Principles and Standards for Video Coding

Fachhochschule Osnabrück
University of Applied Sciences

Prof. Dr. Karsten Morisse
kamo@fhos.de
http://www.et.fh-osnabrueck.de/~kamo

Part I:

Principles for video coding

Overview

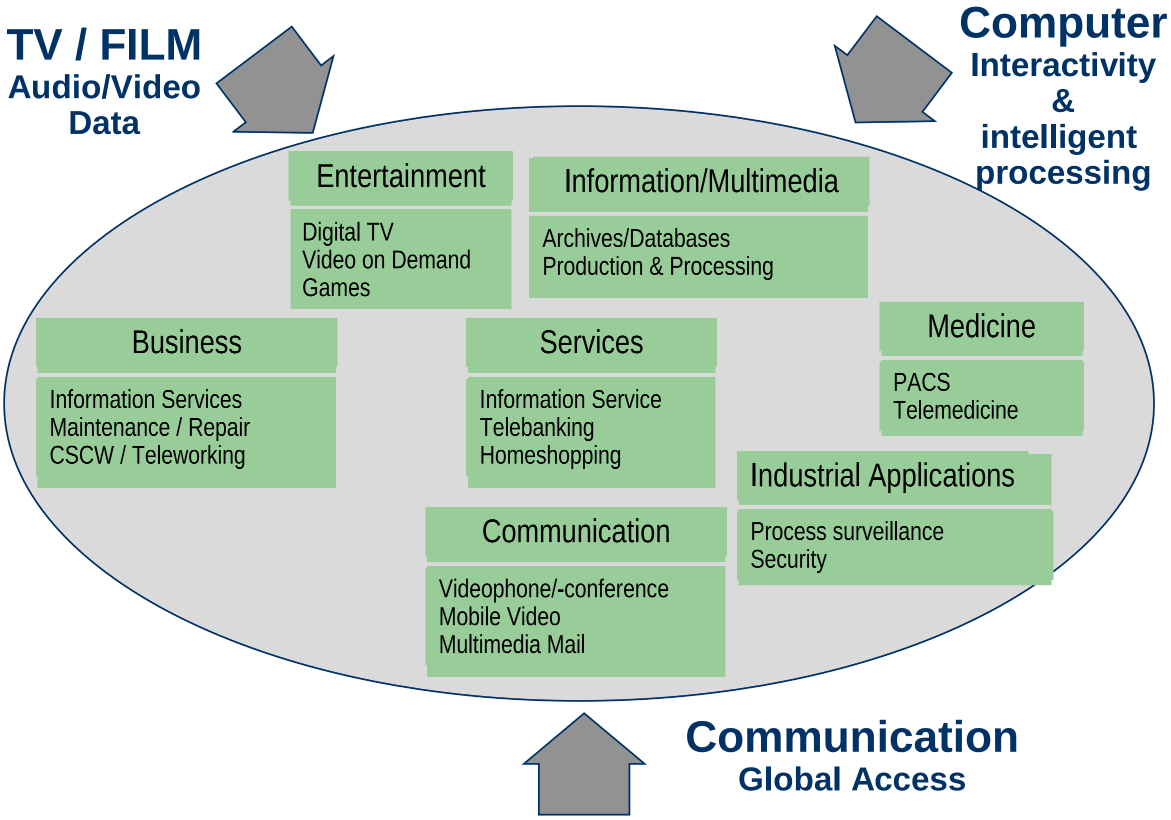
- Scenarios for Multimedia Applications
  - Motivation
  - Requirements
- Coding Principles
  - Redundancy
  - Irrelevancy
- Compression techniques
  - Quantization
  - Entropy Coding
  - Prediction Coding
  - Intraframe Coding
  - Transformation Coding
  - Hybrid Coding
- Video coding - Standards
  - JPEG
  - H.261, H.263
  - MPEG-1, MPEG-2, MPEG-4, MPEG7

Literature

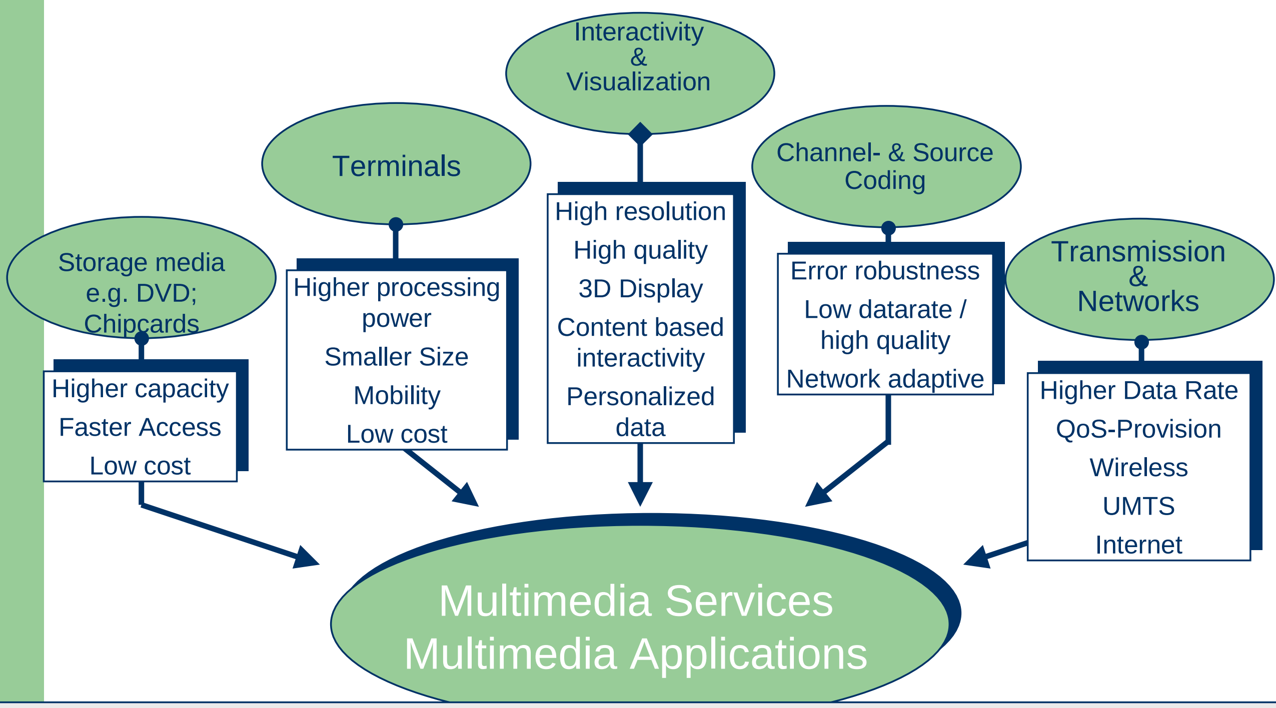
- R.Steinmetz: *Multimedia-Technologie*, 2. Auflage, Springer-Verlag, 1998.
- C.Wu, J.Irwin: *Emerging Multimedia Computer Communication Technologie*, Prentice Hall, 1998.
- B. Heinrichs: *Multimedia im Netz*, Springer-Verlag 1996.
- S. Heath: *Multimedia and Communications Technology*, 2nd edition, Focal Press, 1999.
- B. Furth et al.: *Video and Image Processing in Multimedia Systems*, Kluwer Academic Publishers, 1995.
- J.D.Gibson et al.: *Digital Compression for Multimedia*, Morgan Kaufmann Publishers, 1998.
- W.Effelsberg, R. Steinmetz: *Video Compression Techniques*, dpunkt Verlag, 1998.
- K. Jack: *Video Demystified*, 2nd Edition, HighText Interactive, 1996.
- R. Mäusl: *Fernsehtechnik*, 2. erw. Auflage, Hüthig Verlag, 1995
- H. Benoit: *Digital Television*, Arnold, 1997
- J. Mitchell, W. Pennebaker, C. Fogg, D. LeGall: *MPEG Video Compression Standard*, Chapman&Hall, 1996
- F. Jones: *Desktop Digital Video Production*, Prentice-Hall, 1999
- S.J. Huang: *Variable Bitraten-Codierung*, FKT, Nr. 10, 1999
- R. Koenen: Overview of the MPEG-4 Standard (http://cselt.it/mpeg)
- M. Ghanbari: *Video Coding – an introduction to standard codecs*, IEE, 1999

