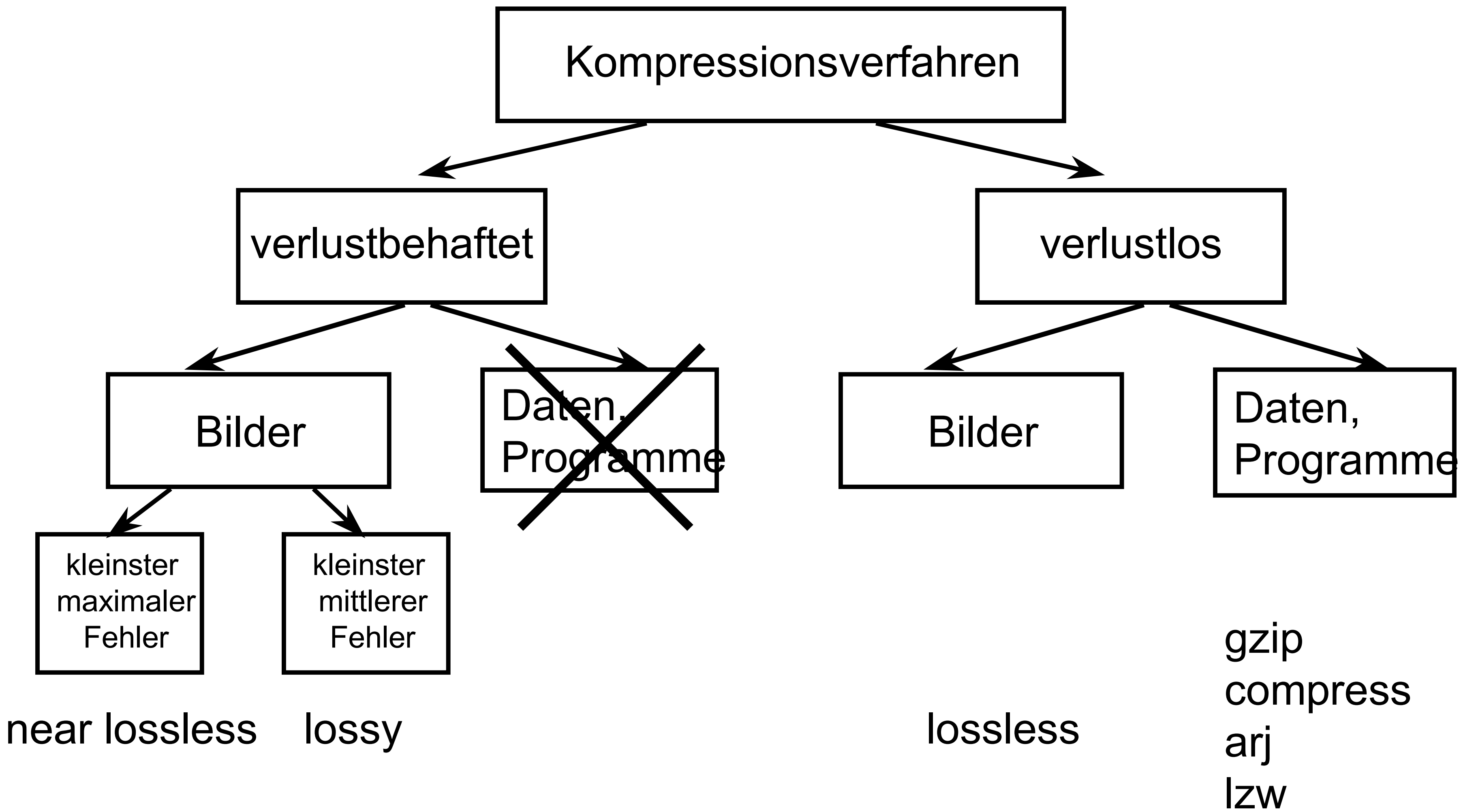
Standbildcodierung

Dipl.-Ing. Guido Heising

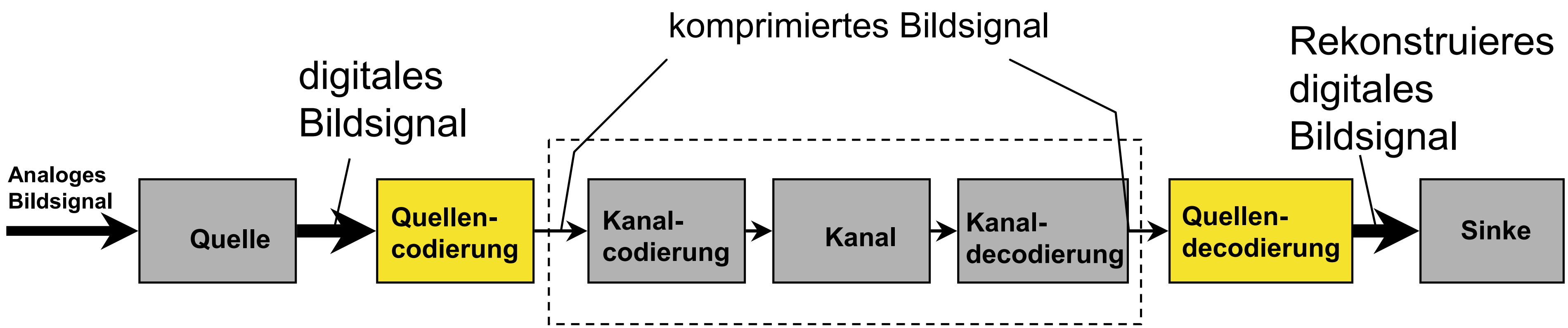
Gliederung der Vorlesung

- Einführung in die Bildcodierung
  - verlustlose/verlustbehaftete Codierung
  - Bestimmung der Codiereffizienz
  - Redundanz/Irrelevanz
- Prädiktive Codierung
  - örtliche Prädiktion
- Entropiecodierung
  - Informationsgehalt/Entropie
  - Huffman-Codierung
- 1D- und 2D-Transformationscodierung
  - 1D-Diskrete-Cosinustransformation (DCT)
  - 2D-DCT
- Lauflängencodierung

Übersicht Kompressionsverfahren



Digitale Bildübertragungsstrecke



Bilderfassung	Coder	Speichermedium	Decoder	Bildausgabe
Scanner	-	Harddisk	-	Filmprinter
Digitale Kamera	JPEG	Floppy	JPEG	Monitor
Filmscanner	MPEG	Internet	MPEG	Drucker
	...	...	...	...

Speichermedien

Medium	Größe
Floppy	1.4 MB
CD	650 MB
RAM	16 - 256 MB
Harddisk	1 - 100 GB (Giga Bytes)
Storage Systems	einige TB (Terra Bytes)

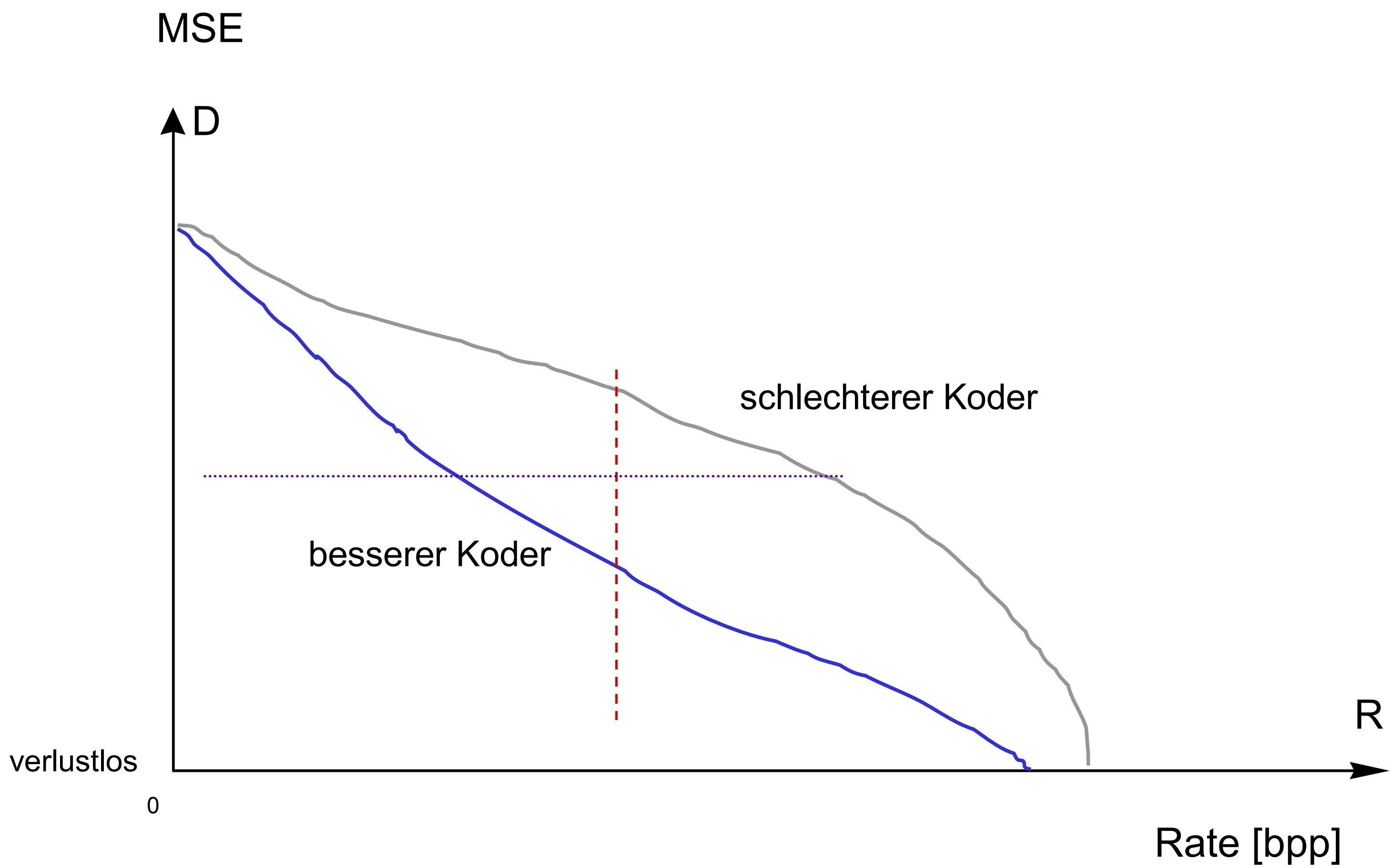
Kanäle (Übertragungsmedien)

