

Rolf Block

Grundlagen der Bilddatenreduktion

- **Physiologische Grundlagen des Sehens**
- **Transformationscodierung mit der DCT**
- **Teilbandcodierung**
- **Waveletcodierung**

Grundlagen der Bilddatenreduktion

Physiologische Grundlagen des Sehens

Eigenschaften des menschlichen Auges:

Das Auge ist ein Sensor, der aus Blende und Linse besteht.

Lichtsensitive Zellen (Photorezeptoren) befinden sich auf der Netzhaut (Retina)

Stäbchen (ca. 120 Mio.):

für das Schwarz-Weiß-Sehen (auch bei geringer Helligkeit/Nachtsehen), max. bei 554nm (Tageslicht), 513nm (Dunkeladaptation)

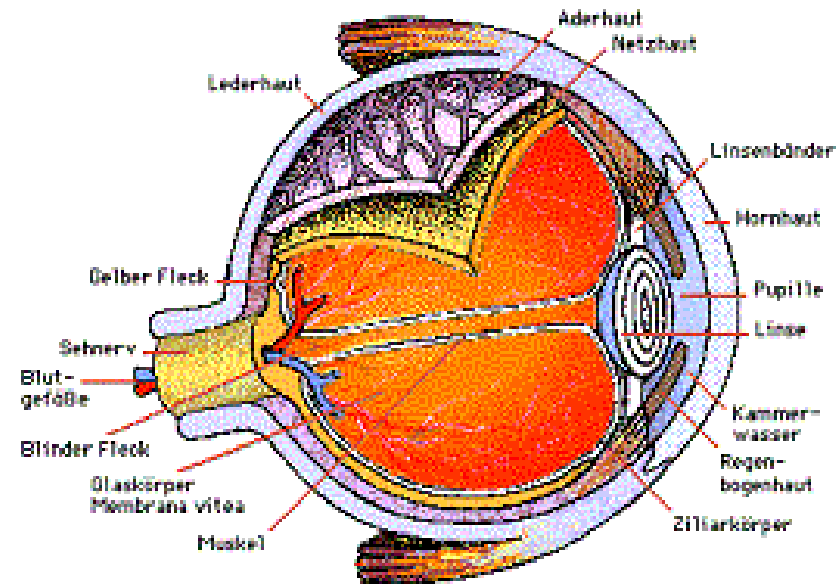
Zäpfchen (ca. 7 Mio.)

für das Farbsehen

blau-sensitiv 4% der Zäpfchen bei 430 nm max. Sensitivität

grün-sensitiv 32% der Zäpfchen bei 530 nm max. Sensitivität

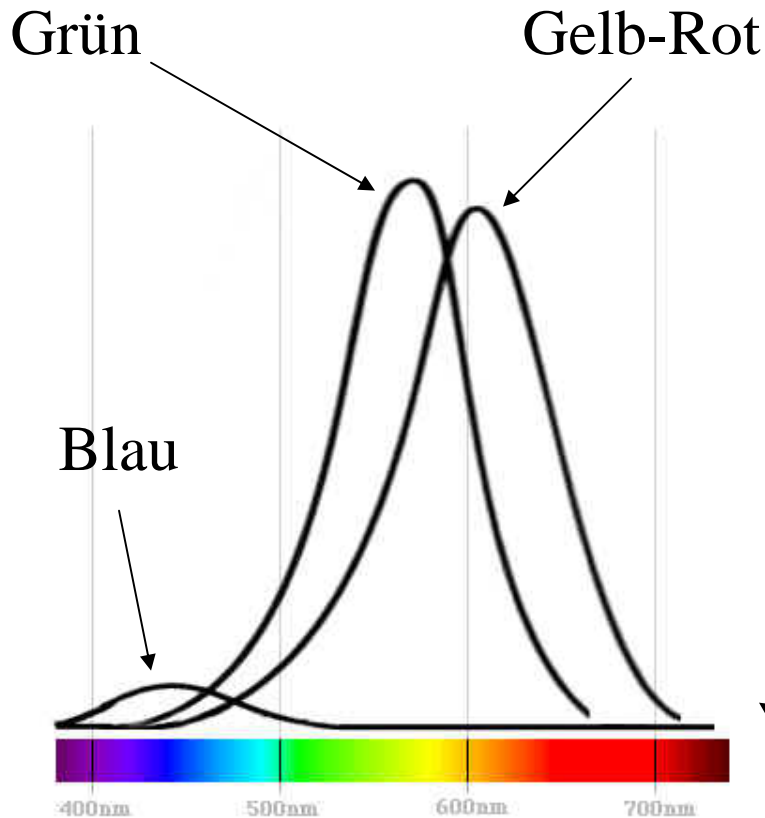
rot-sensitiv 64% der Zäpfchen bei 560 nm max. Sensitivität



Stäbchen und Zäpfchen setzen das Abbild in ein elektrisches Signal um, das von nachgeschalteten Neuronen weiterverarbeitet wird.