Scenarios for Multimedia Applications

Applications: Overview / Trends

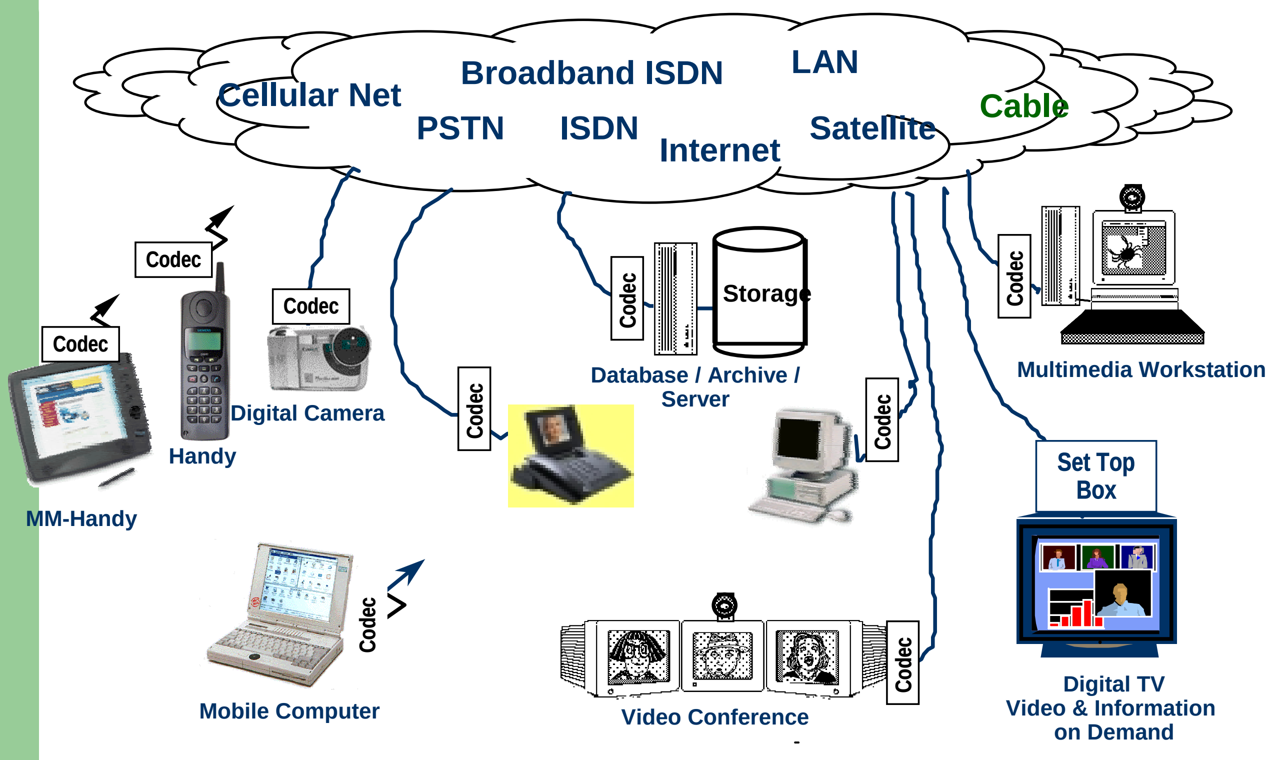


Technology Trends in Multimedia



Boundaries between Communication, Computer & TV, Entertainment disappear and favour new applications:
Mobile Communication, Broadband Access Networks, Interactive TV, DVB, Internet Access, E-Commerce,...

Multimedia Environments



Network Requirements for Multimedia

	Distribution	Retrieval	Messaging	Dialog
Application / Service	• Τς; Ραδιο; • Παψ Τς • Ωεβ Τς • Βυσινεσσ Τς	• Ωεβ Αχχεσσ • ςοΔ • Ε−Χομμερχε • Τραινινγ/σερϖιχε • Ιντεραχτιϖε Τς	• Μυλτιμεδια μεσσαγινγ • ε−μαιλ	• ςιδεοπηονε • ςιδεοχονφερενχε • Συρϖειλλανχε
Traffic	Unidirectional Realtime Streaming	Bidirectional Asymmetric Non realtime Streaming	Bidirectional Asymmetric Non realtime Streaming	Bidirectional Symmetric/ Asymmetric Realtime
User Terminal	TV-Set TV STB PC	PC, (TV) Special terminals Mobile terminals	PC, (TV) Special terminals Mobile terminals	PC, (TV) Special terminals Mobile terminals
Typical Datarates	Video: 1 – 8 Mbit/s Audio: 64–384 kbit/s	10 Mbit/s peak 64 kbit/s – 1Mbit/s	64 kbit/s–1Mbit/s	64 kbit/s–1Mbit/s
QoS	High Resolution High quality	High to reduced resolution/quality	Reduced quality	Reduced quality
Bit error rate	10^{-8} – 10^{-10}	10^{-4} to 10^{-6}	0 (ARQ)	10^{-4} to 10^{-6}
Delay / Jitter	Extremely low	Non critical	Non critical	Very low; critical
Network Protocol	MPEG TS (ATM, UDP/IP)	UDP/IP;IETF ITU-H.32x	TCP/IP; IETF ITU-H.32x	UDP/IP, IETF ITU-H.32x
Network	Terrestrial Antenna Satellite Cable Network	Internet xDSL Mobile (UMTS) ISDN	Internet, xDSL Mobile (UMTS) ISDN	Internet, Mobile (UMTS) ISDN

Video Requirements for MM-Applications

Image Type	Typical Applications	Typical Bitrates	Quality Demands	Compression Standard
Photographic Still Images	Studio Press Desktop	< 24 bit/pixel 1 - 2 bit/pixel < 1 bit/pixel	High High Medium	PCM JPEG JPEG
Video	TV Studio	166 Mbit/s	High	PCM
	TV Broadcast	2 - 6 Mbit/s	Medium	MPEG-2
	Video on Demand	< 5 Mbit/s	Medium	MPEG 1 & 2
	Desktop Video	384 kbit/s to 2 Mbit/s	Medium	MJPEG H.263 MPEG 1 & 2
	Videophone & Videoconference	20 kbit/s to 1 Mbit/s	Low to Medium	H.261/H.263
	Internet-Terminal	5 kbit/s to 1 Mbit/s	Low to Medium	MPEG 4

Coding Principles

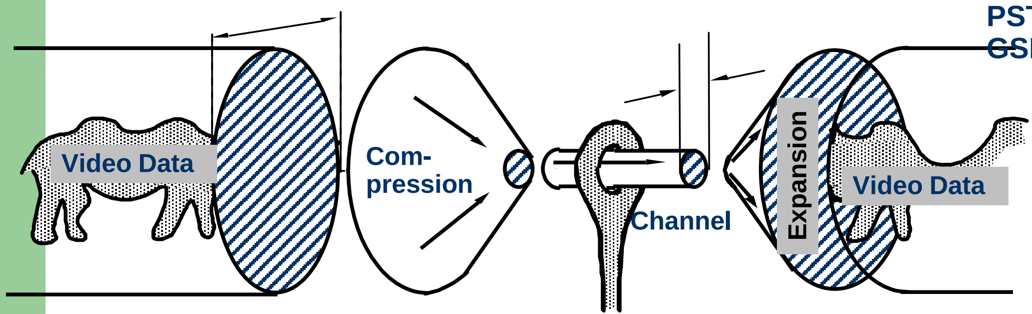
Overview

- Scenarios for Multimedia Applications
  - Motivation
  - Requirements
- **Coding Principles**
  - **Redundancy**
  - **Irrelevancy**
- Compression techniques
  - Quantization
  - Entropy Coding
  - Prediction Coding
  - Intraframe Coding
  - Transformation Coding
  - Hybrid Coding
- Video coding - Standards
  - JPEG
  - H.261, H.263
  - MPEG-1, MPEG-2, MPEG-4, MPEG7

Digital Image Transmission

ATM	**>100**	**Mbit/s**
LAN	**10-100**	**Mbit/s**
ISDN	**px64**	**kbit/s**
PSTN	**28**	**kbit/s**
GSM	**8**	**kbit/s**

Digital Video CCIR-601 166 Mbit/s



Camel and pinhole problem

Goal: Data reduction of digitized images and videos

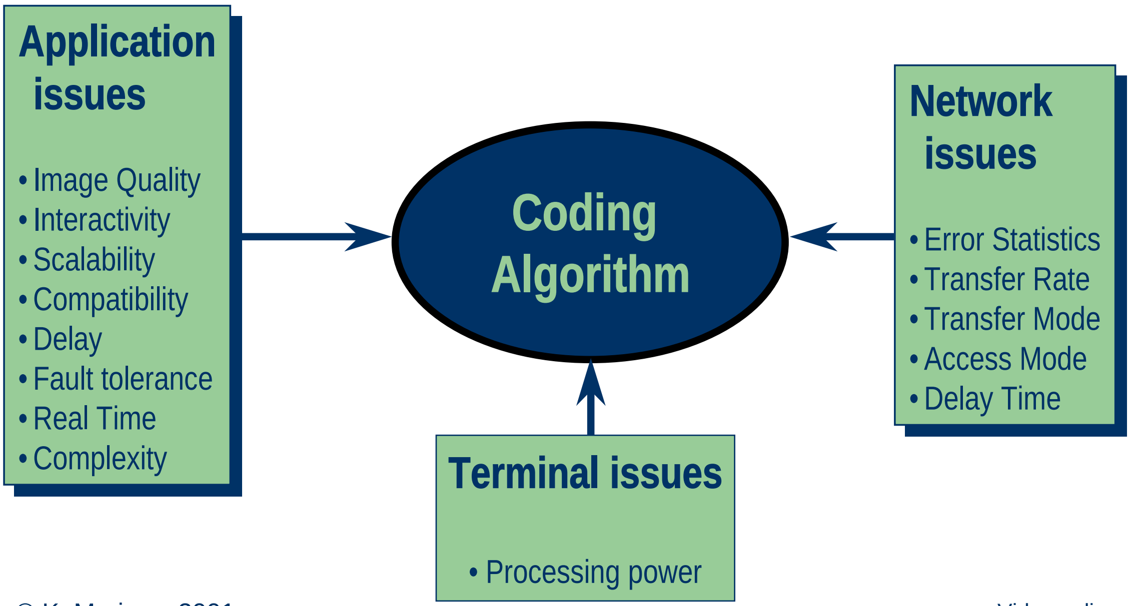
Constraints: Image quality adapted to application
- No loss of relevant information
- Tolerable impairment of image quality

Benefit: Economical usage of image data
- Fast transmission over available channels
- Economical storage