Kanal	Datenraten
• GSM	4 - 9.6 kb/s
• Telefon	14.4 - 56 kb/s
• ISDN	64 - 128 kb/s
• zukünftige Mobilnetze	bis 2 Mb/s (variabel)
• Ethernet	bis 4 Mb/s (variabel)
• Fast Ethernet	bis 40 Mb/s (variabel)

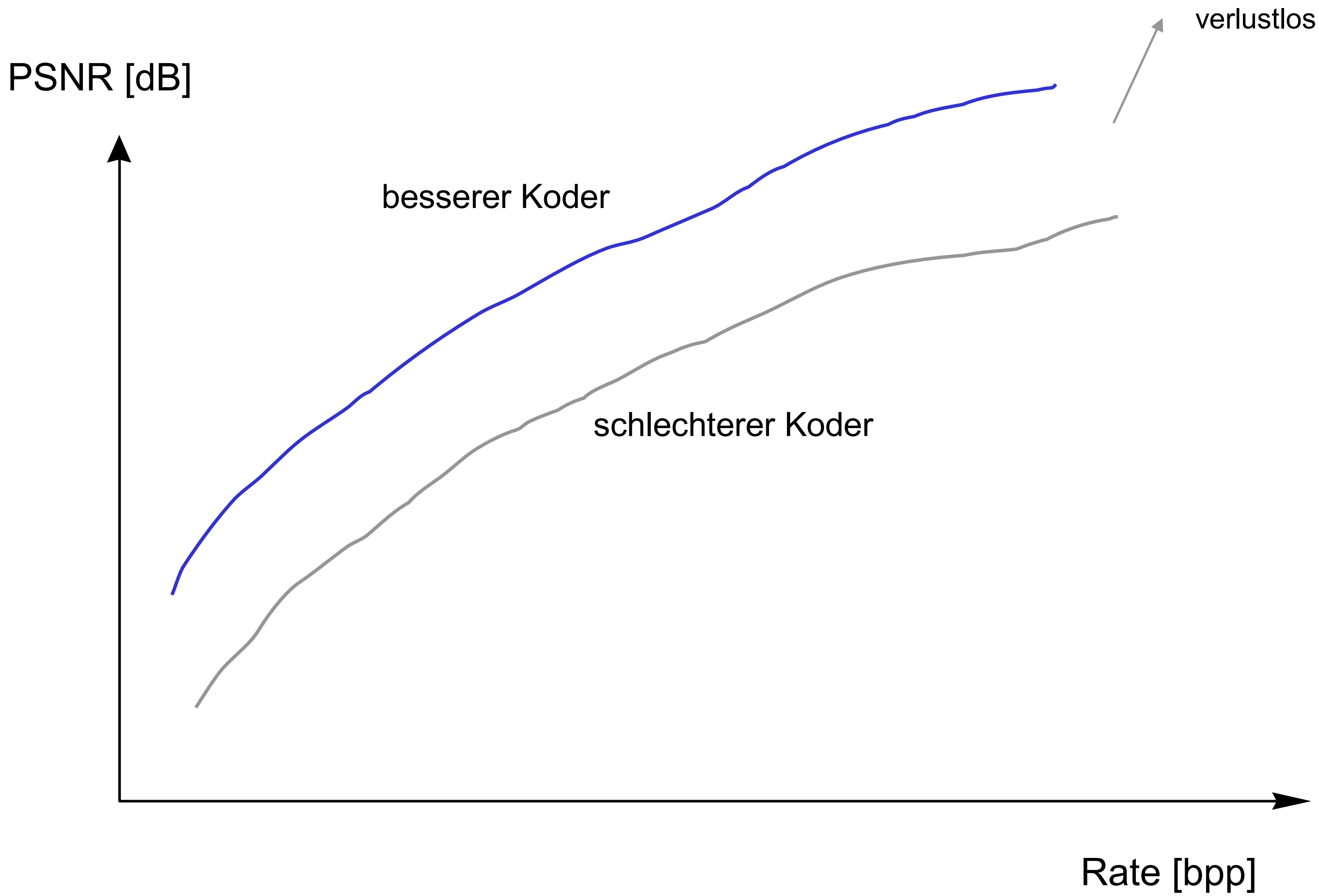
Parameter eines Kompressionsverfahrens

- Datenrate / Kompressionsfaktor
- Verzerrung, Bildqualität
- Komplexität
- Verzögerungszeit

Codierungsvergleich (MSE/Rate)



Codierungsvergleich (PSNR/Rate)



Beispiel: Video über ISDN (64 kbit/s)

Format	Spalten	Zeilen	Bildfrequenz [Hz]	Bilddateigröße [kB]
CIF	352	288	30	152

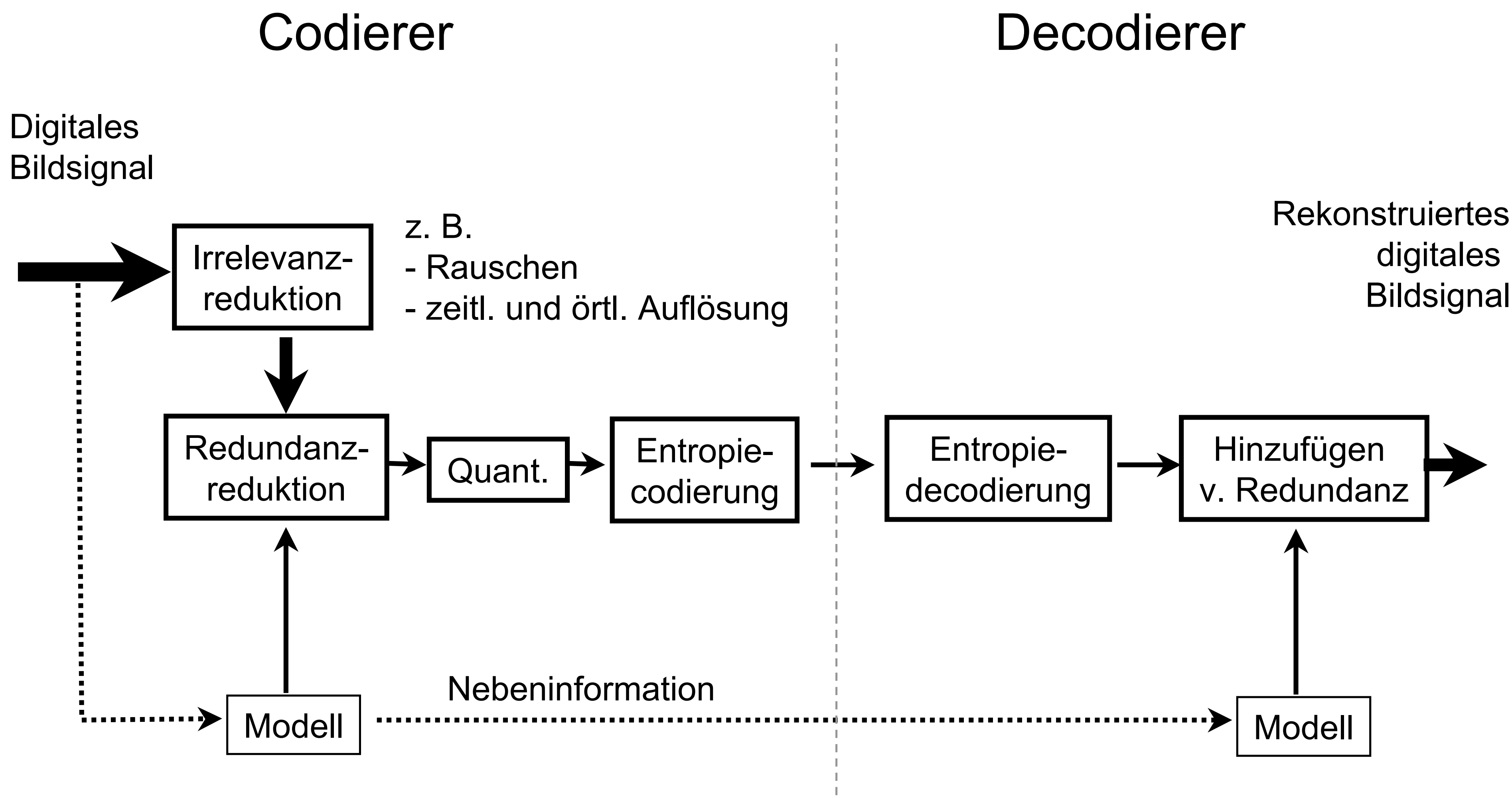
(4:2:0 Format: 8 bit Luminanz + 4 bit Chrominanz pro Pixel)

