Grundlagen der Bildkompression (Teil 1)

Dr. Detlev Marpe
Fraunhofer Institut
für Nachrichtentechnik HHI

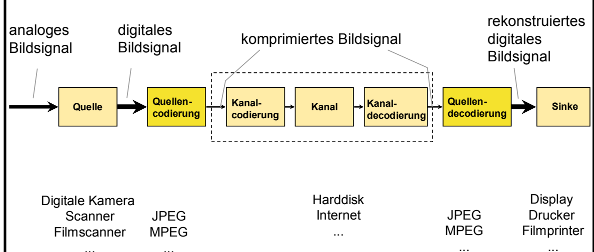
fhtw Fachhochschule für Technik und Wirtschaft Berlin University of Applied Sciences

Fraunhofer Institut Nachrichtentechnik Heinrich-Hertz-Institut

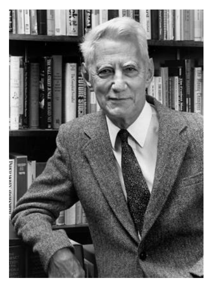
Themen Medientechnik II

- Grundlagen der statistischen Signalverarbeitung
  - Mittelwert, Varianz, SNR, ...
- **Bildcodierung, Bildcodierstandards und Bildformate**
  - **Grundlagen**
  - GIF, PNG
  - JPEG, JPEG-LS, JPEG 2000
- Videocodierung, Videoverarbeitung und Videocodierstandards
  - Grundlagen
  - Videocodierstandards H.261, MPEG1/2, H.263, MPEG4, H.264/MPEG4-AVC
- (Audiocodierung)

Digitale Bildübertragungsstrecke



Claude E. Shannon (1916 – 2001)



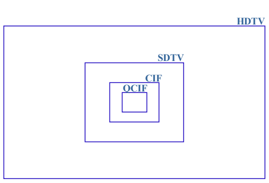
"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."

aus "A Mathematical Theory of Communication"[1948]

Datenmengen und Kanalkapazitäten

Format	Spalten	Zeilen	Bildfrequenz [Hz]	Bildgröße [kB]	Datenrate [Mb/s]
QCIF	176	144	5-15	38	1 - 4
CIF	352	288	10-30	152	10 - 30
SDTV	720	576	25i	311	124
HDTV	1920	1080	25i	1555	622

25i(nterlace) = 50 Halbbilder



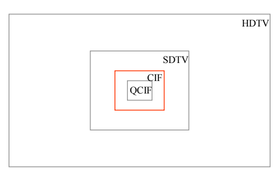
GSM	4 - 9.6 kb/s
ISDN	64 - 128 kb/s
DVB-H (pro Kanal)	~ 384 kb/s
UMTS	bis 2 Mb/s (var)
xDSL (downstream)	bis 50 Mb/s (var)
Ethernet	bis 4 Mb/s (var)
Fast Ethernet	bis 40 Mb/s (var)

Beispiel CIF über DVB-H (384 kb/s)

Format	Spalten	Zeilen	Bildfrequenz [Hz]	Bildgröße [kB]	Datenrate [Mb/s]
CIF	352	288	25	152	

(4:4:4) $352 \cdot 288 \cdot 3 \cdot 25 \cdot 8$ = 60.825.600 b/s

(4:2:0) $352 \cdot 288 \cdot 1{,}5 \cdot 25 \cdot 8$ = 30.412.800 b/s

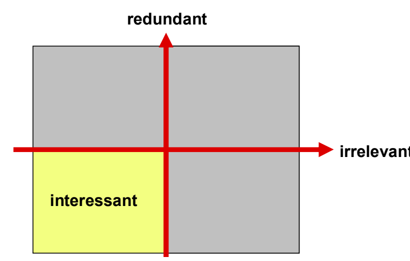


<u>erforderliche Kompressionsfaktoren</u>:
60.825.600 / 384.000 = **158 : 1**
bzw.
30.412.800 / 384.000 = **79 : 1**

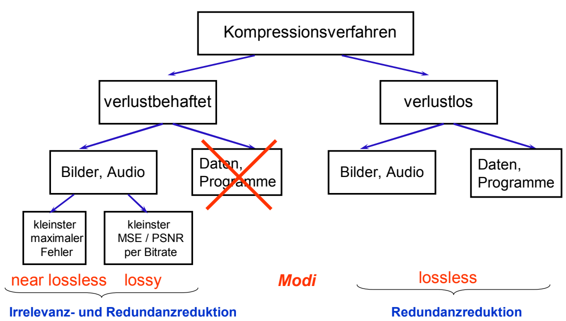
Bild- und Videodaten sind redundant !