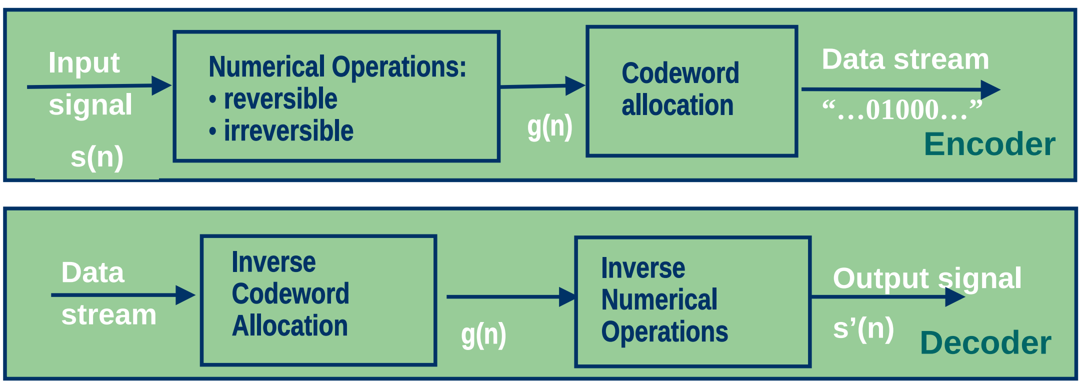
Applications: Videocommunication and Multimedia
- Real-time transmission (Live)
- Image storage and archiving

Requirements for video coding systems

Application, Network and Terminals define the Requirements for a Coding System



Coding process (simplified)



Goal: Find Code-Wort-Allocation for input signal with
- Minimal number of bits
- Unique reconstruction
- Best similarity between input signal and reconstructed signal

Problem: Find optimal numerical operations and codeword tables

Important features of input signal:
- Signal resolution: spatial, temporal, grey & color values
- Signal statistic: number and distribution of samples and values

Typical Compression ratios

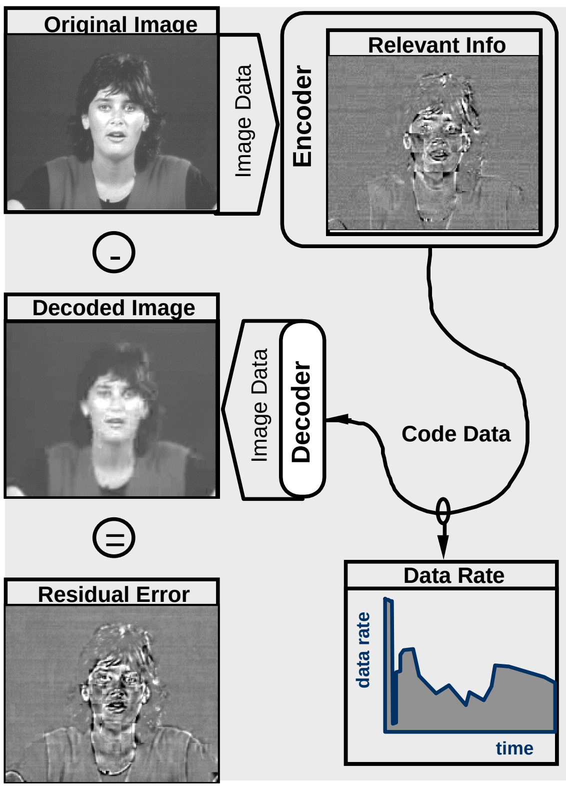
Image format	CCIR 601	CIF	CIF	QCIF
Image size	720x576;50Hz	352x288;25Hz	352x288;12.5Hz	176x144;8.3Hz
Data rate in Mbit/s	166	31	15	2.5
Subsampling	1:1	1: 5	1:11	1: 66
Compression	1:83	1:81	1:120	1:40/90/312
Total Compression	1:83	1:400	1:1320	1: 2640/5960/21000
Transmission rate	2 Mbit/s	384kbit/s	128 kbit/s	64/28 / 8 kbit/s
Network	ATM	ISDN	ISDN	ISDN/PSTN/GSM

Principles of image compression

- **Redundancy:**
  Information that is known from spatial or temporal similarities or statistical dependencies
  ➡ Reversible Compression
  Losless Compression
  (factor ~2)

- **Irrelevancy:**
  Information that is not visible or not important for the receiver
  ➡ Irreversible Compression
  Lossy Compression
  (Factor ~ 40)

Compression algorithms remove redundancies and extract relevant information



Redundancy

Removing Redundancy

Principle:
- exploiting temporal dependency
  likeness of successive frames in a videosequence
- exploiting spatial dependency
  dependencies of neighboured pixel values
- exploiting statistical dependency of coded values

Entropiecoding:
- **Runlength Coding**
- **Huffman Coding**
- **Arithmetic Coding**
- **Bit plane Coding**

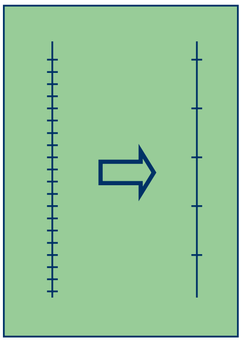
Decorrelation:
- **Transformation (DCT)**
- **Subband Coding**
- **Wavelets**

Prediction:
- **2-D DPCM**
- **Temporal DPCM**
- **Motion-compensation**

Irrelevancy

Removing Irrelevancy

Features:
- Resolution reduction: removing of information which is unimportant for receiver
- exploitation of limitations of human visual system



Methods:
- Thresholding
- Temporal subsampling
- Spatial subsampling
- Quantization of coefficients
- Vector quantization

Removing irrelevancy produces errors!

can be shown in difference picture

⇒ Removing information only where no damages occurs

Human visual system

- Resolution limit
  - Eye has only a limited resolution capacity: information with frequencies higher than the limit of the eye is not recognizable
- Mach-Effect – Contrast recognition at edges
  - Eye is optimized/trained on recognition of contours
  - Main picture information not in regular surfaces but at edges
  - Visual system enhances contrast changes at edges automatically
  - Thus: coarse quantization of edges
- Oblique Effekt – Perception of horizontal and vertical structures
  - Eye is optimzed for recognition of horizontal and vertical structures
  - Diagonal structures are recognized with lower resolution
- Masking effects
  - Background with irregular patterns covers embedded objects
  - Homogeneous afor noise or blocking artefacts
  - Thus: hide coding errors in areas with irregular structure
  - Masking effect is not valid for moving objects
- Resolution of moving objects
  - still objects are recognized much sharper than moving objects
  - slow motion can be shown with low temporal resolution
  - fast motion needs better temporal resolution
- Color resolution
  - Eye resolves color (chrominance) much worse than brightness (luminance)
  - Pixel resolution for color can be lower than for grey value intensity

Compression Techniques

Overview

- Scenarios for Multimedia Applications
  - Motivation
  - Requirements
- Coding Principles
  - Redundancy
  - Irrelevancy
- **Compression techniques**
  - **Quantization**
  - **Entropycoding**
  - **Predictioncoding**
  - **Intraframe Coding**
  - **Transformation Coding**
  - **Hybridcoding**
- Video coding - Standards
  - H.261, H.263
  - MPEG-1, MPEG-2, MPEG-4, MPEG7
  - JPEG

Classes of Coding Systems

- Entropy coding
  - Lossless compression: no loss of data during encoding and decoding, i.e. the source data and the decoded data are identical
  - Can be applied to different media without considering media specific characteristics or the semantics of the data
- Source coding
  - Data are analyzed and the semantics is considered during encoding process
  - Distinction between relevant and irrelevant information (lossy compression)
  - Coding algorithms try to use specific features of the source data
- Hybrid coding
  - Mixture of entropy and source coding

Characteristics of Coding Systems

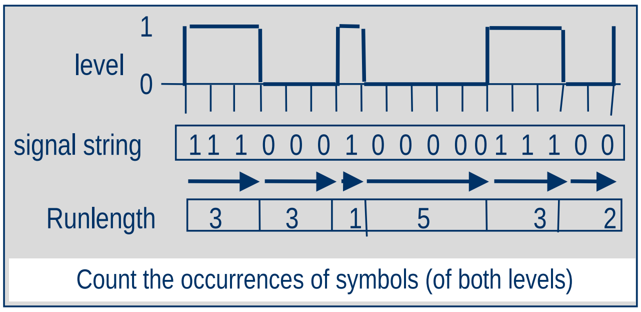
Characteristics	Description
lossless	original data are identically reconstructed
lossy	reconstructed image is degraded compared to original
intraframe	temporal correlation between frames is not considered
interframe	correlation between different frames is considered
symmetric	coding and decoding need about same time
asymmetric	coding needs more time than decoding
realtime	coding is synchronous with source signal
scalable	coded images are in several resolution steps available

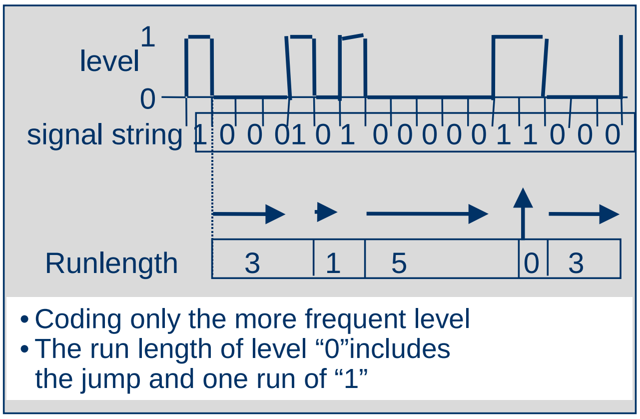
Lossless Compression

- Universal algorithms
  - Work without knowledge of data (e.g. statistic distribution of symbols)
  - Examples
    - Run Length Encoding
    - Static Pattern Substitution
- Statistic algorithms
  - Use statistical characteristics of the data (Symbol distribution)
  - Use VLE (Variable Length Encoding):
    - Frequent symbols $\rightarrow$ short code words
    - Rare symbols $\rightarrow$ long code words
  - Examples: Huffman-Coding, Arithmetic Coding, Shannon/Fano-Coding
- Table based algorithms
  - Generate tables of read strings
  - Substitution principle:
    - Find an input pattern in the table and substitute it with the table index
  - Examples: Lempel-Ziv Algorithms (LZ77, LZ78,LZW)
    (used in Unix compress, gzip, pkzip, winzip,...)

