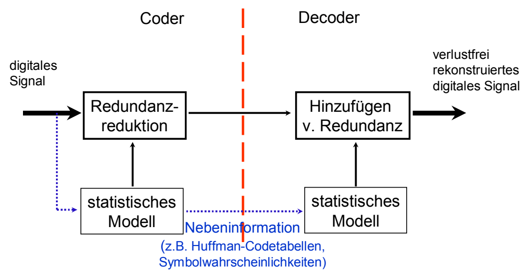
Grundlagen der Bildkompression (Teil 2)

Dr. Detlev Marpe
Fraunhofer Institut
für Nachrichtentechnik HHI

Fachhochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

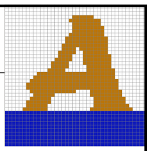
Fraunhofer Institut Nachrichtentechnik Heinrich-Hertz-Institut

Prinzip der Entropiecodierung



Je besser das statistische Modell an das Signal angepasst ist, umso größer ist die erzielte Redundanzreduktion

Beispiele zur statistischen Modellierung



Statistisches Modell 0.ter Ordnung:

3 Farbwerte:	A (**w**eiß)	B (**g**elb)	C (**b**lau)
$p(j)$	0.5	0.25	0.25

$\Rightarrow$ mittlere Codelänge = **1.5 bits/pel** (siehe VL 10.4.06)

Statistisches Modell 1.ter Ordnung:

Zusammenfassung zu Paaren von je zwei Bildpunkten (siehe VL 10.4.06):
3*3=9 Symbole: ww, wg, wb, gg, gb, gw, bb, bw, bg
$\Rightarrow$ mittlere Codelänge = **1.87 bits/Pixelpaar = 0.935 bits/pel**

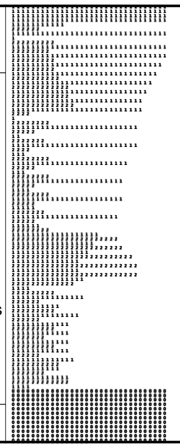
Statistisches Modell *N*.ter Ordnung:

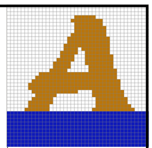
Zusammenfassung zu Blöcken **B** von $N+1$ Bildpkten

$$\frac{1}{N+1}\underbrace{H(\mathbf{B})}_{\substack{\text{Blockentropie}\\ \text{N.ter Ordnung}}} \leq \underbrace{H(B)}_{\substack{\text{Entropie}\\ \text{0.ter Ordnung}}}$$

Je größer die statistische Abhängigkeit, umso kleiner die Blockentropie!

Lauflängencodierung





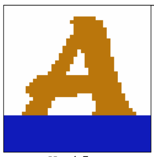
Zusammenfassung zu Blöcken variabler Länge aber gleichen Werts: (Wert, Lauflänge)

Beispiel:
mit 2 bits/Wert und 5 bits/Lauf

(1,32); (1,32); (1,32); (1,32);
(1,11); (2,6); (1,32); (1,1);(2,9);
...
(0, 32); (0, 32); (0, 32);(0,16)

mit 99 Läufen à 7 Bit = 693 Bits ergibt sich eine Rate von
0.433 bits/pel
bzw. eine (minimale)
Dateigröße von **87 Bytes**

Idee der Kontextmodellierung

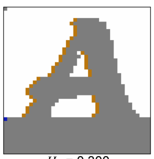
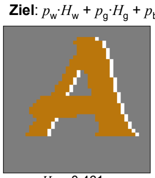
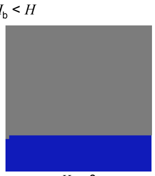


H = 1,5

Das Bild wird abhängig von Kontexten (hier z.B. 3 Kontexte aus der Farbe des vorhergehenden Pixels) in **disjunkte Teilbilder** zerlegt.

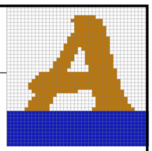
In den Teilbildern ergibt sich dann bei einer guten Kontextmodellierung eine **veränderte Wahrscheinlichkeitsverteilung** der Symbole, die dann eine **verringerte Entropie** besitzt.