Datenrate: $352 \cdot 288 \cdot 12 \cdot 30 =$ **36500 kbit/s**

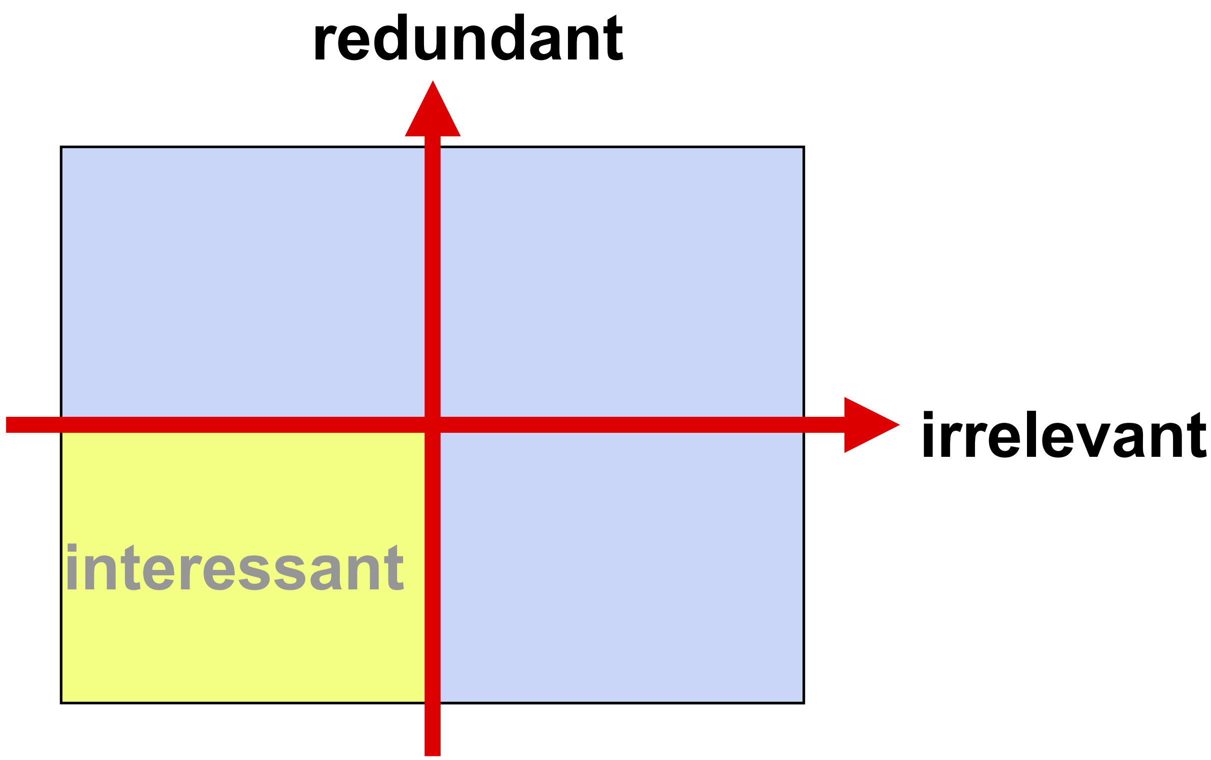
Erforderlicher Kompressionsfaktor für ISDN-Übertragung: 36500 / 64 = **570**

-> **0,021 bit/Pixel**

Prinzip der Quellencodierung



Nachrichtenebene (nach Schouten)



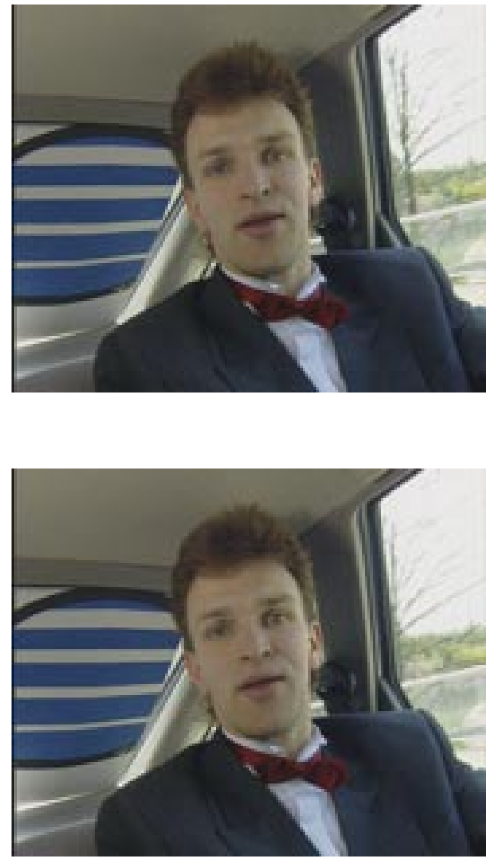
Quellencodierung:

1. Unterdrückung von irrelevanter und redundanter Information
2. Extraktion und Codierung der nicht redundanten Signalanteile

Bild- und Videodaten sind redundant!

Örtlich und zeitlich benachbarte Bildpunkte haben ähnliche Farbwerte

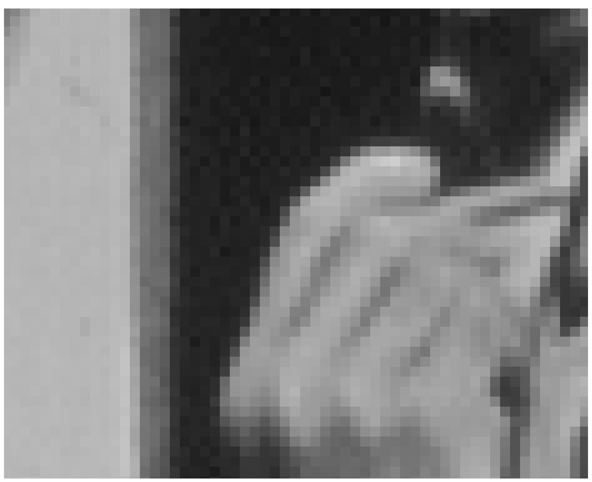
örtlich



zeitlich

Prädiktionscodierung

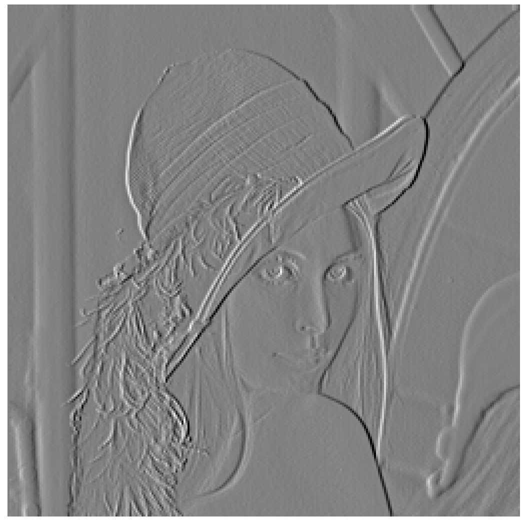
- Redundanzreduktion durch Vorhersage (Prädiktion) des aktuellen Pixels aus bekannten Pixeln aus der Vergangenheit
- Bei der örtlichen Prädiktion im Beispiel unten bedeutet Vergangenheit:
  Prädiktion aus bereits decodierten Pixeln des gleichen Bildes



Beispiel: Örtliche Prädiktion

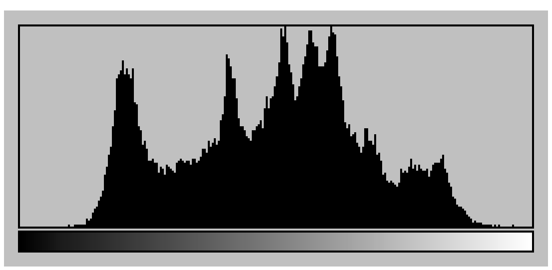
Original

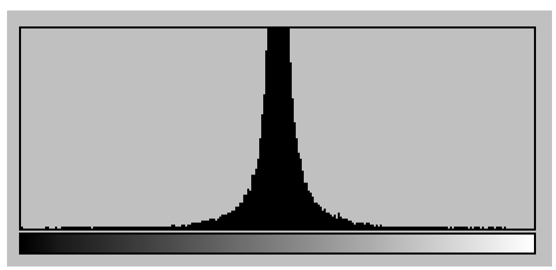
Prädiktionsfehler = Original – Prädiktion