Nachrichtenebene (nach Schouten)

Nachricht: **Information + Redundanz**
Irrelevanz: Information, die für Empfänger (Senke) **ohne Bedeutung**



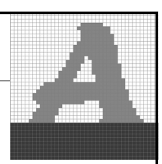
Klassifikation von Kompressionsverfahren



Parameter eines Kompressionsverfahrens

- Datenrate / Kompressionsfaktor
- Verzerrung / Qualität
- Komplexität
- Verzögerungszeit
- Robustheit gegenüber Übertragungsfehlern

Informationsgehalt eines Bildes ?



Graustufenbild $\Rightarrow$ 8 Bits pro Pixel

Geht es besser ? Ja!

Nur N = 3 Intensitäten $\Rightarrow$ 2 Bits pro Pixel reichen auch:

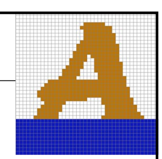
z.B.	dunkelgrau :	00
	mittelgrau:	01
	weiß:	10

Geht es noch besser? Ja!

z.B.	dunkelgrau :	10
	mittelgrau:	11
	weiß:	0

Wie groß ist die mittlere Rate in Bits pro Pixel?

Informationsgehalt eines Bildes ?



3 Symbole A (weiß), B (gelb), C (blau)
p(A) = 800/1600 = ½
p(B) = 400/1600 = ¼
p(C) = 400/1600 = ¼

Code 1 :	A: 0 B: 10 C: 11
Codelänge 1 [bit]:	A: 1 B: 2 C:2
mittlere Codelänge 1:	½·1 + ¼·2 + ¼·2 = 1,5 Bit/Pixel

Code 2 :	A: 10 B: 0 C: 11
Codelänge 2 [bit]:	A: 2 B: 1 C:2
mittlere Codelänge 2:	½·2 + ¼·1 + ¼·2 = 1,75 Bit/Pixel

Können wir den Informationsgehalt rechnerisch bestimmen?

Informationsgehalt I

[Shannon 1948]:

Eine (gedächtnislose*) Nachrichtenquelle mit J Symbolen j und den Wahrscheinlichkeiten $p(j)$ wird dann optimal übertragen, wenn die Länge der einzelnen Codes entsprechend

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

gewählt wird.

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

*Eine gedächntnislose Quelle führt Buch über die Wahrscheinlichkeitsverteilung der Quellsymbole, hat aber kein Gedächtnis bzgl. der Werte vorangeganger Symbole

Entropie H

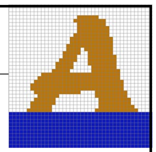
[Shannon 1948]:

Die *Entropie* H gibt die minimale Anzahl von Bits an, die im Durchschnitt zur Übertragung einer Nachricht mit J unterschiedlichen Symbolen mit den Wahrscheinlichkeiten $p(j)$ nötig ist:

$$H = -\sum_{j=1}^{J} p(j) \cdot \log_2 p(j)$$

$$= \sum_{j=1}^{J} p(j) \cdot I(j) \left[\frac{\text{bit}}{\text{symbol}}\right]$$

Entropie eines Bildes (1. Beispiel)



3 Symbole j	A (**w**eiß)	B (**g**elb)	C (**b**lau)
$p(j)$	0.5	0.25	0.25
$-\log_2(p(j))$	1	2	2
Code 1	0	10	11
Codelänge	1	2	2