Run Length Encoding

- Sequences of identical values are represented by the value and the number of repetitions
- Run length coding is most efficient for highly correlated signals
- Applications are especially in
  - bi-level images like text, graphics, maps,
  - differential coding (DPCM)
  - bit level coding
- Sensitive to transmission errors

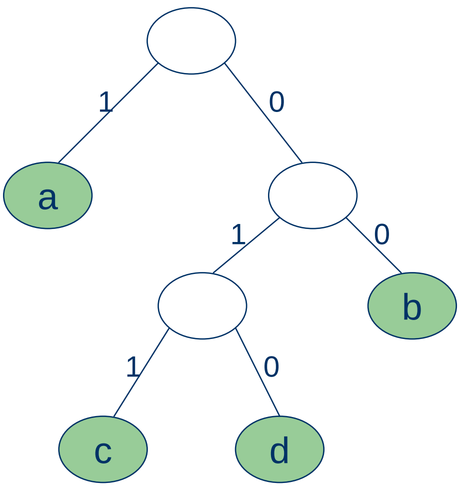




Huffman Coding (1952)

- Huffman algorithm generates Code trees
- For an alphabet, the Code tree represents the code words:

Symbol ≡ Leaf of the tree
Code ≡ Path from root to leaf



Symbol	Code
a	1
b	00
c	011
d	010

Huffman Coding

- Huffman coding generates code tree bottom-up:

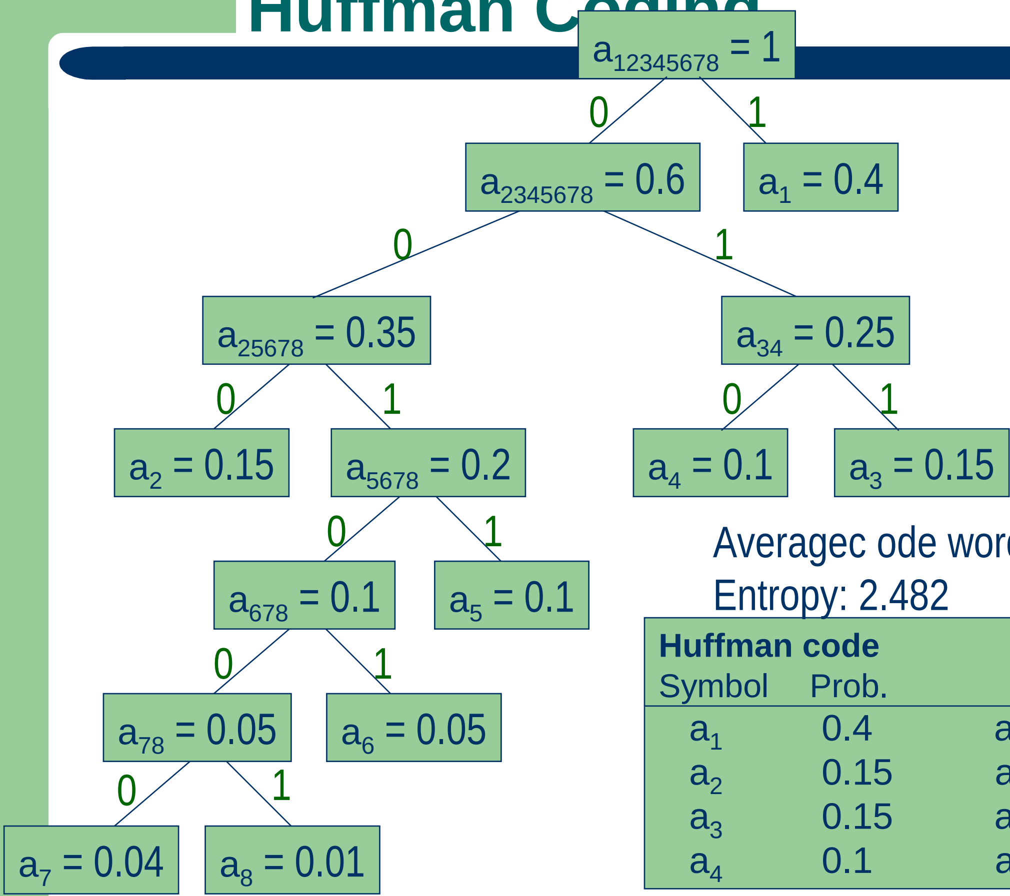
Huffman coding

```
Assemble nodes with every symbol of the input alphabet and
give them weights corresponding to the probabilities

while not_a_tree
      Let L0, L1 be the nodes with smallest weight
      Assemble father L01 with sons L0 and L1
            // assign code symbols (e.g. 0, 1) to the edges
            // weight(L01) = weight(L0) + weight(L1)
end
```

- Huffman codes are prefix codes
- However:
  - Code is not optimal, because the number of bits for each code word is integer (every symbol is mapped to an individual code word)

Huffman Coding



Averagec ode word length: 2.55

Entropy: 2.482

Huffman code

Symbol	Prob.		
a_1	0.4	a_5	0.1
a_2	0.15	a_6	0.05
a_3	0.15	a_7	0.04
a_4	0.1	a_8	0.01

Huffman Coding – Features

- Principles
  - Huffman Coding is the mostly used entropy coding scheme
  - Probability of the input signal is computed and variable length code words are assigned to the signal
    Code table is needed within encoder and decoder
  - Code is nearly optimal, i.e. free from redundancies
  - Code has variable word length and thus a variable bitrate
  - No control symbols (e.g. for separation) are necessary:
    Decoder recognizes begin and end of a code word
  - Mostly used in combination with RLE to treat long code words properly
  - Uses two runs through the input signal
    - Calculation of probabilities
    - Encoding
  - Possible: Dynamic Huffman Coding:
    - Start with estimated probabilities
    - Adopt probabilities during encoding process

- Advantages
  - Huffman Codes are prefix codes, i.e. no code word is a prefix of a longer code word
  - With fixed probabilities and if the code word length are integer values Huffman coding is optimal, i.e. there is no better coding algorithm.
  - Compression ratio for natural images is ~ 1:3

- Disadvantages
  - Huffman codes are very sensible against transmission errors

Compression

- We have seen:

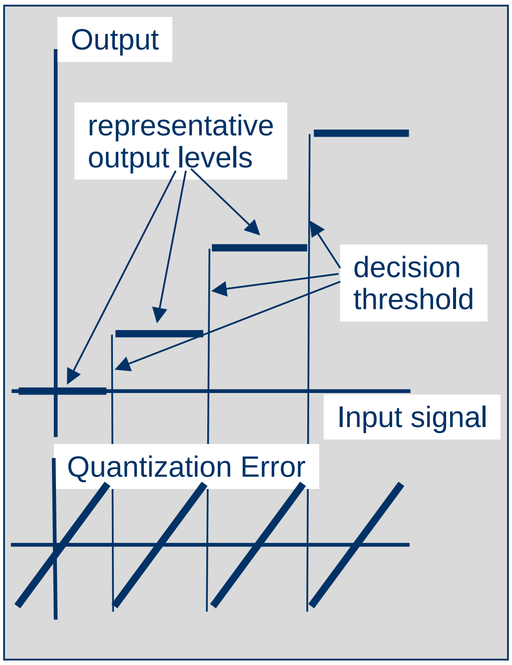
Entropy coding
(lossless compression)

Universal Algorithms	Statistical Algorithms	Table based Algorithms
• Run Length Encoding • Statistical Pattern Substitution	• Morse-Code • Huffman-Code • Arithmetic Coding	• Lempel-Ziv-Algorithm

- What we have not seen
  - Source Coding / Transformation Coding
  - Hybrid Coding

Quantization principle

Quantization is most important means of Irrelevancy Reduction

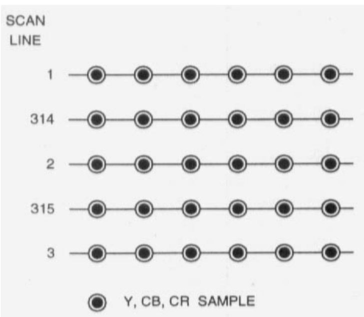


Quantization of a continuous signal gives digital representation within a certain range

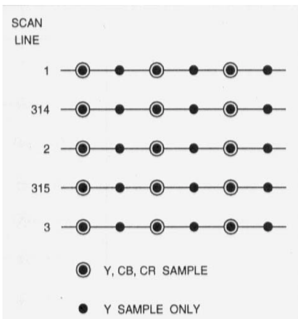
Principle:
- Representative output levels are attributed to continuous range of input values
- Construction of quantizer according to
  - error criteria (maximum error, total error,…) or
  - entropy criteria

For irrelevancy reduction in coding systems quantization is performed on color values, prediction signals, transform coefficients, ...