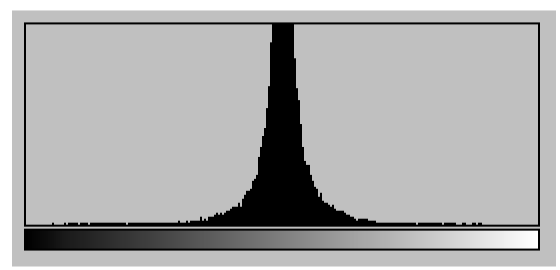


aus horizontalen Nachbarn

aus horizontalen + vertikalen Nachbarn



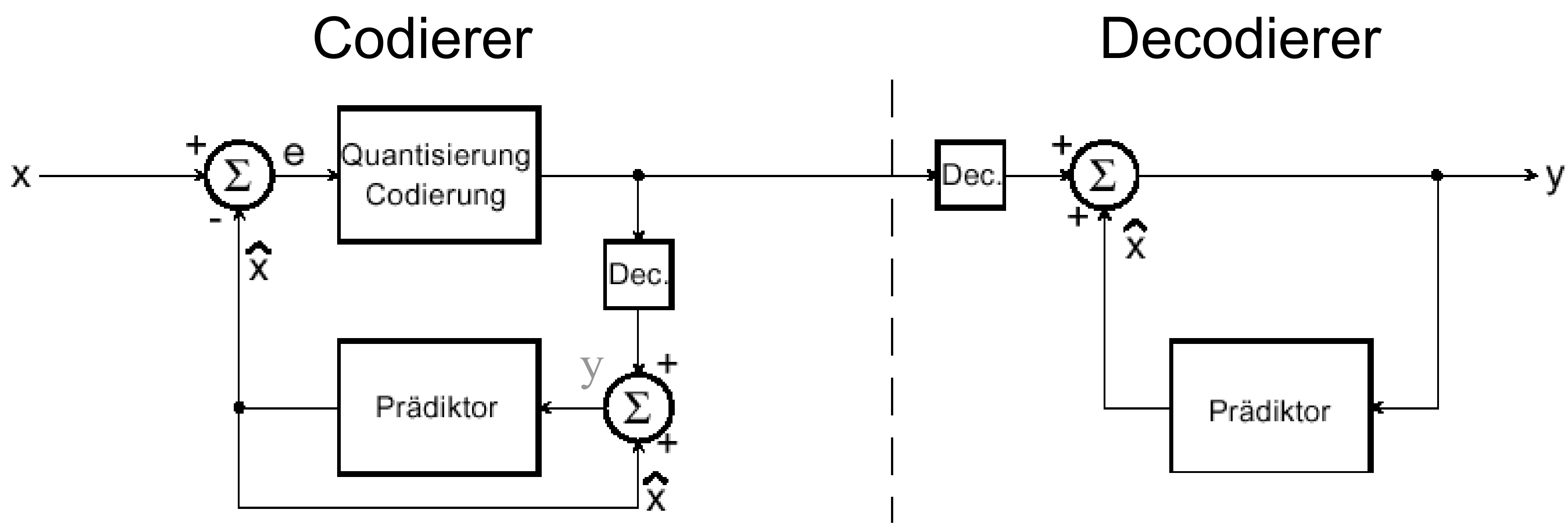




Prädiktionsfehlercodierung

Prädiktionsfehler = Original – Prädiktion

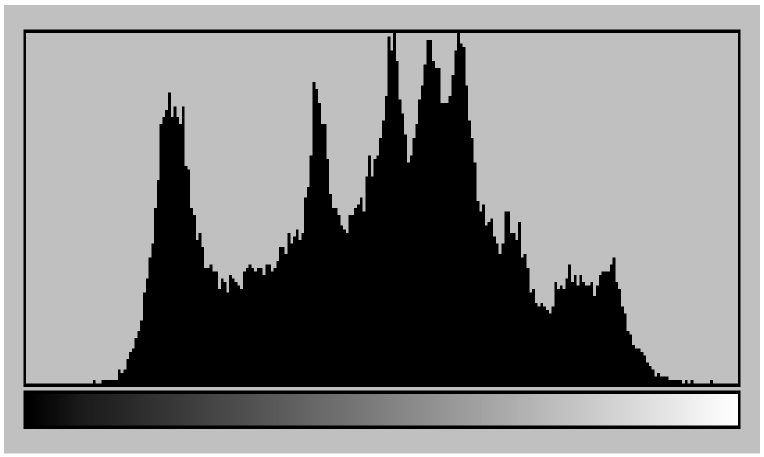
$$e(x) = x - \hat{x}$$



Vorteile der Prädiktion für die Codierung

1. Prädiktionsgewinn: $G_{präd} = \sigma_x^2 / \sigma_e^2$ (Varianz$_{Original}$/Varianz$_{Präd.\text{-}Fehler}$)

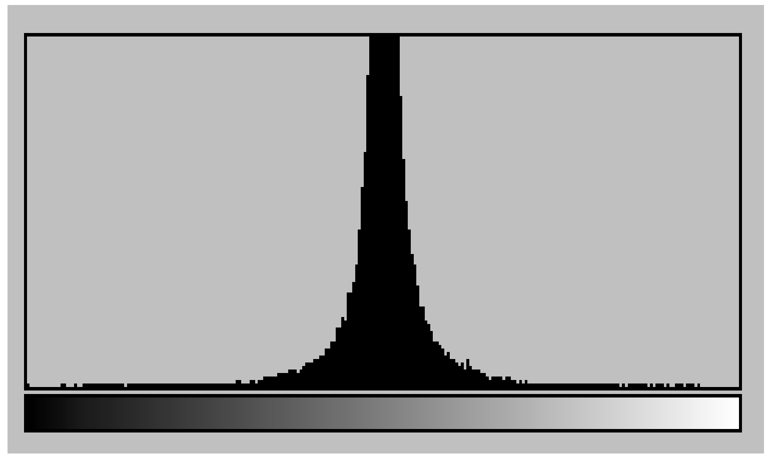
2. Günstigere Amplitudendichteverteilung (ADV):



ADV des Originalbildes *x*

Wertebereich: 0 – 255

-> 8 Bit / Pixel



ADV des Prädiktionsfehlerbildes *e*

Wertebereich: -255 bis 255

-> 9 bit/Pixel, Praxis: < 8 bit/Pixel,

da ungleiche Verteilung der Werte für die Codierung vorteilhaft ist

Entropiecodierung

Codierung mit einem Code variabeler Länge (*VLC*)

-> Häufig auftretende Werte werden mit einem kurzen Codewort repräsentiert, seltene mit langen Codeworten

-> Die mittlere Codewortlänge kann dadurch um so stärker reduziert werden, je ungleichmäßiger die Häufigkeiten verteilt sind

Informationsgehalt *I*, Entropie *H*

- Der ***Informationsgehalt I*** gibt die Infomationsmenge des Symbols *x* in *bit* an

$$I(x) = -\log_2 p(x) \ [\text{bit}]$$

- Die ***Entropie H*** gibt den *mittleren Informationsgehalt I* einer Signalquelle an sowie die *minimale Bitzahl*, die im Durchschnitt zur Übertragung eines Signals mit J unterschiedl. Symbolen mit den Wahrscheinlichkeiten *p*(*x*) nötig ist

$$H = -\sum_{x=1}^{J} p(x) \cdot \log_2 p(x) \ \left[\frac{\text{bit}}{\text{symbol}}\right]$$

- p(x) können z.B die Wahrscheinlichkeiten sein, mit der die einzelnen Farbwerte x im Bild vorkommen

Entropie (Beispiel)