Ziel: $p_w \cdot H_w + p_g \cdot H_g + p_b \cdot H_b < H$

=  +  + 

H_w = 0,300 | H_g = 0,461 | H_b = 0

Kontextmodellierung (der bedingten Wahrscheinlichkeiten)



Kontext w (p_w = 0.5):
$p_{w|w}$ = 759/800 = 0.94875
$p_{g|w}$ = 40/800 = 0.05
$p_{b|w}$ = 1/800 = 0.00125 $\quad H_w$ = 0.300

Kontext g (p_g = 0.25):
$p_{w|g}$ = 39/400 = 0.0975
$p_{g|g}$ = 361/400 = 0.9025
$p_{b|g}$ = 0/400 = 0 $\quad H_g$ = 0.461

Kontext b (p_b = 0.25):
$p_{w|b}$ = 0/399 = 0
$p_{g|b}$ = 0/399 = 0
$p_{b|b}$ = 399/399 = 1 $\quad H_b$ = 0

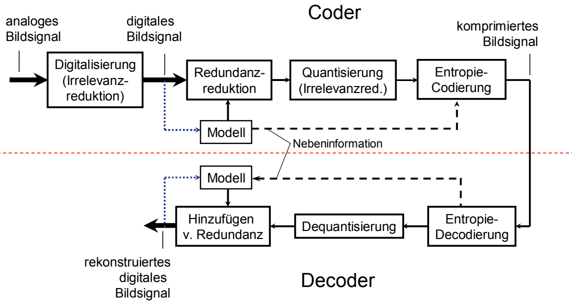
$H = p_w \cdot H_w + p_g \cdot H_g + p_b \cdot H_b = 0.2653$

Es ergibt sich eine theoretische Rate

von **0.2653 bits/pel**

bzw. eine (minimale) Dateigröße von **53 Bytes** !

Prinzip der verlustbehafteten Codierung

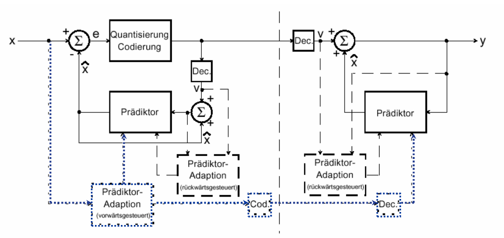


Prädiktive Codierung (Pixelbasiert)

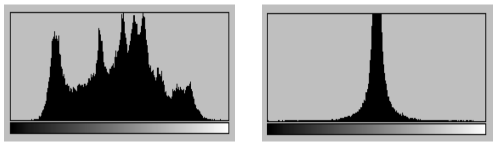
Idee: Redundanzreduktion durch **Vorhersage** des aktuellen Pixelwerts aus bekannten Pixelwerten der Vergangenheit

DPCM (Differential Pulse Code Modulation)

Coder Decoder



Wie lässt sich Prädiktionsqualität messen?



Ausgangsbildsignal:
Wertebereich: 0 – 255
⇒ 8 bits / pel

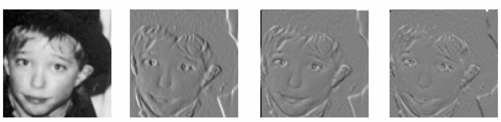
Prädiktionsfehlersignal:
Wertebereich: -255 bis 255
⇒ theoretisch 9 bits / pel

Auch ohne Quantisierung (verlustlose Codierung) erzielte Rate praktisch < 8 bits / pel, da **extrem ungleiche Verteilung** für Entropiecodierung von Vorteil

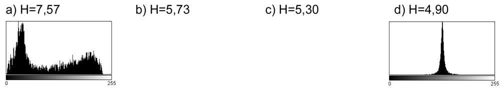
Beispiel: örtliche Prädiktion

Prädiktionsfehlerbilder

Originalbild | horizontal | horizontal + vertikal | adaptiv



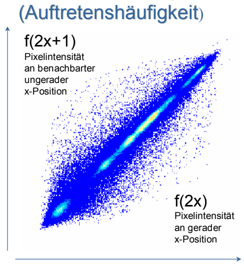
a) H=7,57 b) H=5,73 c) H=5,30 d) H=4,90



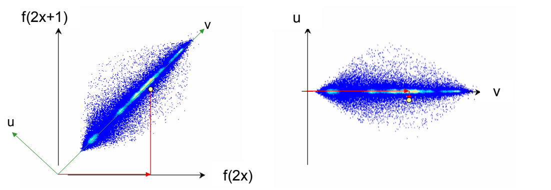
1D Transformationscodierung

Graustufenbild

Verbundverteilungsdichte (Auftretenshäufigkeit)



