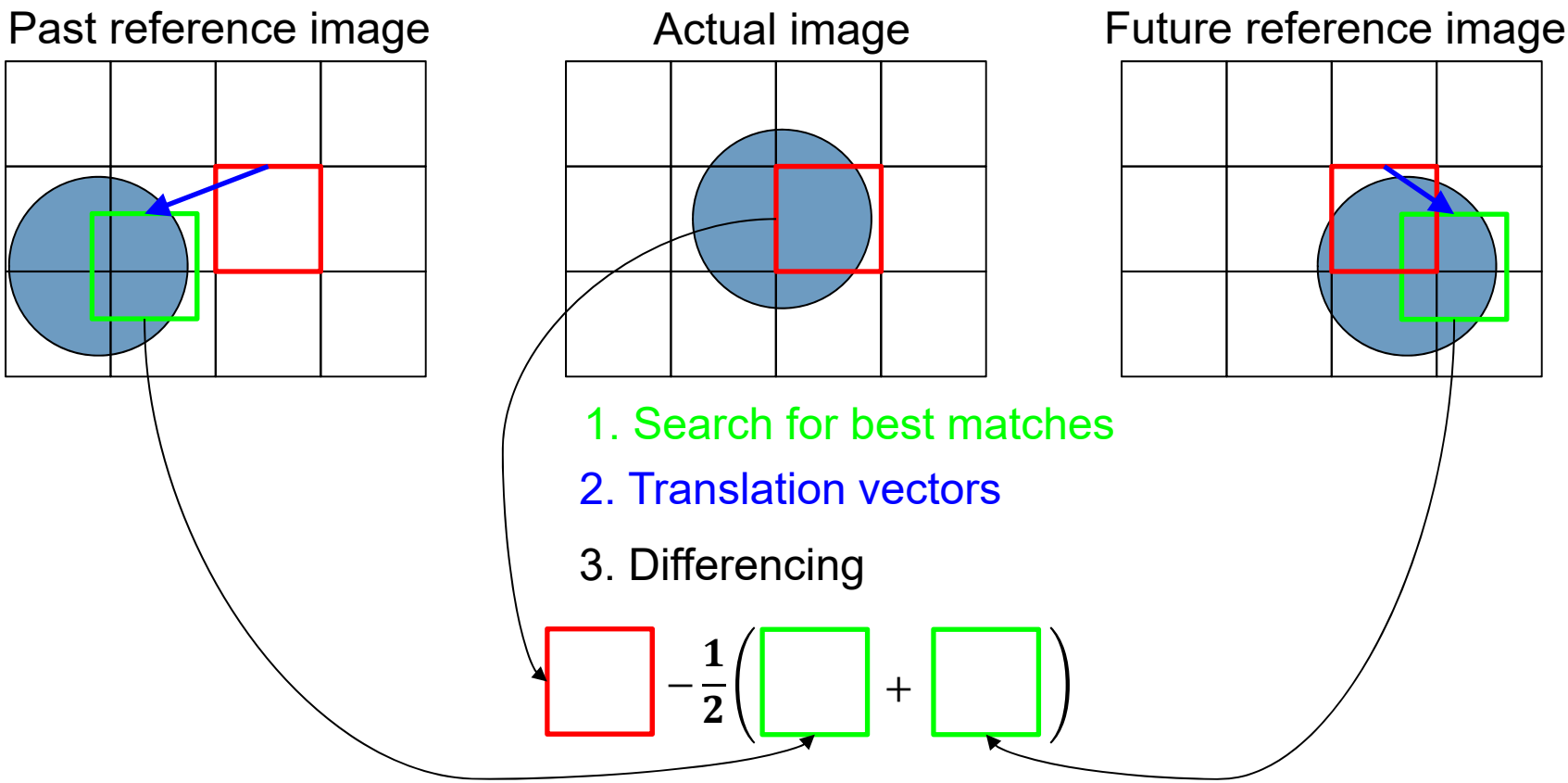
P-Frame-Coding



1. Search for best match
2. Translation vector
3. Differencing

MPEG-1 (6)