3 Symbole A, B, C mit Auftretenswahrscheinlichkeiten:

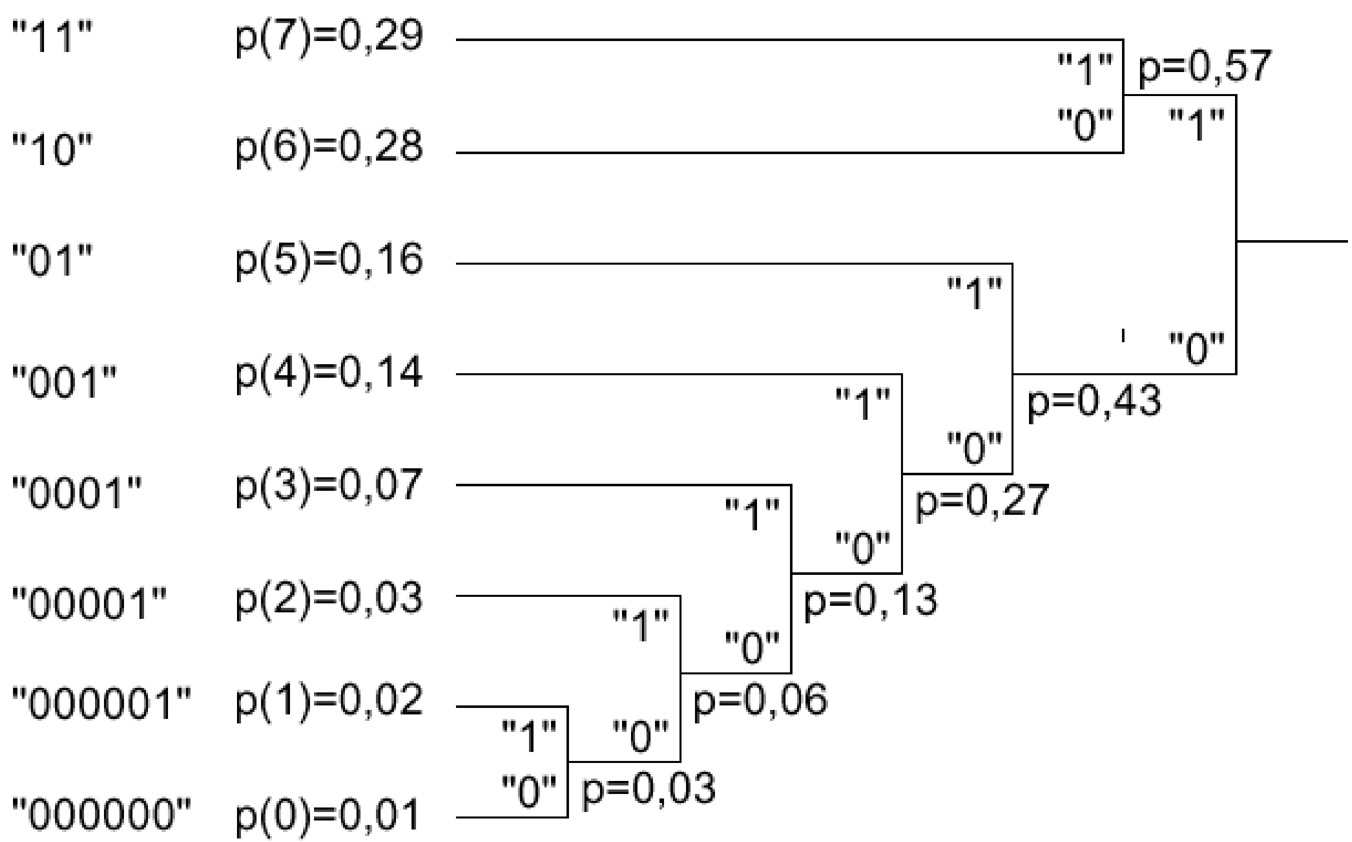
p(A) = 0,5   p(B) = 0,25   p(C) = 0,25

Entropie:	H = -[0,5·(-1) + 0,25 ·(-2) + 0,25 ·(-2)]
$$H = -\sum_{x=1}^{J} p(x) \cdot \log_2 p(x)$$	= 0,5 + 0,5 + 0,5 = 1,5 [bit]
Binärcode 1 :	A: 0 B: 10 C: 11
Codelänge [bit]:	A: 1 B: 2 C:2
mittlere Codelänge:	½·1 + ¼·2 + ¼·2 = 1,5 [bit]
Binärcode 2 :	A: 10 B: 0 C: 11
Codelänge [bit]:	A: 2 B: 1 C:2
mittlere Codelänge:	½·2 + ¼·1 + ¼·2 = 1,75 [bit]

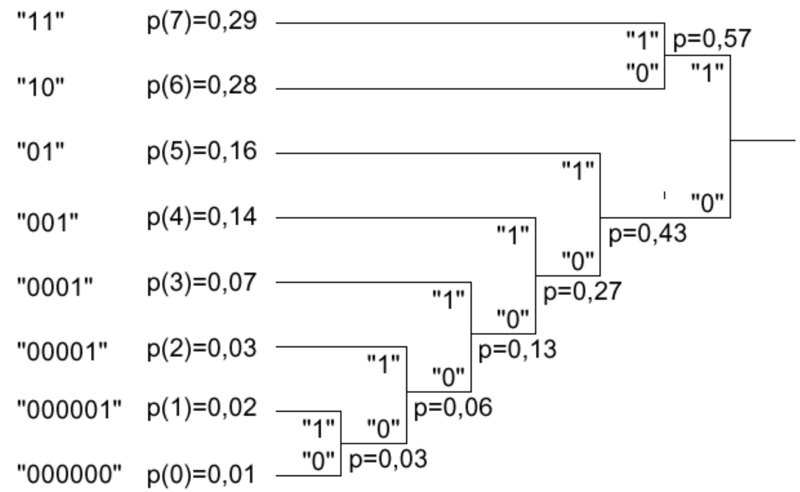
Wie lässt sich ein Code bestimmen?

Ziel: Je wahrscheinlicher ein Symbol, desto kürzer das soll das zugehörige Codewort sein

Verfahren: z. B. mit Huffman-Codierung



Wie lässt sich ein Code bestimmen?



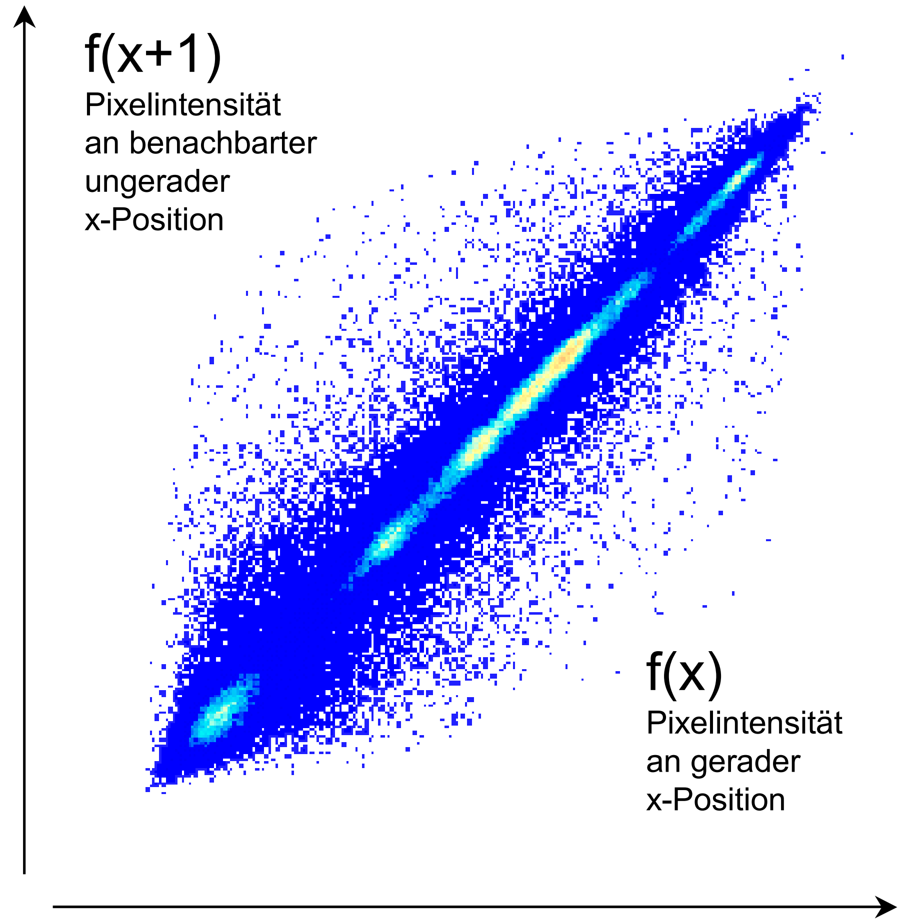
Huffman-Codierung

1. Ordnung der Symbole x nach ihrer Wahrscheinlichkeit p(x)
2. Zusammenfassen der beiden Symbole bzw. Symbolgruppen mit der geringsten Wahrscheinlichkeit bzw. Gruppenwahrscheinlichkeit
3. Dem Code der zusammengefassten Symbole wird eine 0 bzw. 1 hinzugefügt
4. Solange zusammenfassen, bis alle Symbole im Codebaum enthalten sind