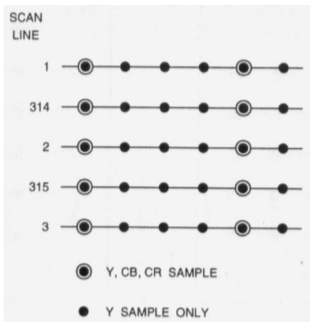
Color Subsampling



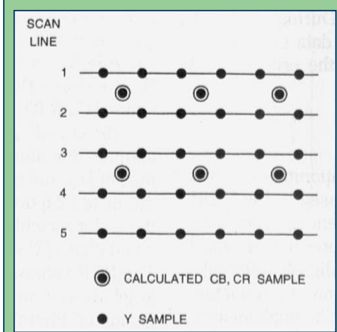
4:4:4 YCbCr Format

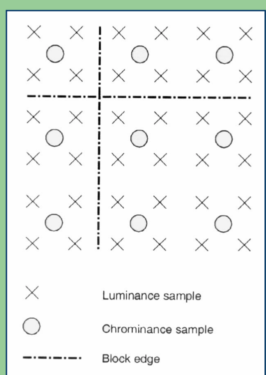


4:2:2 YCbCr Format



4:1:1 YCbCr Format

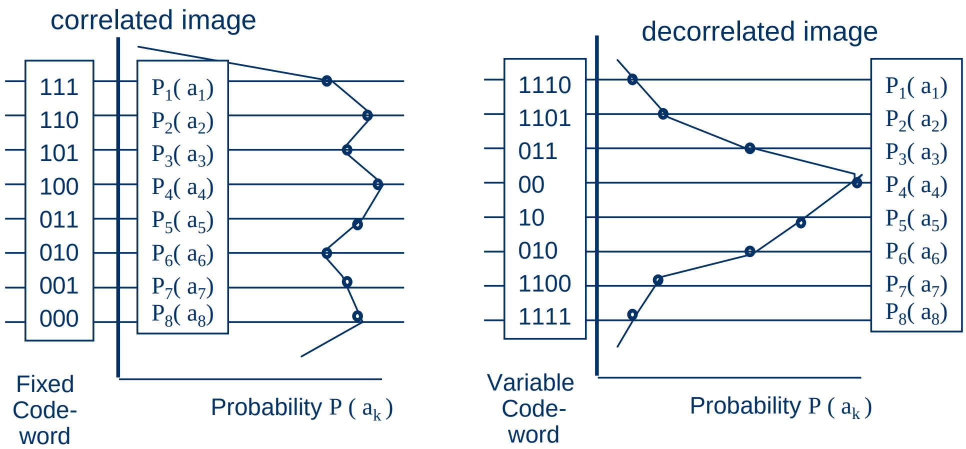




4:2:0 YCbCr Format

Variable Word Length Coding

Statistical or Entropy Encoding means encoding with variable wordlength
$\ell(a_k) \sim P(a_k)$
Frequent values --> short code words
Rare values --> long code words

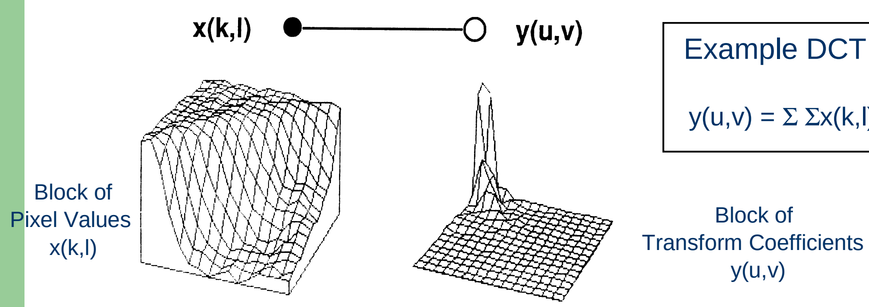


Entropy: $E = -\sum_{i=1}^{N} P(a_k) \log P(a_k)$ $\Rightarrow$ correlated E ~ 3 bit; decorrelated E ~ 2.5 bit

Fixed word length:	$\ell(a_k) = 3$ bit	--> $\ell_{AV} = 3$ bit
Variable word length:	$\ell(a_k) \sim P(a_k)$	--> $\ell_{AV} = 2.5$bit

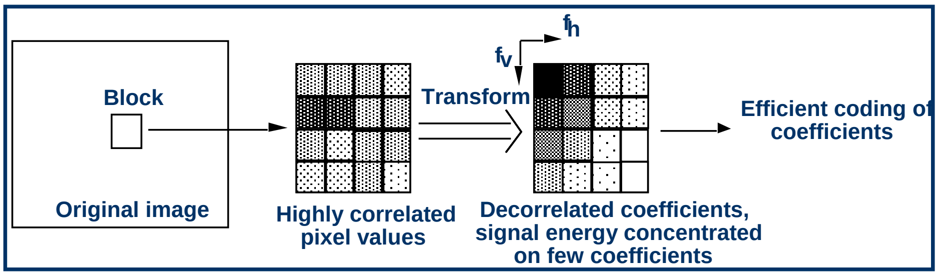
Redundancy Reduction by Transformation

Linear transformation and linear filter decorrelate image signal, concentrate information (energy) on few elements and reduce redundancy



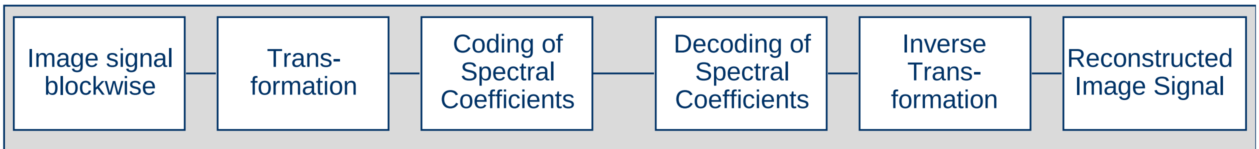
Example DCT:

$$y(u,v) = \Sigma\, \Sigma x(k,l) \cos \frac{š(2l+1)}{2N} u \cos \frac{š(2k+1)}{2N} v$$

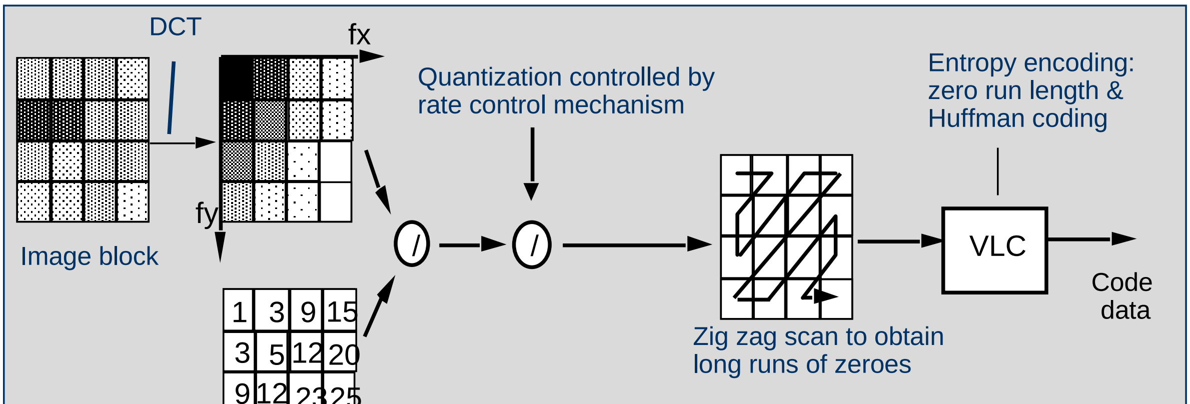


Transform Coding Principle

Coding Problem:
Distribute a limited number of bits among NxN transform coefficients such that the resulting distortion is minimum

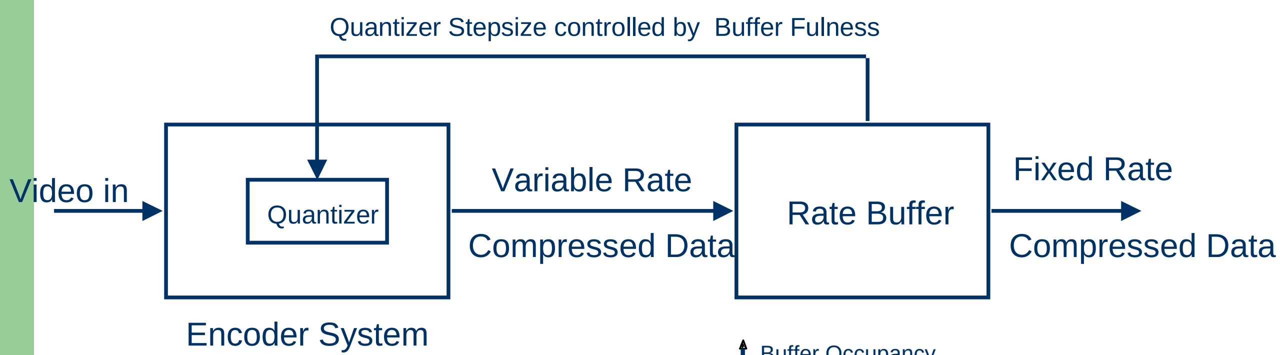


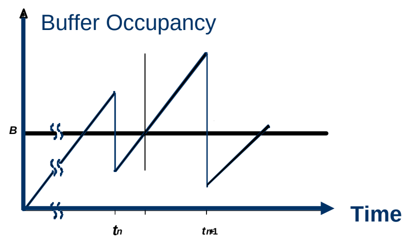
Block diagram of transform coding



Coefficient weighting according to visual sensitivity: high spatial frequencies can be coded coarsely

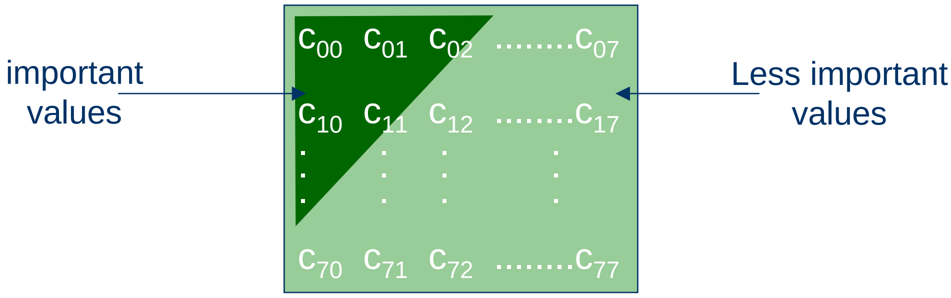
Rate Control of Encoder





Diskrete Cosinus Transformation

- Data compression?
- Important values: top left,
  less important values bottom right