Grundlagen der Bilddatenreduktion

Physiologische Grundlagen des Sehens



- Die für das menschliche Auge empfindlichsten Farben sind :
- Grün (bei ca. 545 nm)
- Gelb-Rot (bei ca. 580nm)
Empfindlichkeit: 5% unter Grün
- Violettblau (bei ca. 440nm)
Empfindlichkeit: 1/30 Grün

Das nebenstehende Diagramm zeigt die verschiedenen Empfindlichkeitsspektren der Zapfchen.

Grundlagen der Bilddatenreduktion

Physiologische Grundlagen des Sehens

Die Retina spielt eine wichtige Rolle bei der Signalverarbeitung

Bevor das Signal in das Sehzentrum des Gehirns geleitet wird, erfolgt im Zwischenhirn eine Farbverarbeitung: aus den Signalen RGB werden drei neue Signale gebildet:

Differenzsignal	Rot-Grün = $R-G$	zur Rot-Grün- Unterscheidung
Differenzsignal	Gelb-Blau= $Y-B$	zur Blau-Gelb-Unterscheidung
Summensignal	Rot+Grün= Gelb	
	$R+G=Y$	zur Helligkeit- und Gelbwahrnehmung

Die Farbe Blau spielt für die Helligkeitswahrnehmung keine Rolle, es ist nicht möglich, Farbtöne rötlich-grün, bläulich-gelb wahrzunehmen

[Demo](#)

Grundlagen der Bilddatenreduktion

Physiologische Grundlagen des Sehens

Eigenschaften des menschlichen Auges:

- Örtliche Auflösung bei Farbsehen ist deutlich geringer als beim Helligkeitssehen (Farbkanten können nur schlecht wahrgenommen werden)
- Hohe örtliche Aktivität im Bild führt zu verringerter Gradationsauflösung (Amplitude kann vergrößert dargestellt werden)
- Horizontale und vertikale Kanten werden besser wahrgenommen als schräge Strukturen
- Auflösungsvermögen ist bei ruhenden Bildern am größten
 - Eine volle örtliche Auflösung wird nur bei ruhenden oder vorhersehbaren bewegten Bildinhalten benötigt.
 - Eine zeitliche volle Auflösung ist nur bei stark bewegten Bildinhalten notwendig.
 - hohe zeitliche Aktivität verringert ebenfalls die Amplitudenauflösung (örtlicher und zeitlicher Maskierungseffekt)
 - nach Szenenwechsel sind Störungen im Bild für ca. 1/15 Sekunden nicht wahrnehmbar

Farbsysteme

YUV-Modell:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

YCbCr-Modell

abgeleitet von YUV für das digitale Fernsehen:

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & -0.500 \\ 0.500 & -0.419 & -0.081 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$0 \leq Y \leq 255 \quad -127,5 \leq Cb, Cr \leq 127,5$$

Y-Komponente: im Bereich [16...235], schwarz bei 16

Cb und Cr werden durch offset in positiven Bereich verschoben

Bereich [16...240] mit Nullpegel bei 128

subsampling

Eigenschaften des menschlichen Auges:

Hohe örtliche Aktivität im Bild führt zu verringerter Gradationsauflösung

Horizontale und vertikale Kanten werden besser wahrgenommen
als schräge Strukturen

Transformationscodierung

Ein Pixelpaar x_1, x_2 lässt sich durch Überlagerung von 2 Basisblöcken beschreiben. In der üblichen Beschreibung sind die Basisblöcke Einheitsvektoren:

$$\begin{array}{c}
 \boxed{9} \ \boxed{7} \\
 \\
 \boxed{1} \ \boxed{0} \\
 \boxed{} \ \boxed{\times}
 \end{array}
 = 9 *
 \begin{array}{c}
 \boxed{0} \ \boxed{1} \\
 \boxed{\times} \ \boxed{}
 \end{array}
 + 7 *
 \begin{array}{c}
 \boxed{1} \ \boxed{0} \\
 \boxed{\times} \ \boxed{}
 \end{array}
 \qquad
 \begin{pmatrix} 9 \\ 7 \end{pmatrix} = 9 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 7 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Die Transformationscodierung bildet statt der Einheitsvektoren Basisbilder (Basisfunktionen), deren Überlagerung das ursprüngliche Pixelpaar bildet. Die Abbildung der Werte (x_1, x_2) auf die Koeffizienten (c_0, c_1) heißt Transformation.