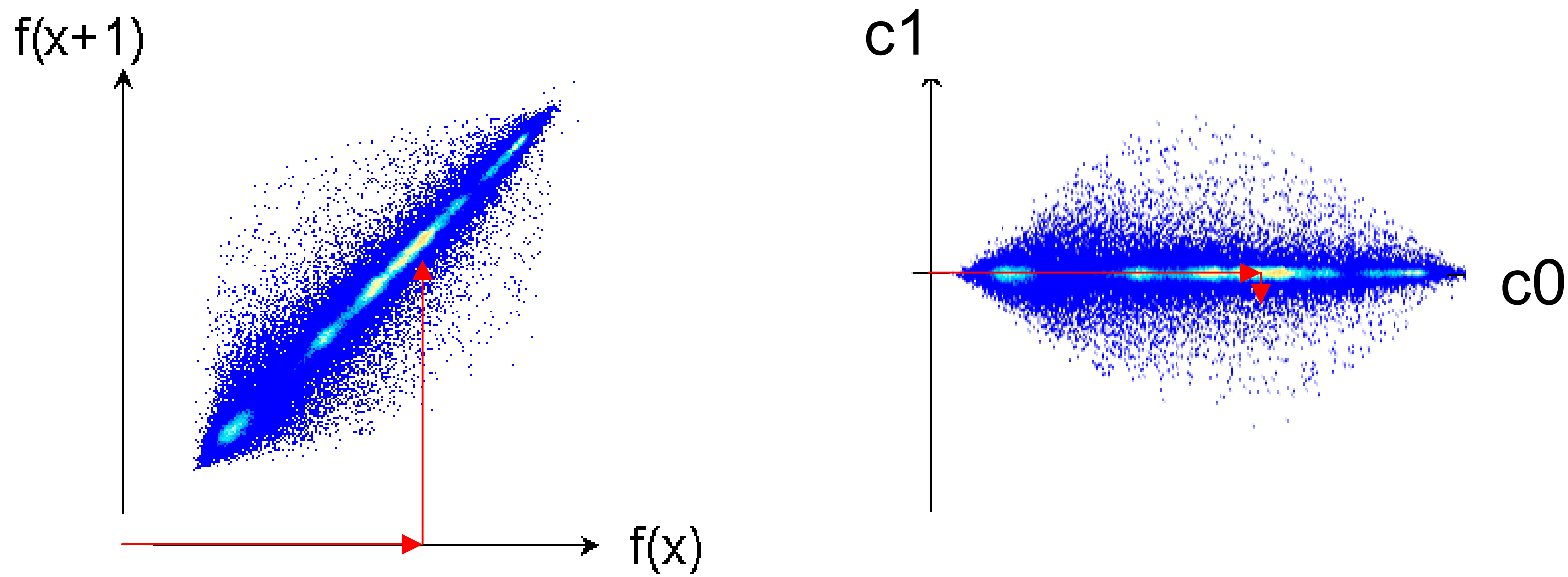
1D-Transformationscodierung

Graustufenbild

Verbundverteilungsdichte (Auftretenshäufigkeit)



1D-Transformationscodierung



Redundanzreduktion durch Transformation des Bildes
in ein neues Koordinatensystem

1D-Transformationscodierung

Jedes Pixelpaar (f(x),f(x+1)) kann durch die Transformations-koeffizienten (c0,c1) neu beschrieben werden:

$$\boxed{9 \mid 7} = 9 * \boxed{1 \mid 0} + 7 * \boxed{0 \mid 1}$$

(f(x),f(x+1))

$$\boxed{9 \mid 7} = 8 * \boxed{1 \mid 1} - 1 * \boxed{-1 \mid 1}$$

-> (8,-1)= (c0,c1)

2D-Transformationscodierung

$$\begin{array}{|c|c|}\hline 4 & 2 \\ \hline 2 & 4 \\ \hline\end{array} = 4 * \begin{array}{|c|c|}\hline 1 & 0 \\ \hline 0 & 0 \\ \hline\end{array} + 2 * \begin{array}{|c|c|}\hline 0 & 1 \\ \hline 0 & 0 \\ \hline\end{array} + 2 * \begin{array}{|c|c|}\hline 0 & 0 \\ \hline 1 & 0 \\ \hline\end{array} + 4 * \begin{array}{|c|c|}\hline 0 & 0 \\ \hline 0 & 1 \\ \hline\end{array}$$

$$\begin{array}{|c|c|}\hline 4 & 2 \\ \hline 2 & 4 \\ \hline\end{array} = c_0 * \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 1 & 1 \\ \hline\end{array} + c_1 * \begin{array}{|c|c|}\hline 1 & 1 \\ \hline -1 & -1 \\ \hline\end{array} + c_2 * \begin{array}{|c|c|}\hline 1 & -1 \\ \hline 1 & -1 \\ \hline\end{array} + c_3 * \begin{array}{|c|c|}\hline 1 & -1 \\ \hline -1 & 1 \\ \hline\end{array}$$

- Jeder 2x2 Block lässt sich durch Überlagerung von vier 2x2 Basisblöcken beschreiben
- Das obere Beispiel zeigt die übliche Beschreibung mittels Einheitsvektoren, das untere die Beschreibung mittels Basisbildern
- Die Abbildung der Werte (4,2,2,4) auf (c_0,c_1,c_2,c_3) heißt *Transformation*

2D-Transformationscodierung

$$\begin{array}{|c|c|}\hline 4 & 2 \\ \hline 2 & 4 \\ \hline\end{array} = 4 * \begin{array}{|c|c|}\hline 1 & 0 \\ \hline 0 & 0 \\ \hline\end{array} + 2 * \begin{array}{|c|c|}\hline 0 & 1 \\ \hline 0 & 0 \\ \hline\end{array} + 2 * \begin{array}{|c|c|}\hline 0 & 0 \\ \hline 1 & 0 \\ \hline\end{array} + 4* \begin{array}{|c|c|}\hline 0 & 0 \\ \hline 0 & 1 \\ \hline\end{array}$$

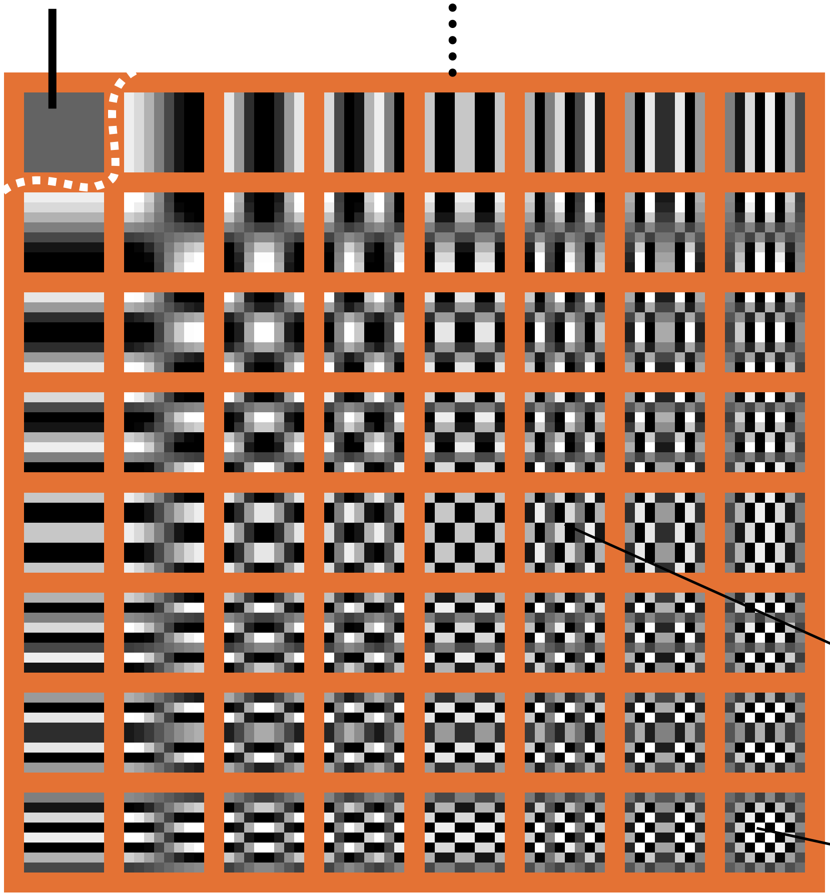
$$\begin{array}{|c|c|}\hline 4 & 2 \\ \hline 2 & 4 \\ \hline\end{array} = 3 * \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 1 & 1 \\ \hline\end{array} + 0 * \begin{array}{|c|c|}\hline 1 & 1 \\ \hline -1 & -1 \\ \hline\end{array} + 0 * \begin{array}{|c|c|}\hline 1 & -1 \\ \hline 1 & -1 \\ \hline\end{array} + 1* \begin{array}{|c|c|}\hline 1 & -1 \\ \hline -1 & 1 \\ \hline\end{array}$$

- Jeder 2x2 Block lässt sich durch Überlagerung von vier 2x2 Basisblöcken beschreiben
- Das obere Beispiel zeigt die übliche Beschreibung mittels Einheitsvektoren (Basisbildern), das untere die Beschreibung mittels Basisbildern
- Die Abbildung der Werte (4,2,2,4) auf (3,0,0,1) heißt *Transformation*