⇒ Code 1 ist optimal, da mittlere Codelänge = Entropie

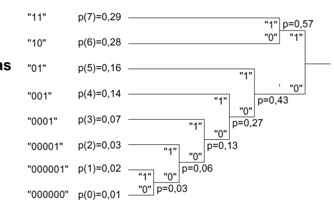
Symbol	N	p	Code	Codelänge	p*Codelänge
w	800	0.5	0	1	0.5
g	400	0.25	10	2	0.5
b	400	0.25	11	2	0.5

mittlere Codelänge = **1.5 bits/pel**

Dateigröße = mittlere Codelänge * Symbolanzahl / 8 = 1.5 * (40*40)/8
≥ **300 Bytes** (Gleichheitszeichen gilt nur dann, wenn der Aufwand für die Übertragung der Codetabelle vernachlässigt wird)

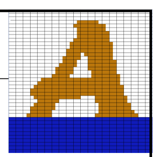
Code minimaler mittlerer Codelänge: Huffman Code

Prinzipien:
- **Je wahrscheinlicher** ein Symbol, **desto kürzer ist das** zugewiesene **Codewort**
- Kein Codewort darf Präfix eines anderen Codeworts sein **(Präfixfreier Code)**



Verfahren:
1. Ordnung der Symbole j nach ihrer Wahrscheinlichkeit $p(j)$
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

Entropie eines Bildes, 2. Beispiel a)

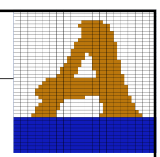


Zusammenfassung von je zwei Symbolen:
9 Symbole: ww, wg, wb, gg, gb, gw, bb, bw, bg

Symbol	$p_k \cdot p_l$	p_{kl}	Code	Codelänge	$p_{kl} \cdot$ Codelänge
ww	0.5 · 0.5	0.25	00	2	0.5
wb	0.5 · 0.25	0.125	010	3	0.375
wg	0.5 · 0.25	0.125	011	3	0.375
gw	0.25 · 0.5	0.125	100	3	0.375
gb	0.25 · 0.25	0.0625	1100	4	0.25
gg	0.25 · 0.25	0.0625	1111	4	0.25
bw	0.25 · 0.5	0.125	101	3	0.375
bb	0.25 · 0.25	0.0625	1110	4	0.25
bg	0.25 · 0.25	0.0625	1101	4	0.25

H = 3 bit/symbol, mittlere Codelänge = 3 bit/symbol
mittlere Codelänge = **1.5 bit/pel** Dateigröße ≥ **300 Bytes**

Entropie eines Bildes, 2. Beispiel b)

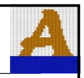


Zusammenfassung von je zwei Symbolen:
9 Symbole: ww, wg, wb, gg, gb, gw, bb, bw, bg

Symbol	N	p	Code	Codelänge	p · Codelänge
ww	372	0.465	0	1	0.465
wb	0	0			
wg	21	0.02625	1000	4	0.105
gw	19	0.02375	1001	4	0.095
gb	0	0			
gg	188	0.235	101	3	0.705
bw	0	0			
bb	200	0.25	11	2	0.5
bg	0	0			

H = 1.77 bit/symbol, mittlere Codelänge = 1.87 bit/symbol
mittlere Codelänge = **0.935 bit/pel** Dateigröße ≥ **187 Bytes**

Verhältnis mittlere Codelänge / Entropie



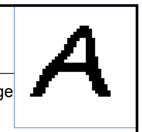
Symbol	N	p(j)	Code	Codelänge	p(j)*Codelänge
ww	372	0.465	0	1	0.465
wg	21	0.02625	1000	4	0.105
gw	19	0.02375	1001	4	0.095
gg	188	0.235	101	3	0.705
bb	200	0.25	11	2	0.5

mittlere Codelänge = **1.87 bit/symbol**

Symbol	N	p	I(j)	p*I(j)
ww	372	0.465	1.1047	0.51368428
wg	21	0.02625	5.2515	0.13785289
gw	19	0.02375	5.3959	0.12815331
gg	188	0.235	2.0893	0.49097782
bb	200	0.25	2.0000	0.5