$$\begin{array}{c}
 \boxed{9} \ \boxed{7} \\
 \\
 \boxed{1} \ \boxed{1} \\
 \boxed{} \ \boxed{}
 \end{array}
 = c_0 *
 \begin{array}{c}
 \boxed{1} \ \boxed{1} \\
 \boxed{} \ \boxed{}
 \end{array}
 + c_1 *
 \begin{array}{c}
 \boxed{-1} \ \boxed{1} \\
 \boxed{\blacksquare} \ \boxed{}
 \end{array}
 \qquad
 \begin{pmatrix} 9 \\ 7 \end{pmatrix} = 8 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$= 8 *
 \begin{array}{c}
 \boxed{1} \ \boxed{1} \\
 \boxed{} \ \boxed{}
 \end{array}
 - 1 *
 \begin{array}{c}
 \boxed{-1} \ \boxed{1} \\
 \boxed{} \ \boxed{}
 \end{array}$$

Die Koeffizienten können als Gewichtungsfaktoren für die Basisbilder gedeutet werden. Die Überlagerung der gewichteten Basisbilder wird als Rücktransformation bezeichnet und liefert das ursprüngliche Pixelpaar.

Die Erweiterung auf die 2-dimensionale Transformationscodierung zeigt die folgende Abbildung. Die Beschreibung durch Einheitsmatrizen ergibt:

$$\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} = 4 * \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2 * \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 2 * \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 4 * \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Die Beschreibung mit Basisbildern ergibt:

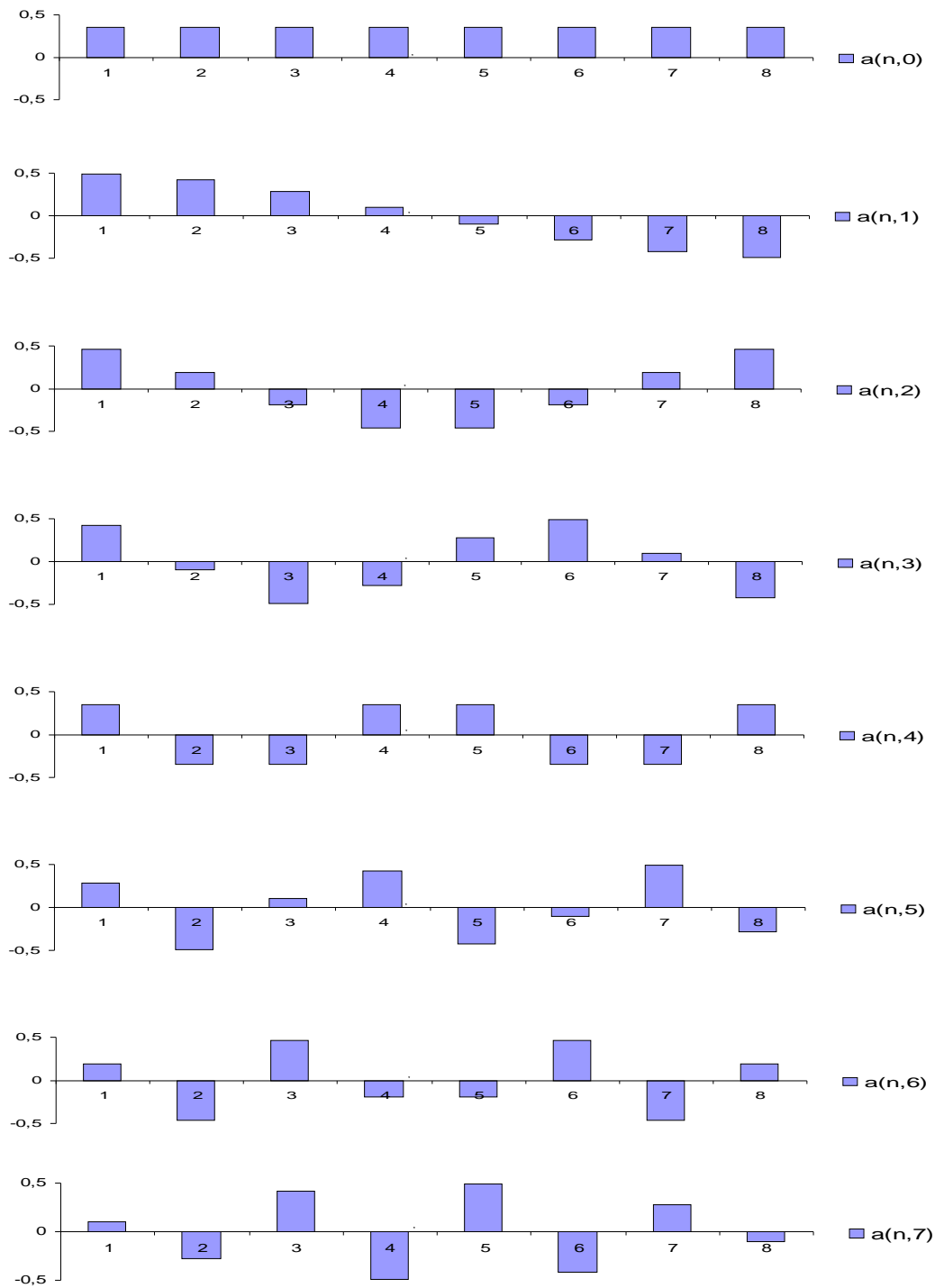
$$\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} = c_0 * \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + c_1 * \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} + c_2 * \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} + c_3 * \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= 3 * \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + 0 * \begin{bmatrix} \square & \square \\ \blacksquare & \blacksquare \end{bmatrix} + 0 * \begin{bmatrix} \square & \blacksquare \\ \square & \blacksquare \end{bmatrix} + 1 * \begin{bmatrix} \square & \blacksquare \\ \blacksquare & \square \end{bmatrix}$$