Approximation durch Basisbilder/-funktionen

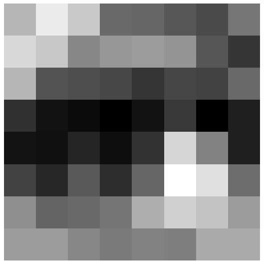
Gleichanteil DC

Wechselanteil AC

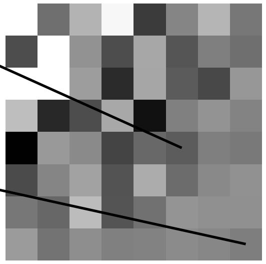


Diskrete Cosinus Transformation (DCT) - Basisbilder

Bildblock 8x8

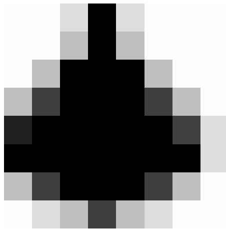


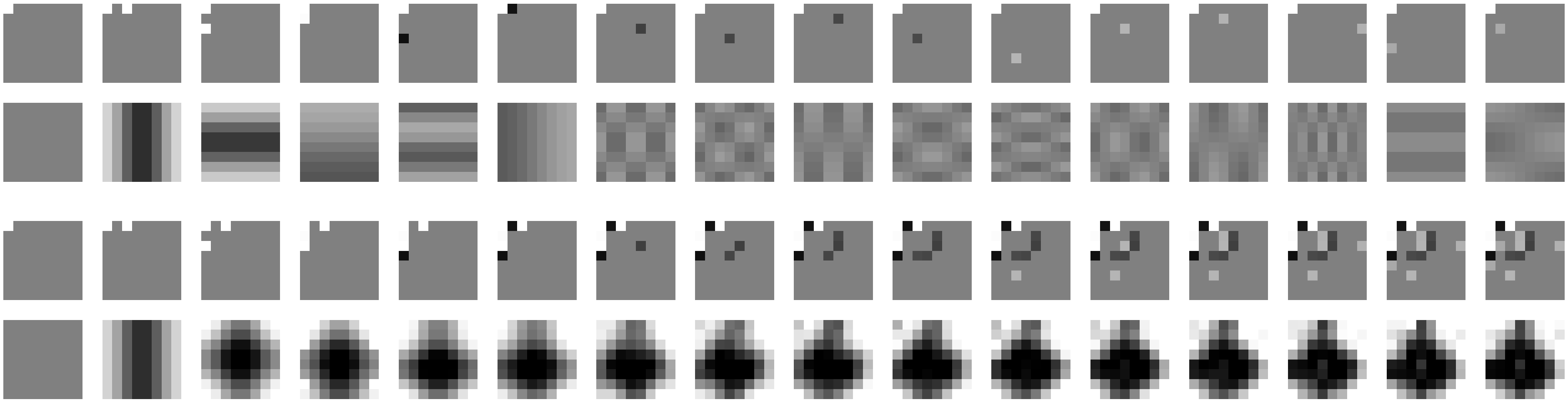
8x8 Transformationskoeffizienten geben Gewichtung der 8x8-Basisbilder an (Je heller der Koeffizient, desto größer der Wert)



Approximation durch sukzessive Überlagerung von Basisbildern

- Ein Bildblock kann häufig durch wenige Koeffizienten (*n*) relativ gut approximiert werden.
- Um den Fehler möglichst gering zu halten, werden die *n* Koeffizienten mit höchster Energie ausgewählt.





Approximation mit *n*=16 von maximal N=64 Koeffizienten

Vorteile der Transformation für die Codierung

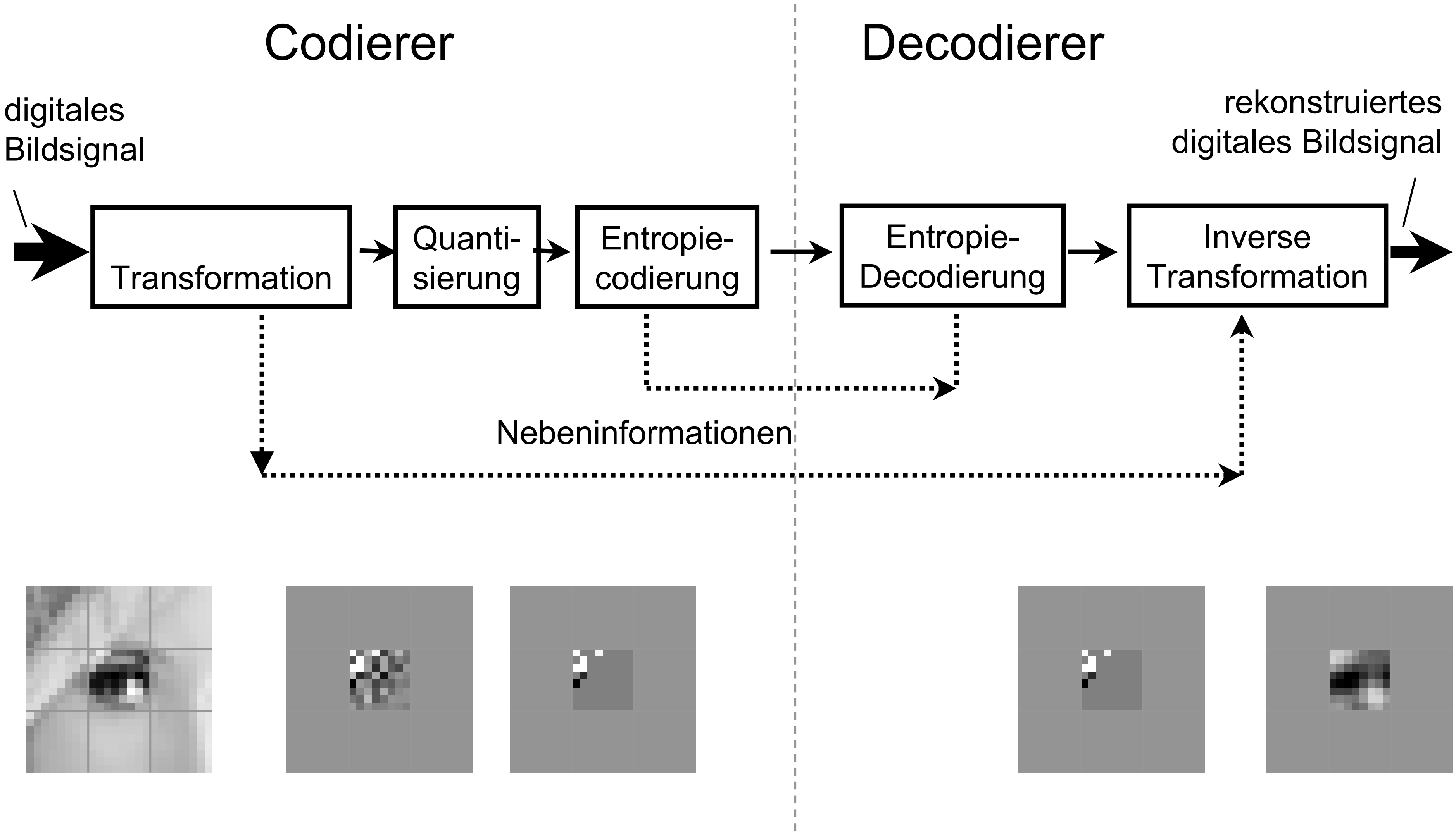
- Konzentration der Varianz des Bildsignals auf wenige, i.d.R. tieffrequente DCT-Koeffizienten

  -> Weniger Koeffizienten müssen codiert werden

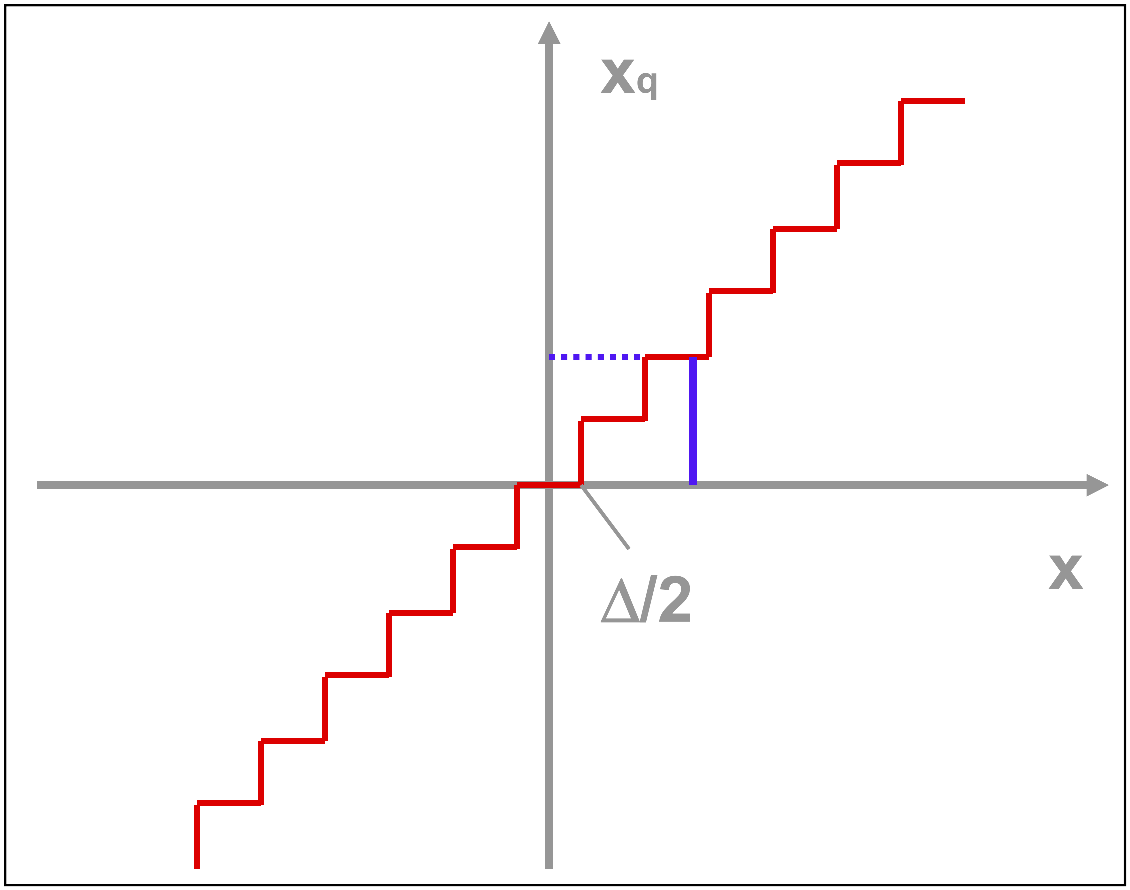
  -> Der Wertebereich der hochfrequenten Amplituden kann ohne große Verluste stark eingeschränkt werden, z.B.