1D Transformationscodierung



Idee: Redundanzreduktion durch Transformation des Bildes in ein neues Koordinatensystem

1D Transformationscodierung

Jedes Pixelpaar kann nun neu beschrieben werden
(99,97) = 99 • (1,0) + 97 • (0,1)
wird
(99,97) = 98 • (1,1) - 1 • (-1,1)

$$\boxed{99\,|\,97} = 99 \cdot \boxed{1\,|\,0} + 97 \cdot \boxed{0\,|\,1}$$

$$\boxed{99\,|\,97} = 98 \cdot \boxed{1\,|\,1} - 1 \cdot \boxed{-1\,|\,1}$$

2D Transformationscodierung

Jeder 2x2-Block lässt sich durch Superposition (Überlagerung) von vier 2x2-Basisblöcken beschreiben

$$\begin{bmatrix}44 & 46\\48 & 50\end{bmatrix} = 44 \cdot \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + 46 \cdot \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + 48 \cdot \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} + 50 \cdot \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

grau = 0
weiß = 1
schwarz = -1

$$\begin{bmatrix}44 & 46\\48 & 50\end{bmatrix} = 47 \cdot \begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} - 2 \cdot \begin{bmatrix}1 & 1\\-1 & -1\end{bmatrix} - 1 \cdot \begin{bmatrix}1 & -1\\1 & -1\end{bmatrix} + 0 \cdot \begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}$$

Beispiel in Vektorschreibweise

$$\begin{bmatrix}44 & 46\\48 & 50\end{bmatrix} \qquad \begin{pmatrix}44\\46\\48\\50\end{pmatrix} = 44 \cdot \begin{pmatrix}1\\0\\0\\0\end{pmatrix} + 46 \cdot \begin{pmatrix}0\\1\\0\\0\end{pmatrix} + 48 \cdot \begin{pmatrix}0\\0\\1\\0\end{pmatrix} + 50 \cdot \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

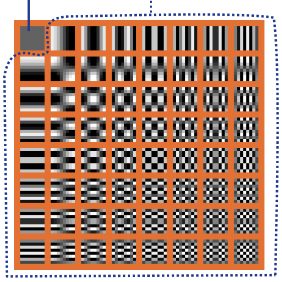
2x2 Bildblock

$$\begin{pmatrix}44\\46\\48\\50\end{pmatrix} = 47 \cdot \begin{pmatrix}1\\1\\1\\1\end{pmatrix} - 2 \cdot \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix} - 1 \cdot \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix} + 0 \cdot \begin{pmatrix}1\\-1\\-1\\1\end{pmatrix} \qquad \begin{bmatrix}47 & -1\\-2 & 0\end{bmatrix}$$

2x2 **Transformationskoeffizienten**

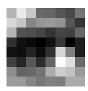
Diskrete Cosinus Transformation (DCT)

Gleichanteil DC   Wechselanteil AC



8x8 DCT-Basisbilder

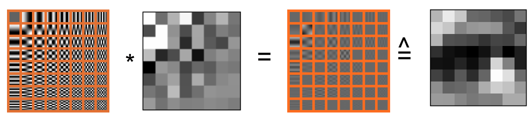
8x8 Bildblock



8x8 Transformations-koeffizienten

geben Gewichtung der 8x8 DCT-Basisbilder an

Funktionsweise und Vorteile einer Transformation

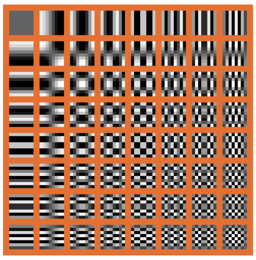


Basisbilder | Transformations-koeffizienten | gewichtete Basisbilder | Überlagerung ergibt Bildblock