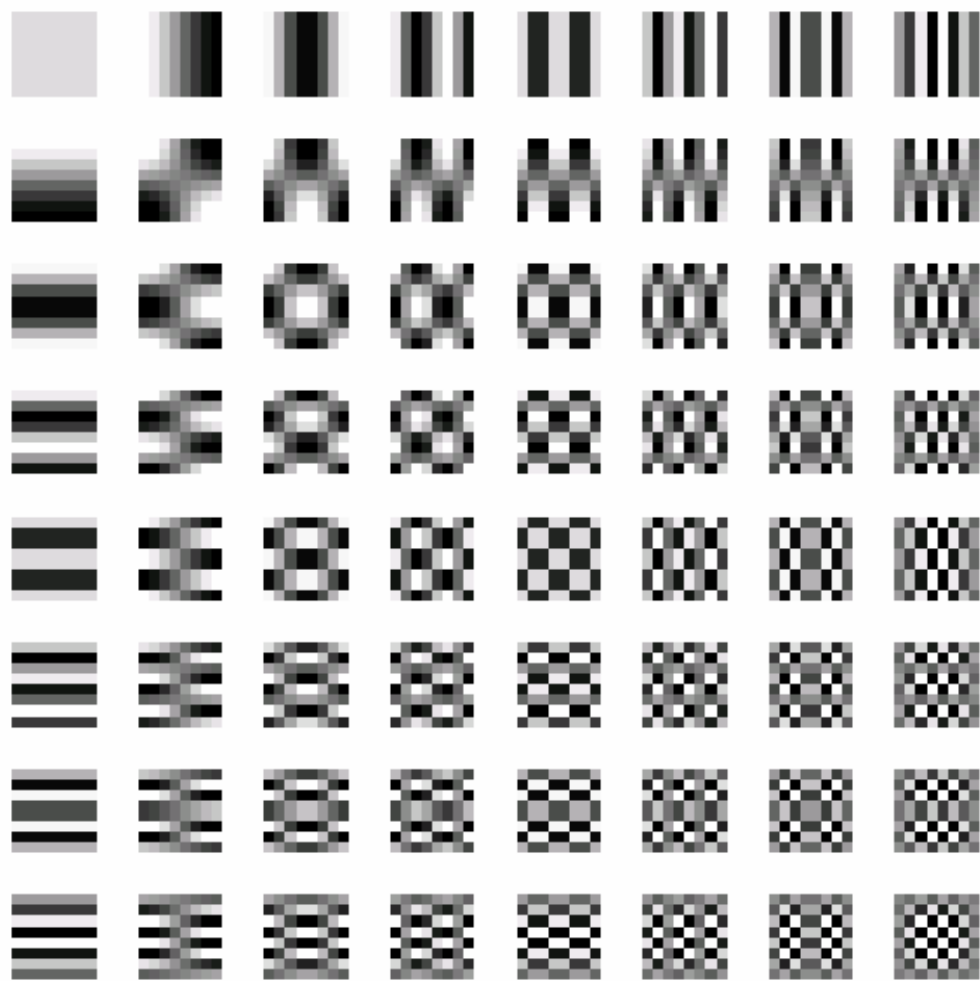
Die Basisbilder (Basisfunktionen) können durch Rücktransformation gewonnen werden, indem jeweils 1 Koeffizient zu Eins gesetzt wird und alle anderen zu Null gemacht werden.