$\Rightarrow$ Elimination of less important values

Discrete Cosinus Transformation

$$\begin{bmatrix} 95, & 88, & 87, & 95, & 88, & 95, & 95, & 95 \\ 143, & 144, & 151, & 151, & 153, & 170, & 183, & 181 \\ 153, & 151, & 162, & 166, & 162, & 151, & 126, & 117 \\ 143, & 144, & 133, & 130, & 143, & 153, & 159, & 175 \\ 123, & 112, & 116, & 130, & 143, & 147, & 162, & 189 \\ 133, & 151, & 162, & 166, & 170, & 188, & 166, & 128 \\ 160, & 168, & 166, & 159, & 135, & 101, & 93, & 98 \\ 154, & 155, & 153, & 144, & 126, & 106, & 118, & 133 \end{bmatrix}$$

Intervall shift to [-128, 127] by subtracting 128

$$\begin{bmatrix} -33, & -40, & -41, & -33, & -40, & -33, & -33, & -33 \\ 15, & 16, & 23, & 23, & 25, & 42, & 55, & 53 \\ 25, & 23, & 34, & 38, & 34, & 23, & -2, & -11 \\ 15, & 16, & 5, & 2, & 15, & 25, & 31, & 47 \\ -5, & -16, & -12, & 2, & 15, & 19, & 34, & 61 \\ 5, & 23, & 34, & 38, & 42, & 60, & 38, & 0 \\ 32, & 40, & 38, & 31, & 7, & -27, & -35, & -30 \\ 26, & 27, & 25, & 16, & -2, & -22, & -10, & 5 \end{bmatrix}$$

Discrete Cosinus Transformation

$$DCT(i,j) = \begin{cases} 1/\mathrm{sqrt}(N), & \text{if } i = 0 \\ \mathrm{sqrt}(2/N) * \cos(((2j+1)\ i\ \pi) / (2N)), & \text{if } i > 0 \end{cases}$$

with N=8, $i,j \in \{0,1,\ldots,7\}$

0.3535	0.3535	0.3535	0.3535	0.3535	0.3535	0.3535	0.3535
0.4904	0.4157	0.2777	0.0975	-0.0975	-0.2777	-0.4157	-0.4904
0.4619	0.1913	-0.1913	-0.4619	-0.4619	-0.1913	0.1913	0.4619
0.4157	-0.0975	-0.4904	-0.2777	0.2777	0.4904	0.0975	-0.4157
0.3535	-0.3535	-0.3535	0.3535	0.3535	-0.3535	-0.3535	0.3535
0.2777	-0.4904	0.0975	0.4157	-0.4157	-0.0975	0.4904	-0.2777
0.1913	-0.4619	0.4619	-0.1913	-0.1913	0.4619	-0.4619	0.1913
0.0975	-0.2777	0.4157	-0.4904	0.4904	-0.4157	0.2777	-0.0975

Discrete Cosinus Transformation

93.12	2.12	-8.45	-7.49	3.125	0.65	1.47	-2.34
-37.51	-57.62	10.62	17.37	-3.34	4.86	5.28	-2.92
-84.07	63.17	-1.08	-17.09	1.68	7.14	-4.34	-0.008
-50.87	-37.30	-10.36	12.98	-9.98	5.36	-0.62	-4.15
-84.87	-41.54	49.55	-8.36	17.62	-4.76	-0.90	0.72
-62.55	65.92	-13.44	-0.59	1.88	-6.02	-1.62	-2.33
-16.45	14.49	-36.84	17.61	-11.92	4.24	3.33	-3.08
-52.58	30.61	-7.41	-9.83	23.17	-0.40	1.68	1.66

Rounding

93	2	-8	-7	3	1	1	-2
-38	-58	11	17	-3	5	5	-3
-84	63	-1	-17	2	7	-4	0
-51	-37	-10	13	-10	5	-1	-4
-85	-42	50	-8	18	-5	-1	1
-63	66	-13	-1	2	-6	-2	-2
-16	14	-37	18	-12	4	3	-3
-53	31	-7	-10	23	-1	2	2

Transformation Coding

- Quantisation matrix

```
for (i=0; i < 8; i++) {
  for (j=0; j < 8; j++) {
     qij := 1 + ((1+i+j)*quality);
  }}
```

quality = 2

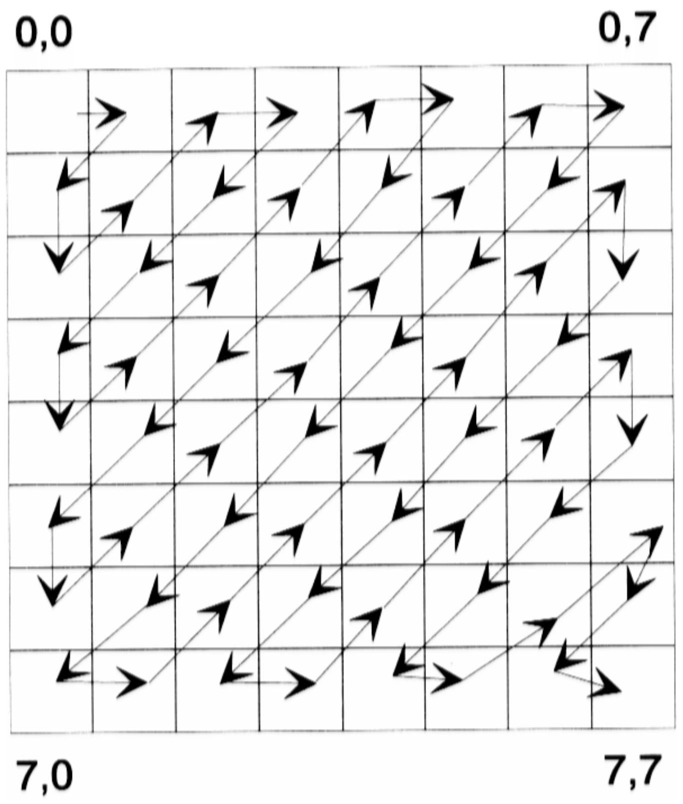
$$\begin{pmatrix} 3, & 5, & 7, & 9, & 11, & 13, & 15, & 17 \\ 5, & 7, & 9, & 11, & 13, & 15, & 17, & 19 \\ 7, & 9, & 11, & 13, & 15, & 17, & 19, & 21 \\ 9, & 11, & 13, & 15, & 17, & 19, & 21, & 23 \\ 11, & 13, & 15, & 17, & 19, & 21, & 23, & 25 \\ 13, & 15, & 17, & 19, & 21, & 23, & 25, & 27 \\ 15, & 17, & 19, & 21, & 23, & 25, & 27, & 29 \\ 17, & 19, & 21, & 23, & 25, & 27, & 29, & 31 \end{pmatrix}$$

Transformation Coding

- Quantized Coefficients

$$\begin{pmatrix} 31, & 0, & -1, & 0, & 0, & 0, & 0, & 0 \\ -7, & -8, & 1, & 1, & 0, & 0, & 0, & 0 \\ -12, & 7, & 0, & -1, & 0, & 0, & 0, & 0 \\ -5, & -3, & 0, & 0, & 0, & 0, & 0, & 0 \\ -7, & -3, & 3, & 0, & 0, & 0, & 0, & 0 \\ -4, & 4, & 0, & 0, & 0, & 0, & 0, & 0 \\ -1, & 0, & -1, & 0, & 0, & 0, & 0, & 0 \\ -3, & 1, & 0, & 0, & 0, & 0, & 0, & 0 \end{pmatrix}$$

Transformation Coding



Transformation Coding

Source image

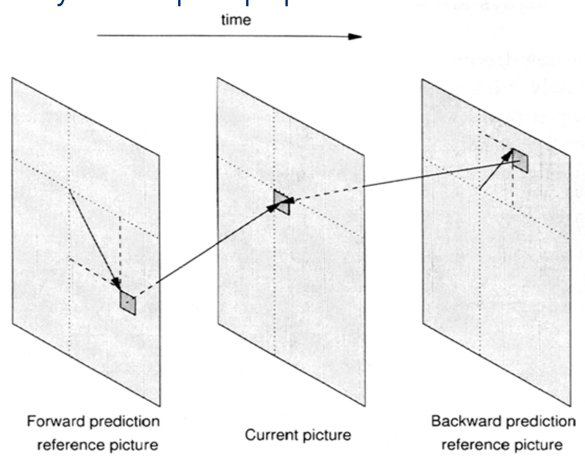
95,	88,	87,	95,	88,	95,	95,	95
143,	144,	151,	151,	153,	170,	183,	181
153,	151,	162,	166,	162,	151,	126,	117
143,	144,	133,	130,	143,	153,	159,	175
123,	112,	116,	130,	143,	147,	162,	189
133,	151,	162,	166,	170,	188,	166,	128
160,	168,	166,	159,	135,	101,	93,	98
154,	155,	153,	144,	126,	106,	118,	133

Reconstructed Picture

98,	95,	91,	89,	90,	95,	101,	106
140,	143,	149,	156,	163,	167,	168,	167
146,	149,	154,	159,	159,	151,	137,	126
149,	142,	136,	137,	145,	156,	163,	166
119,	117,	118,	125,	140,	157,	170,	176
137,	147,	160,	170,	172,	166,	157,	150
166,	167,	164,	152,	132,	112,	99,	93
151,	153,	150,	139,	125,	118,	119,	123

Video Coding

- Intraframe coding
  - Similarities within one single image frame
  - Coding uses data from the actual image
  - Data reduction by using spatial redundancy and spatial properties of human vision
- Interframe coding
  - Similarities in temporally successive frames
  - Coding uses data from actual and other images
  - Data reduction by using temporal redundancy and temporal properties of human vision
  - Important technique:
    - Motion compensated prediction
    - Motion compensated interpolation