Grauwerte x:	$0< x <255$	benötigen 8 bit/Pixel
Koeffizienten c:	$0< c <16$	benötigen 4 bit/Koeff.

Transformationscodierung (JPEG, MPEG)

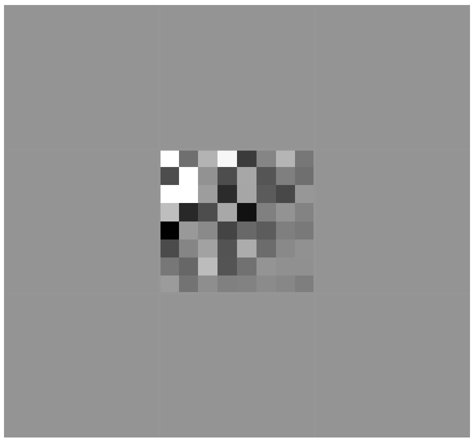


Quantisierungskennlinie

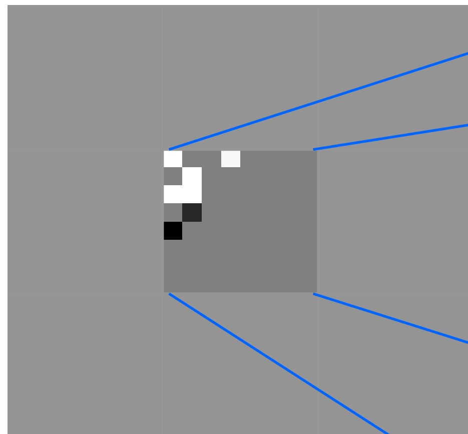


Gleichförmige Quantisierung mit Quantisiererstufenhöhe Δ
Bsp: Werte $3\Delta/2 < x \leq 5\Delta/2$ werden mit $x_q = 2\Delta$ rekonstruiert

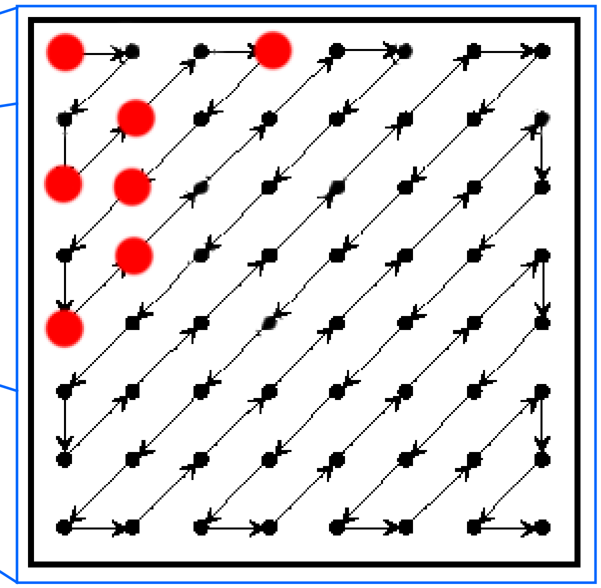
Entropiecodierung von JPEG (1992)



DCT-Koeffizienten
eines 8x8 Blocks



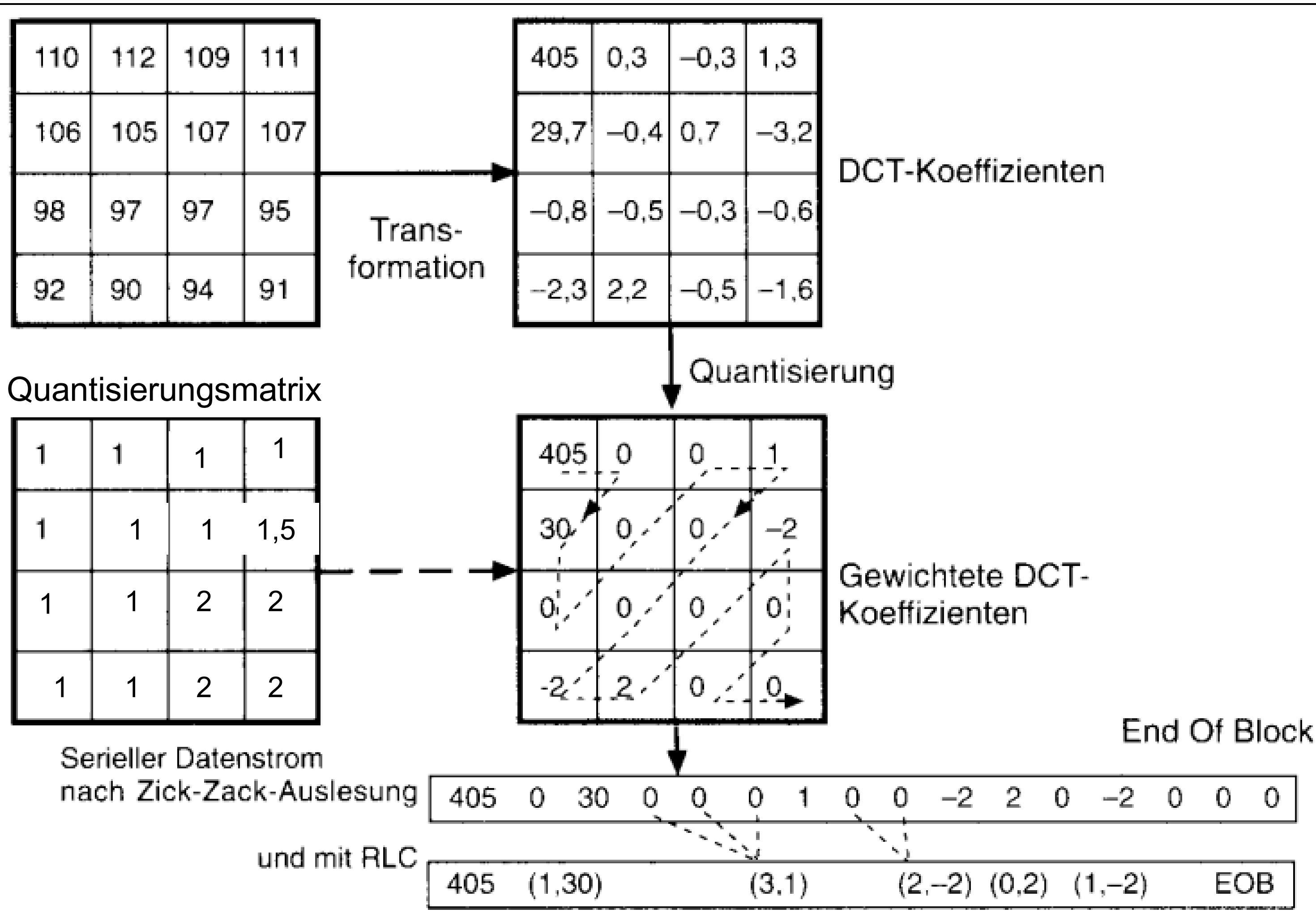
quantisierte
Koeffizienten



Zigzag Scan

Viele Koeffizienten werden auf den Wert *Null quantisiert.*
Eine **Lauflängencodierung** überführt die 2D-Anordnung der Koeffizienten mittels *Zigzag-Scan* in eine 1D-Reihenfolge, wobei der Abstand zwischen Koeffizienten ungleich *Null* und der Wert des Koeffizienten bei JPEG gemeinsam codiert.

Entropiecodierung von JPEG (1992)



Lauflängencodierung (RLC) mit Wertepaaren (Länge,Quantisiererindex)