Die verschiedenen Transformationen unterscheiden sich hauptsächlich durch die Basisfunktionen. Wenn das Originalsignal mit möglichst wenigen Basisfunktionen approximiert werden kann, ist eine starke Dekorrelation zu erzielen und damit eine effiziente Codierung.

Basisfunktionen der DCT

$$A = \begin{bmatrix} 0,35 & 0,35 & 0,35 & 0,35 & 0,35 & 0,35 & 0,35 & 0,35 \\ 0,49 & 0,42 & 0,28 & 0,10 & -0,10 & -0,28 & -0,42 & -0,49 \\ 0,46 & 0,19 & -0,19 & -0,46 & -0,46 & -0,19 & 0,19 & 0,46 \\ 0,42 & -0,10 & -0,49 & -0,28 & 0,28 & 0,49 & 0,10 & -0,42 \\ 0,35 & -0,35 & -0,35 & 0,35 & 0,35 & -0,35 & -0,35 & 0,35 \\ 0,28 & -0,49 & 0,10 & 0,42 & -0,42 & -0,10 & 0,49 & -0,28 \\ 0,19 & -0,46 & 0,46 & -0,19 & -0,19 & 0,46 & -0,46 & 0,19 \\ 0,10 & -0,42 & 0,42 & -0,49 & 0,49 & -0,42 & 0,28 & -0,10 \end{bmatrix}$$


Basisbilder der DCT



Quantisiermatrizen

Luminanzmatrix:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	72	95	98	112	100	103	99

Chrominanzmatrix:

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

$$l_{i,j} = \lceil T_{i,j} / Q_{i,j} + 0,5 \rceil$$

Beispiel: DC-Luminanzkomponente $T_{0,0} = 39,88$

$$l_{0,0} = \lceil T_{0,0} / Q_{0,0} + 0,5 \rceil = \lceil 39,88 / 16 + 0,5 \rceil = 2$$

Die Rekonstruktion r liefert:

$$r_{0,0} = l_{0,0} \cdot Q_{0,0} = 2 \cdot 16 = 32 .$$

Demo

Teilbandcodierung

Die Teilbandcodierung soll an einem einfachen Beispiel erläutert werden. Dazu wird das Eingangssignal mit zwei Filtern in einen niederfrequenten und einen hochfrequenten Anteil zerlegt. Der **niederfrequente** Anteil wird durch Mittelwertbildung (Integration) berechnet:

$$y[n] = (x[n] + x[n-1]) / 2 .$$

Der **hochfrequente** Anteil wird durch Differenzbildung (Differentiation) berechnet:

$$z[n] = (x[n] - x[n-1]) / 2 .$$

Das ursprüngliche Signal kann vollständig aus jedem 2. Wert der Folge $y[n]$ und $z[n]$ berechnet werden:

$$x[n-1] = z[n] - y[n]$$

$$x[n] = z[n] + y[n]$$

Ein **Beispiel** zeigt folgende Tabelle:

$x[n]$	10	14	10	12	14	8	14	12	10	8	10	12
$(x[n]-x[n-1])/2$	5	2	-2	1	1	-3	-3	-1	-1	-1	1	1
$(x[n]+x[n-1])/2$	5	12	12	11	13	11	11	13	11	9	9	11
	10	14	10	12	14	8	14	12	10	8	10	12

Nach Aufteilung in 2 Bänder braucht also nur jeder zweite Wert (in den grau unterlegten Feldern) in jedem Teilband übertragen werden. Für beide Teilbänder können unterschiedliche Codierer verwendet werden. Die Mittelwertbildung glättet das Signal, so dass DPCM-Techniken (LPC-Verfahren) gute Kompressionsergebnisse liefern. Die Differenzenfolge zeigt eine kleinere Varianz und kann ggf. mit einem anderen Codierverfahren (z.B. Vektorquantisierung) optimaler verarbeitet werden.

Teilbandcodierung für zweidimensionale Signale

Beispiel:

10	14	10	12	14	8	14	12
10	12	8	12	10	6	10	12
12	10	8	6	8	10	12	14
8	6	4	6	4	6	8	10
14	12	10	8	6	4	6	8
12	8	12	10	6	6	6	6
12	10	6	6	6	6	6	6
6	6	6	6	6	6	6	6

Das Bild wird zunächst zeilenweise tiefpassgefiltert (Mittelwertbildung) und hochpassgefiltert (Differenzbildung).