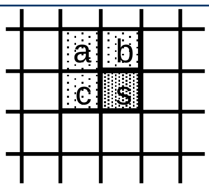


Source: Mitchell et al., MPEG Video Compression Standard

Predictive Coding

Prediction causes effective decorrelation of image signal by concentrating energy into few prediction values

Prediction:



Pel *s* is predicted from pels *a*, *b* and *c*

$p = (a + b + c) / 3$

(various more complex formulas are used)

The prediction error e is:

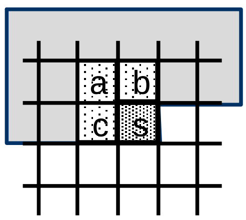
$e = s - p$

Transmit only prediction error *e*

Decorrelation:
- In homogeneous areas e equals 0; at edges e takes high values
- High energy concentration around e = 0 & high entropy reduction

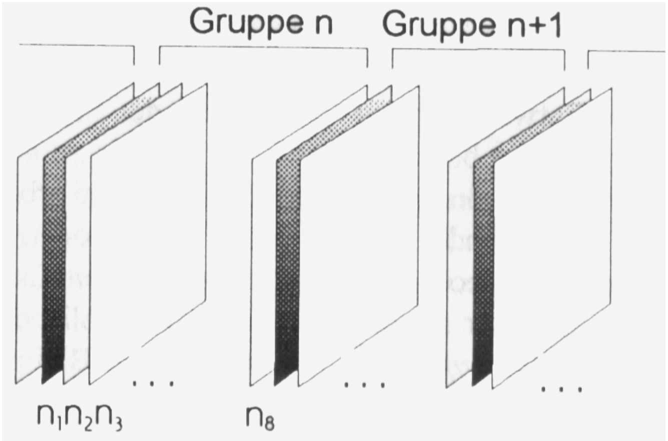
Compression:
- Entropy encoding of prediction error → losless
- Quantization of prediction error → lossy

Coding System:
- Transmit the prediction error e
- Reconstruction at receiver with identical predictor scheme
- Only previously coded information is used for prediction
- In case of transmission error the receiver will reconstruct false values and this failure will propagate depending on used predictor

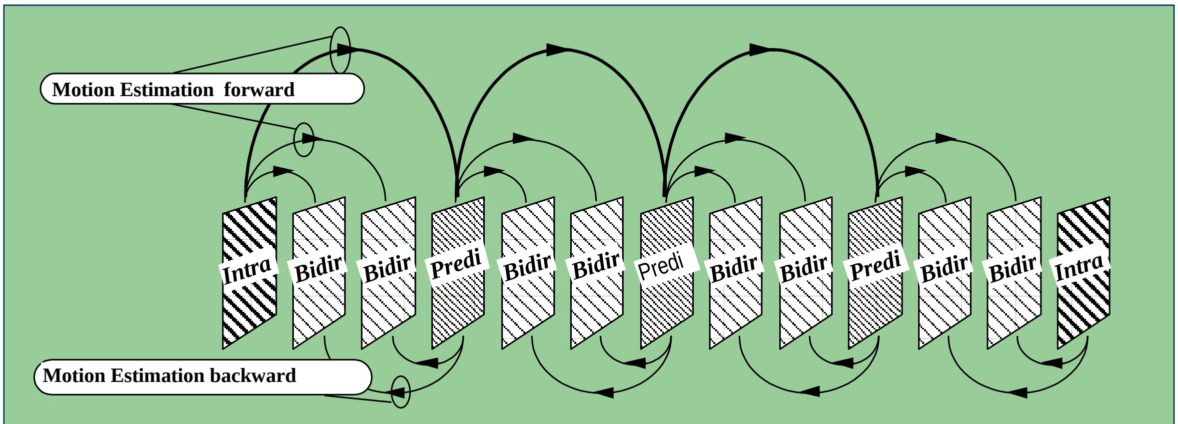


Coding of Image Sequences

- **Idea**:
  Minor changes in image content in successive pictures
- Partition of image sequences in groups
- All groups have the same structure (same number of pictures)
- Coding is done ‚groupwise': Groups are coded together
- Access to individual images not possible



Prediction & Interpolation

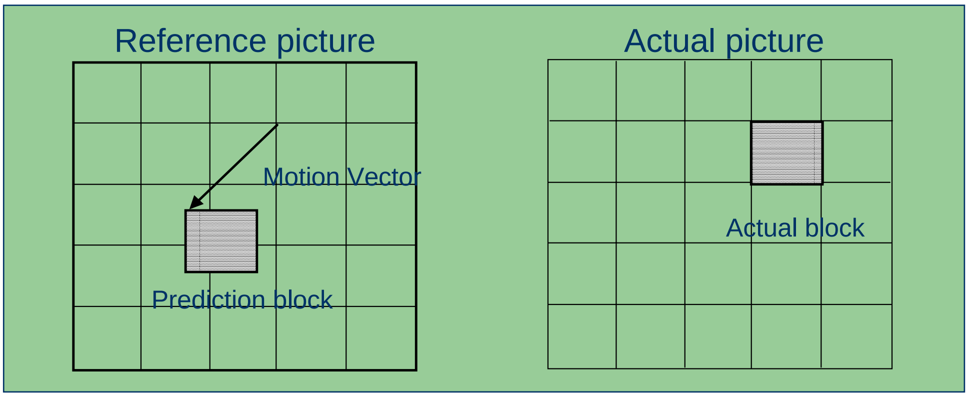


Picture order	1	2	3	4	5	6	7	8	9	10	11	12	13
Coding order	**1**	3	4	**2**	6	7	**5**	9	10	**8**	12	13	**11**

- Intra
  - direct DCT-based transformation coding of the picture (no motion compensation)
  - Support of random access, Fast Forward, Fast Rewind
  - Reduction of error propagation
- Predi
  - Hybrid coding (forward prediction)
- Bidir
  - Hybrid Coding
  - forward-, backward or interpolated prediction
- Bitrate: Original / Intra / Predi / Bidir: 1000 / 150 / 50 – 100 / 7 – 15 kbit/frame

Motion Compensation

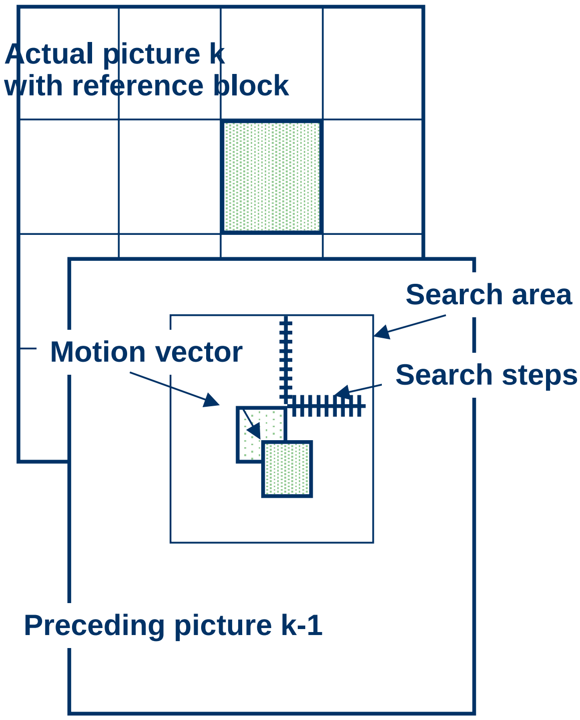
- Goal:
  identify movement of picture blocks
- Prediction block is an identical (or similar) block within the reference picture, but shifted spatially
- Motion vector defines the relative position of the actual block to the corresponding block in the reference picture
- Using prediction
  ⟹ elimination of temporal redundancy



Motion Estimation (ME)

- ME tries to identify the movement in a picture sequence
- ME is needed for coding P- and B-Frames
- A macroblock which is similar to the actual macroblock is searched in one or two reference pictures
  - P-Frame: Reference picture is the preceding I- or P-frame
  - B-Frame: Reference picture
    - preceding I- or P-frame
    - following P-frame
    - Interpolation between preceding I- or P-frame and following P-frame
- For every macroblock of a P- or B-Frame one or two motion vectors are computed
  - P-Frame: 1 motion vector
  - B-Frame: 1 or 2 motion vectors
- Mostly used: Block Matching technique
- ME is the most expensive part within the coding process

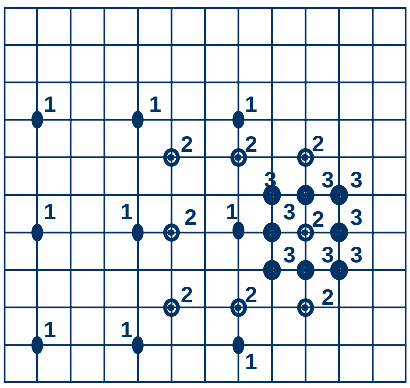
Motion Estimation by Block Matching



- Block Matching:
  - Block of size MxN pels of picture k is compared with all blocks of picture k-1 within a search area
- Matching criteria:
  - Usage of different criteria:
    - Absolute difference, quadratic difference, Cross Correlation function,..
  - Examples
    - Mean Squared Error:
      MSE (dx,dy) = Σ [Σ k(x,y) - Σ k-1 (x+dx,y+dy)]2
    - Sum of Absolute Distortion
      SAD (dx,dy) = Σ | Σ k(x,y) - Σ k-1 (x+dx,y+dy)|
    - Mean Absolute Distortion
      MAD = 1/(M*N) * SAD
- Newer coding schemes use half or quarter pel-resolution

Block Matching

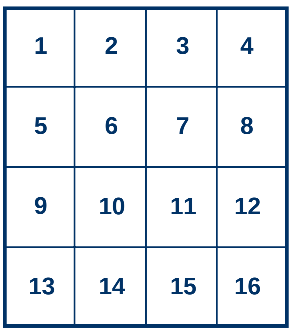
Fast search algorithms by
- Full Search,
- Three Step Search,
- Logarithmic or Spiral search



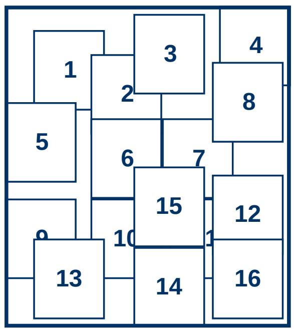
Three Step Search:
Hierarchical Search in 3 steps,
each 9 search positions
with step sizes 3, 2 resp. 1 pels.
Search operations are reduced
from 169 to 27

Block matching does not guarantee that matched blocks are identical, but that the signal difference according to the optimization criterium is minimum

Also sub optimal correlation reduces the information to be coded!