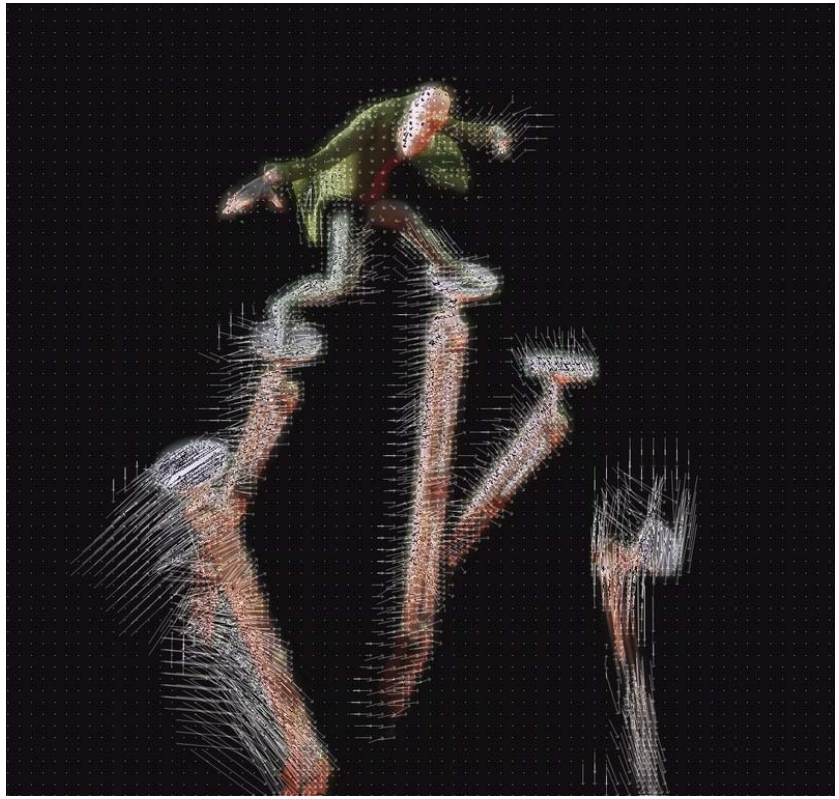
B-Frame-Coding



MPEG-1 (7) Example: Motion prediction

Elephants Dream, Blender, 2006.

MPEG-1 (8) Example: Motion prediction



Source: Wikipedia

Motion vectors of a single frame from *Elephants Dream.*

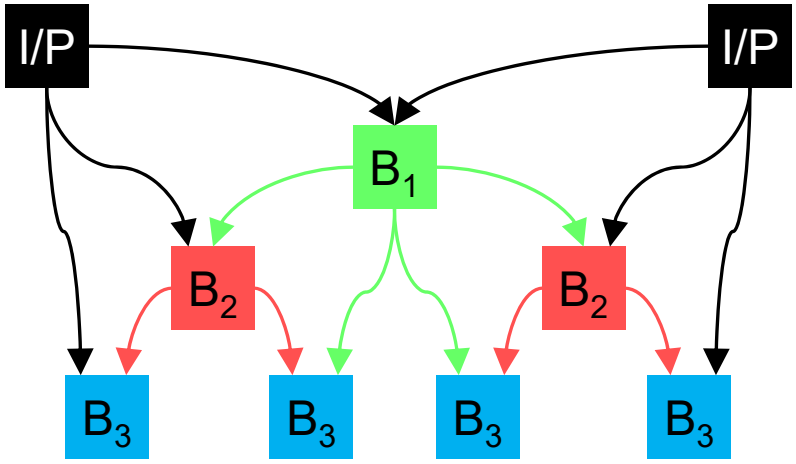
MPEG characteristics (1)

- **MPEG-1**
  - Band width: 1.25 MBit/s Video + 2 Audio channels, i.e. ≤1.82 MBit/s
  - Resolution: 360×288×25 (CIF Europe), 352×240×30 (CIF USA)
  - No interlacing (➡ TV-recordings)

- **MPEG-2**
  - Band width: 2 MBit/s – 80 MBit/s
  - Resolution: up to HDTV (1,920×1,080) theoretically up to 16,383×16,383
  - Subsampling: 4:4:4 or 4:2:2
  - Interlacing possible
  - Compression rate up to 60:1 possible.

MPEG characteristics (2)

- **MPEG-4 AVC, H.264, ISO-14496**
  - Interactive multimedia applications
  - Decomposition of a scene
    - in different audio- and video-media-objects,
    - that are compressed separately and
    - have each their own spatial/temporal position and dimension.
  - Hierarchical B-Frames.
  - Transform blocks for camera panning.
  - Decomposition of a scene in foreground and background and separate compression.



MPEG extensions (1)

- **MVC (multi-view coding)**
  - Extension to mpeg4.
  - **Stereo interlacing:** spatial multiplexing of different views to form one picture (top-bottom, side-by-side, row-wise, column-wise, checkerboard).
  - **Inter-view prediction:** similar to motion prediction but in the spatial dimensions between different views.

MPEG extensions (2)

- **HEVC (high efficiency view coding), H.265, MPEG-H**
  - Enhancement of mpeg4.
  - Block size between 4×4 to 64×64.
  - Four sizes for transform blocks.
  - Current improvement in compression rate at 20%..
  - Many patent claims…

Goal

- What are temporal and spatial redundancy for video data?
- What are the main building blocks of MPEG1-video-coding?
- What are I-, P-, B-Frames?
- What is motion prediction?