$$(x[n]+x[n-1])/2$$

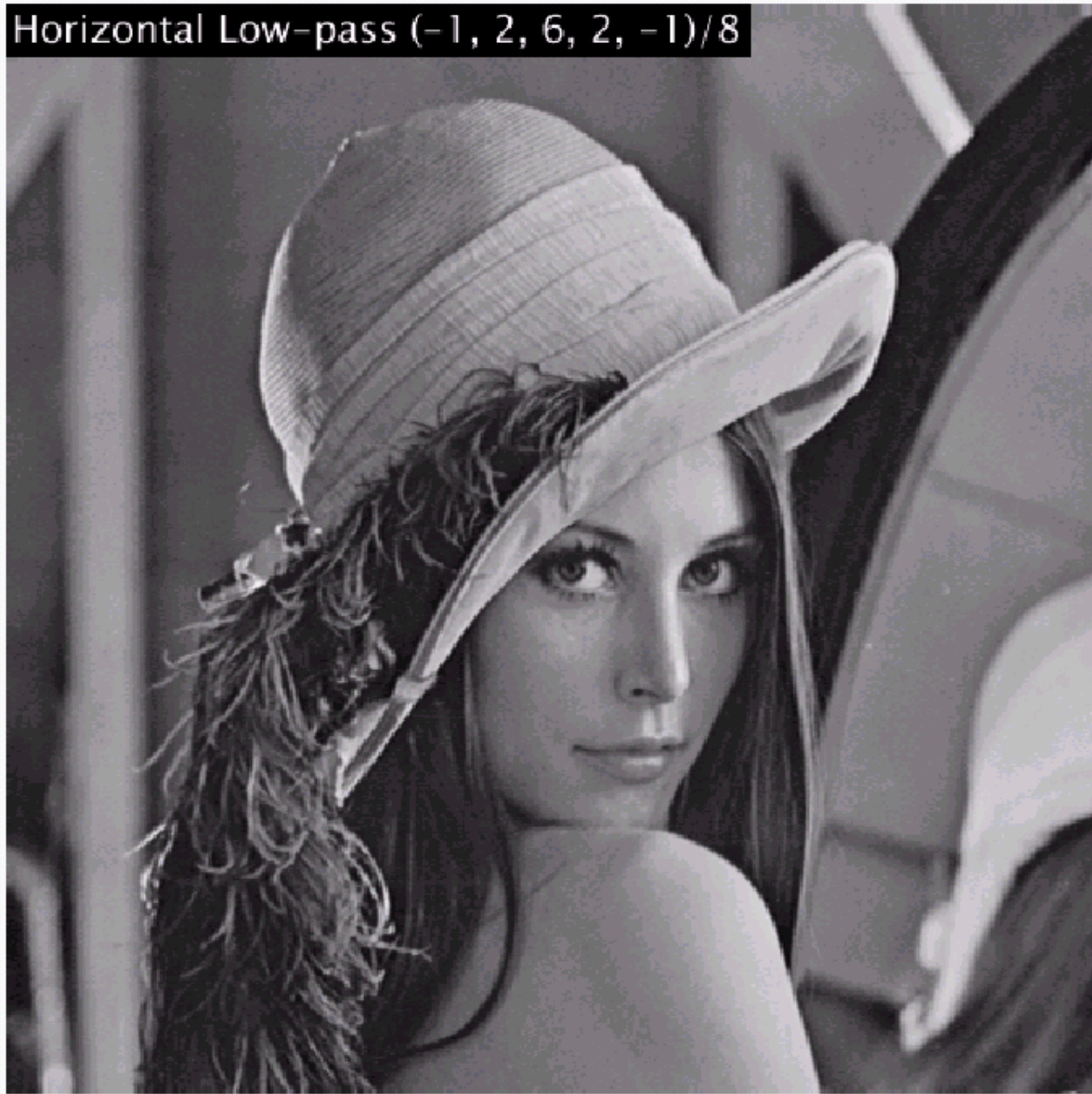
5	12	12	11	13	11	11	13
5	11	10	10	11	8	8	11
6	11	9	7	7	9	11	13
4	7	5	5	5	5	7	9
7	13	11	9	7	5	5	7
6	10	10	11	8	6	6	6
6	11	8	6	6	6	6	6
3	6	6	6	6	6	6	6

$$(x[n]-x[n-1])/2 \quad \longrightarrow n$$

5	2	-2	1	1	-3	3	-1
5	1	-2	2	-1	-2	2	1
6	-1	-1	-1	1	1	1	1
4	-1	-1	1	-1	1	1	1
7	-1	-1	-1	-1	-1	1	1
6	-2	2	-1	-2	0	0	0
6	-1	-2	0	0	0	0	0
3	0	0	0	0	0	0	0