Entropie H = **1.77 bit/symbol**

Verhältnis mittlere Codelänge / Entropie



Symbol	p(j)	Code	Codelänge	p(j)*Codelänge
w	0.875	1	1	0.875
s	0.125	0	1	0.125

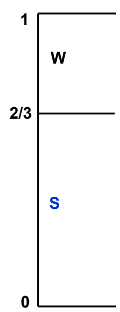
mittlere Codelänge = **1 bit/pel**

Symbol	p	I(j)	p*I(j)
w	0.875	0.1926	0.1686
s	0.125	3	0.375

Entropie H = **0.5436 bit/pel**

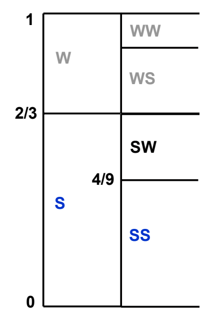
Allgemein gilt: **mittlere Codelänge ≥ 1 bit/symbol**

Arithmetische Codierung (Beispiel)



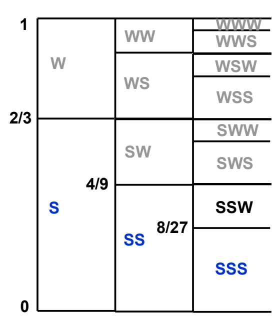
„SSS"
p(**W**)=1/3, p(**S**)=2/3

Arithmetische Codierung (Beispiel)



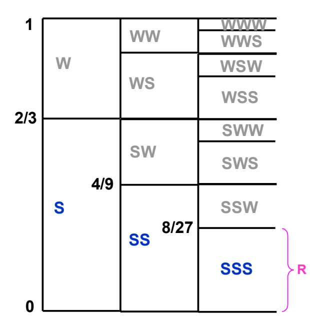
„SSS"
p(**W**)=1/3, p(**S**)=2/3

Arithmetische Codierung (Beispiel)



„SSS"
p(**W**)=1/3, p(**S**)=2/3

Arithmetische Codierung (Beispiel)



„SSS"
p(**W**)=1/3, p(**S**)=2/3

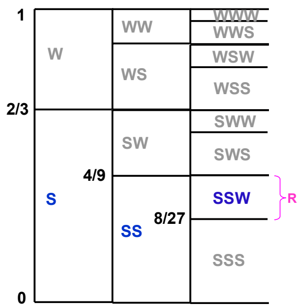
$I(\text{"SSS"}) = -\log_2 p(\text{"SSS"})$
$= -\log_2 (p(S)*p(S)*p(S))$
≈ 1.755 bit

Kürzeste Binärzahl (ohne "0.") im Endintervall = Codewort
→ 8/27 = 0.01001...

0.01

Intervallbreite R:
2/3 × 2/3 × 2/3 = **8/27** = p("SSS")
→ 0 = 0.00000..

Arithmetische Codierung (2. Beispiel)



„SSW“

p(W)=1/3, p(S)=2/3

$I(\text{"SSW"}) = -\log_2 p(\text{"SSW"})$

≈ 2.755 bit

$\rightarrow$ 4/9 = 0.01110...

$\rightarrow$ 8/27=0.01001...

Codewort

0.011

Intervallbreite R:

2/3 $\times$ 2/3 $\times$1/3 = **4/27** = p("SSW")

Zusammenfassung der Grundlagen (Teil 1)

- Einteilung in a) ***verlustlose*** u. b) ***verlustbehaftete*** Kompressionsverfahren
- a) nutzt ***Redundanz*** aus während b) zusätzlich ***Irrelevanz*** reduziert
- [Shannon, 1948] Für jeden Code C einer Nachrichtenquelle P mit J Symbolen j und Wahrscheinlichkeiten $p(j)$ gilt:

  mittlere Codelänge $(C) \geq \; H(P) = -\sum_{j=1}^{J} p(j) \cdot \log_2 p(j)$

  wobei $H(P)$ die ***Entropie*** (mittleren Informationsgehalt) von P bezeichnet.
- Redundanz $R(C)$ eines präfixfreien Codes C:

  $R(C)$ = mittl. Codelänge $(C) - H(P) \geq 0$
- ***Huffman Codes*** lassen sich mit Hilfe eines Codebaums erzeugen
- Unter allen präfixfreien Codes erzielt der Huffman Code minimale Redundanz
- Mittl. Codelänge eines präfixfreien Codes ist größer gleich 1 bit/symbol
- Mit der **arithmetischen Codierung** lässt sich eine Nachrichtenquelle P mit nur *einem* Codewort repräsentieren und zwar so, dass die theoretisch untere Schranke der Entropie $H(P)$ nahezu erreicht wird