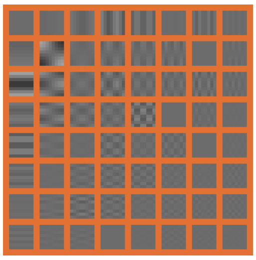
- Konzentration der Energie des Bildsignals auf wenige Transformationskoeffizienten
- Nicht alle Transformationskoeffizienten müssen übertragen werden, um eine ausreichende Bildqualität zu erzielen
- Die übertragenen Transformationskoeffizienten müssen nicht mit voller Genauigkeit codiert werden

Überlagerung von DCT-Basisbildern

DCT-Basisbilder

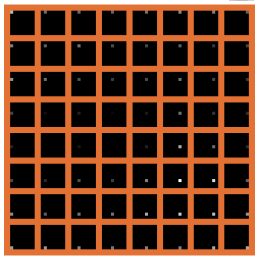


gewichtete Überlagerung

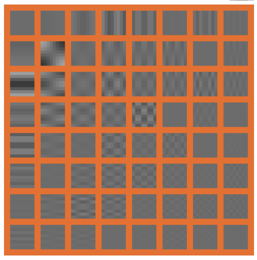


Approximation durch Überlagerung von ...

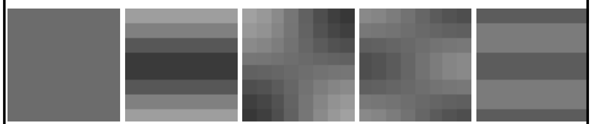
gewichteten 8x8 Ortsfunktionen

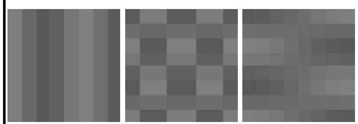


gewichteten 8x8 DCT-Basisfktn

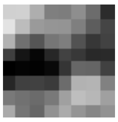


Sukzessive Approximation mit bis zu 8 von 64 Basisfktn





Approximierter 8x8 Bildblock

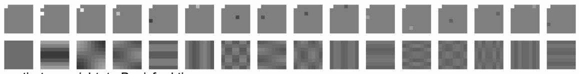


Approximation: eine Näherung mit dem Ziel, den Fehler, der durch die Näherung entsteht, zu minimieren

sukzessiv: aufeinanderfolgend, schrittweise

Sukzessive Approximation mit bis zu 16 von 64 Basisfktn

sortierte einzelne Gewichte / Transformationskoeffizienten



sortierte gewichtete Basisfunktionen

absteigende Energie der Koeffizienten

überlagerte Gewichte / Transformationskoeffizienten

überlagerte gewichtete Basisfunktionen

Um den Approximationsfehler zu minimieren, werden die 16 Koeffizienten mit höchster Energie ausgewählt und in absteigender Energie sortiert bzw. überlagert

Zusammenfassung der Grundlagen (Teil 2)

- Die Effizienz einer **Entropiecodierung** wird durch das zugrunde liegende **statistische Modell** bestimmt, wobei Codiermethoden auf Basis statistischer Modelle höherer Ordnung in der Regel eine erhöhte Effizienz ermöglichen
  - Praktische Ansätze zur Ausnutzung statistischer Bindungen höherer Ordnung liefern z.B. die **Lauflängencodierung** und die **Kontextmodellierung**
- Das Prinzip der **verlustbehafteten Codierung** beinhaltet eine Irrelevanzreduktion, die im allgemeinen in Form einer **Quantisierung** vorgenommen wird
- Die **prädiktive Codierung** (auch DPCM genannt) erlaubt eine Redundanzreduktion durch **Vorhersage** von Signalwerten mit Hilfe bereits codierter Signalanteile
- In der **Transformationscodierung** wird ein gegebenes Signal durch eine Transformation in ein **geeignetes Basissystem** überführt, und zwar so dass ...
  - die Signalenergie **auf wenige Transformationskoeffizienten konzentriert** wird
  - **nicht alle Transformationskoeffizienten** übertragen werden müssen, um eine ausreichende Bildqualität zu erzielen
  - die übertragenen Transformationskoeffizienten **nicht mit voller Genauigkeit** übertragen werden müssen, um eine ausreichende Bildqualität zu erzielen