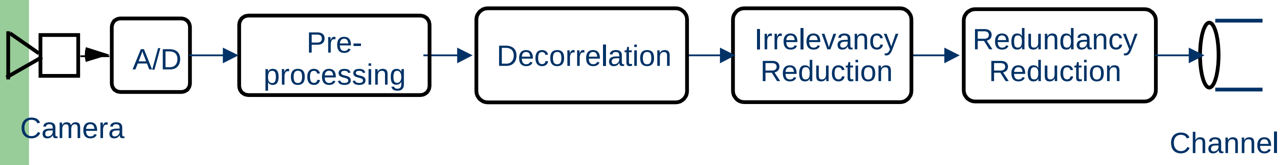


blocks in original image



optimal matching & possible overlap in reference frame

Structure of a codec



Spatial/Temporal Filtering

Subsampling

Analysis for adaptive coding

Transformation DCT

Spatial prediction DPCM

Temporal prediction DPCM

Motion compensation

Subband splitting

Wavelets

Quantization

Vector Quantization

Subsampling

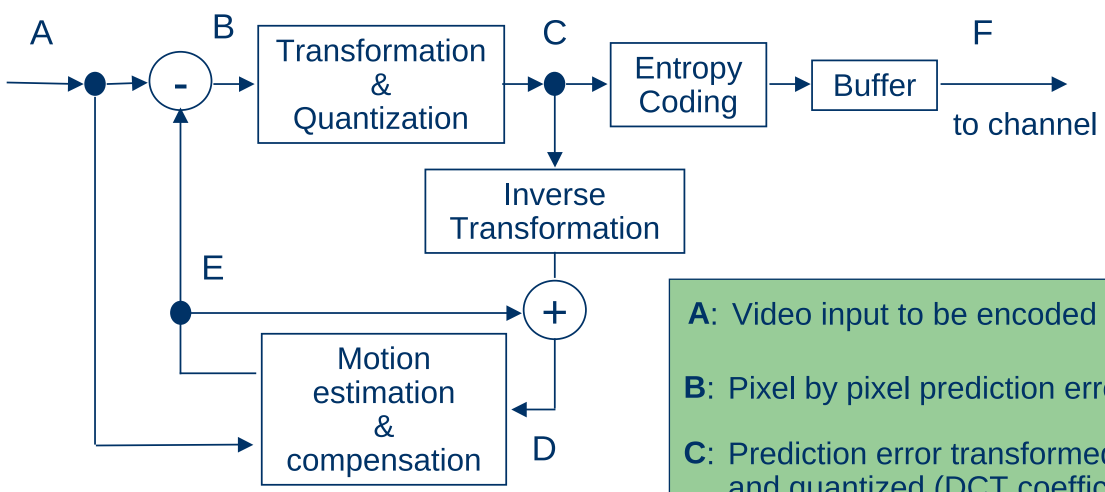
Huffman Coding

Runlength Coding

Arithmetic Coding

Lempel Ziv Coding

Motion Compensated Hybrid Coder



A: Video input to be encoded

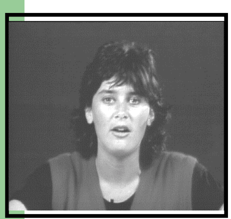
B: Pixel by pixel prediction error

C: Prediction error transformed and quantized (DCT coefficients)

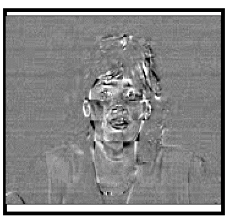
D: Reconstructed pixel values degraded by quantization

E: Motion compensated predicted pixel values

F: Coded video bitstream



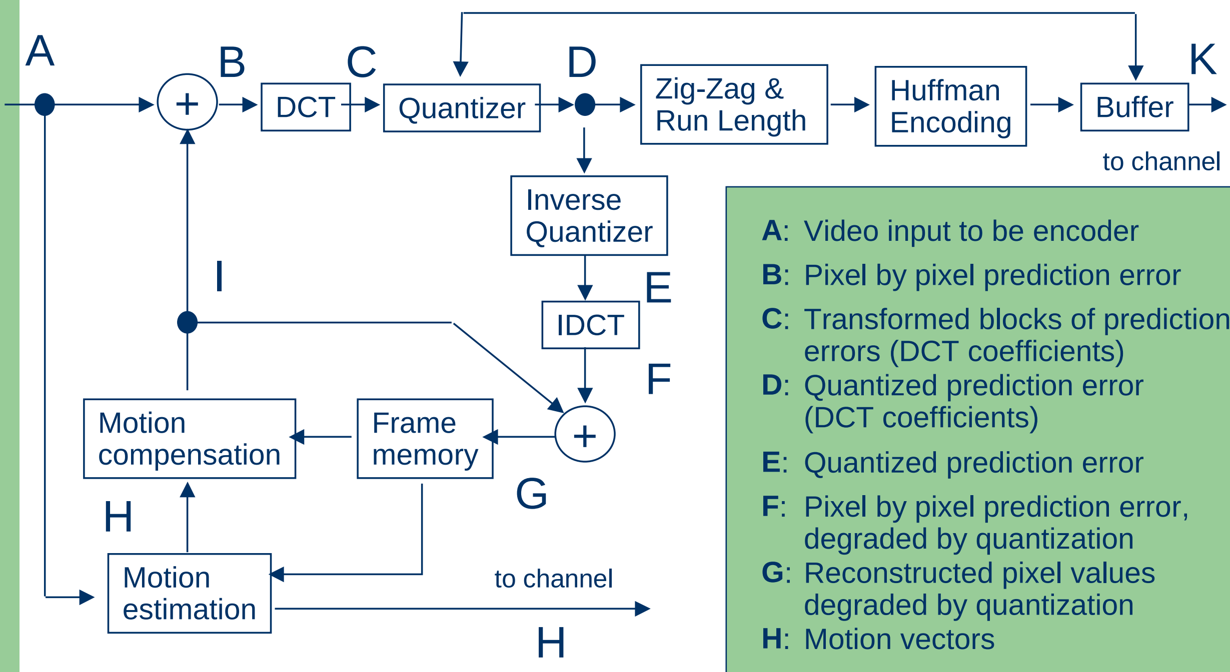
Original Image



Motion compens prediction image

Residual Error

Motion Compensated Hybrid Coder



A: Video input to be encoder

B: Pixel by pixel prediction error

C: Transformed blocks of prediction errors (DCT coefficients)

D: Quantized prediction error (DCT coefficients)

E: Quantized prediction error

F: Pixel by pixel prediction error, degraded by quantization

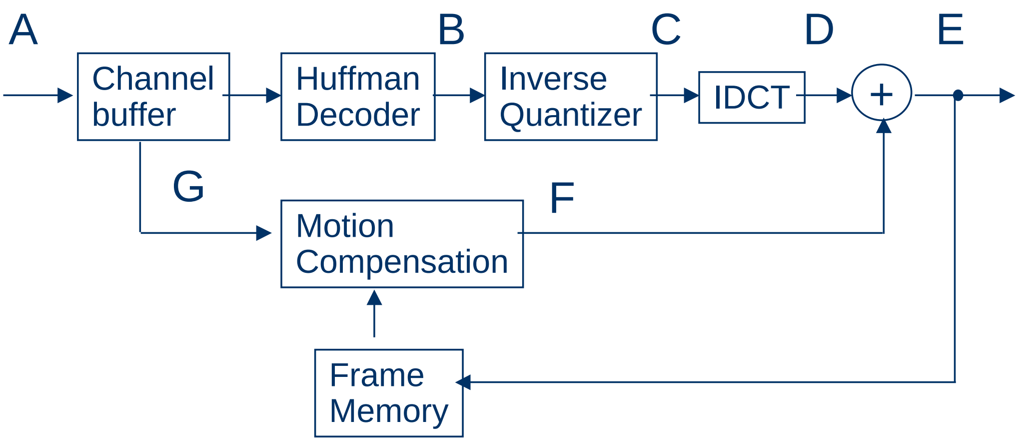
G: Reconstructed pixel values degraded by quantization

H: Motion vectors

I: Motion compensated predicted pixel values

K: Coded video bitstream

Motion Compensated Hybrid Decoder



A: Coded video bit stream

B: Prediction error DCT coefficients

C: Quantized prediction error

D: Pixel by pixel prediction error

E: Reconstructed pixel values

F: Motion compensated predicted pixel values

G: Motion vectors

Compression Artefacts

- Blockiness
  - Visibility of the borders of coding blocks caused by coarse quantization of low detail regions
- Blurriness
  - Loss of spatial detail (sharpness) of edges
- Ringing
  - Periodic line structure with decreasing amplitude.
  - Occurs at edges on flat background if its high frequency components are coarsely quantized
- Color Bleeding
  - Similar to ringing, but at strong chrominance edges
- Moscito Effect
  - Clouds of small spots around moving objects.
    Caused by inproper motion compensation