Horizontal Low-pass $(-1, 2, 6, 2, -1)/8$



Horizontal High-pass $(-1, 2, -1)/2$, Scaled up 4X



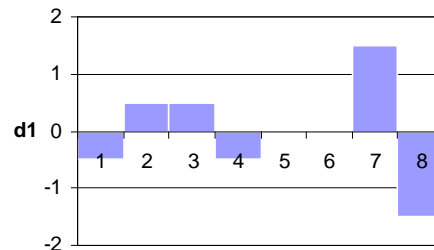
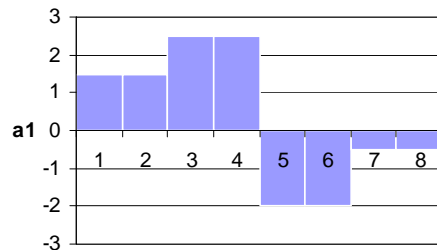
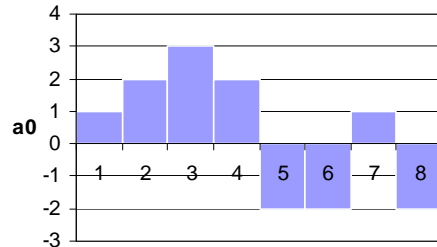




Demo

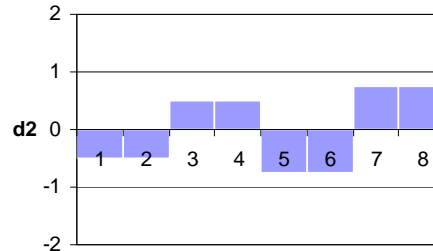
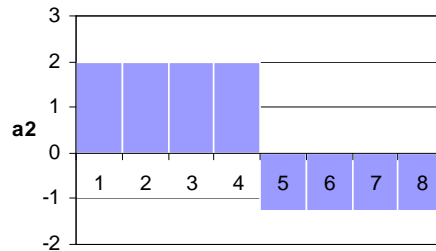
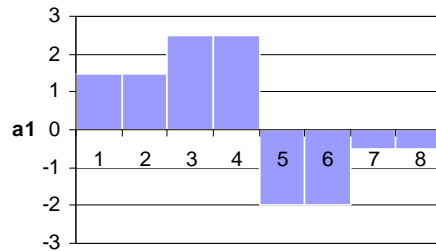
Vom Teilbandfilter zur Wavelettransformation

Ausgangspunkt ist der Signalvektor $a_0 = \{1 \ 2 \ 3 \ 2 \ -2 \ -2 \ 1 \ -2\}$ mit 8 Werten (s. Bild unten).

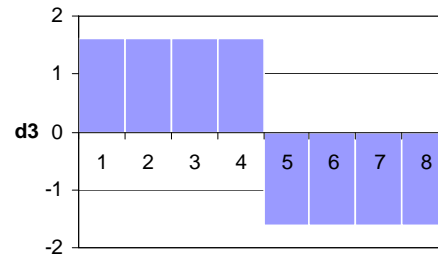
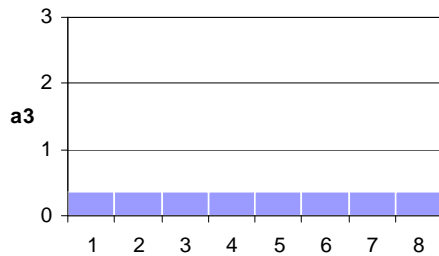
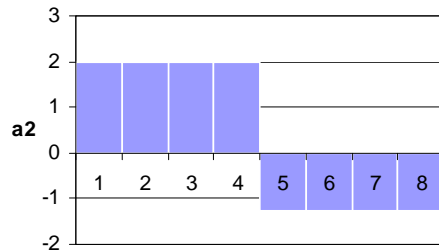


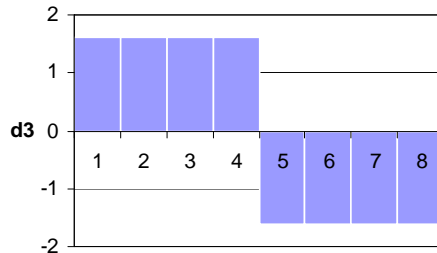
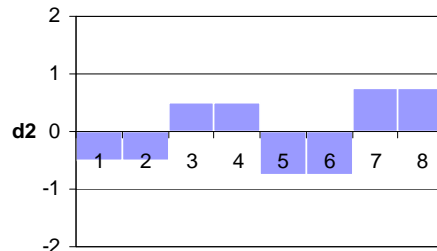
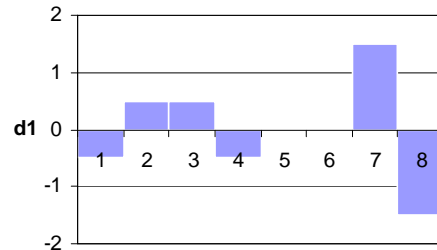
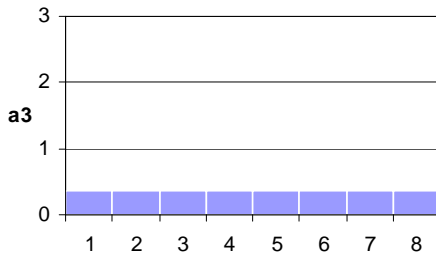
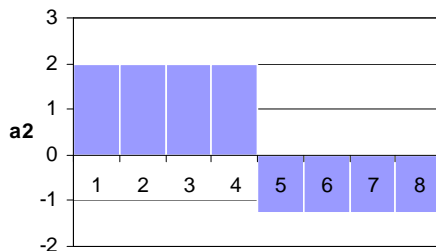
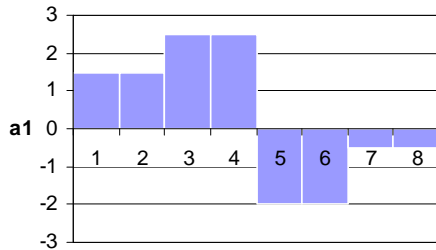
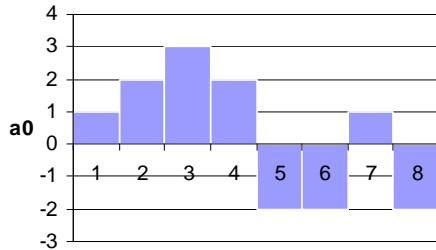
Werden jeweils 2 aufeinanderfolgende Werte gemittelt mit einer Schrittweite 2, so erhält man den Signalvektor $a_1 = \{1,5 \ 1,5 \ 2,5 \ 2,5 \ -2 \ -2 \ -0,5 \ -0,5\}$. Dieses (tiefpassgefilterte) Signal approximiert das Originalsignal mit verminderter Auflösung. Die Differenz zwischen dem Originalsignal a_0 und der Approximation a_1 ist das Detailsignal $d_1 = \{-0,5 \ 0,5 \ 0,5 \ -0,5 \ 0 \ 0 \ 1,5 \ -1,5\}$.

Dieser Algorithmus wird für das Signal a_1 wiederholt. Jeweils 4 aufeinanderfolgende Werte werden gemittelt mit der Schrittweite 4 und ergeben das Approximationssignal $a_2 = \{2 \ 2 \ 2 \ 2 \ -1,25 \ -1,25 \ -1,25 \ -1,25\}$. Dieses Signal bildet wiederum eine Approximation des Signals a_1 mit verminderter Auflösung. Die Differenz zwischen der höheren Approximation a_1 und der geringeren Approximation a_2 ist das Detailsignal $d_2 = \{-0,5 \ -0,5 \ 0,5 \ 0,5 \ -0,75 \ -0,75 \ 0,75 \ 0,75\}$.

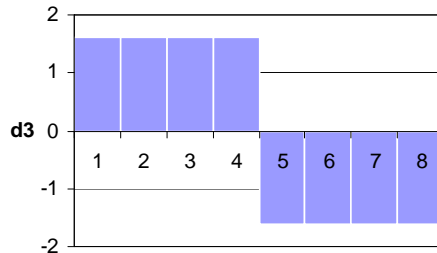
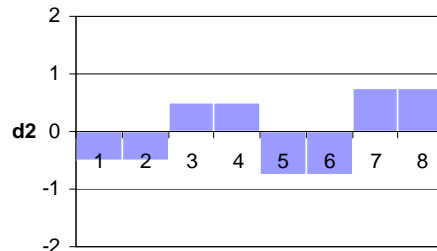
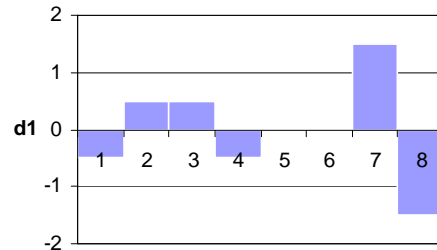
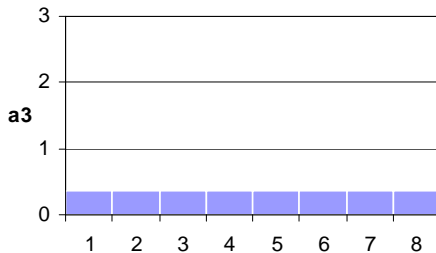
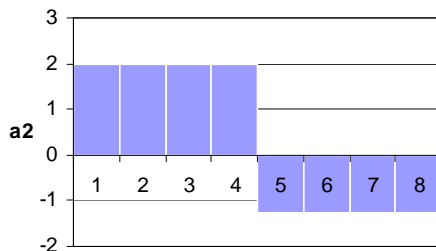
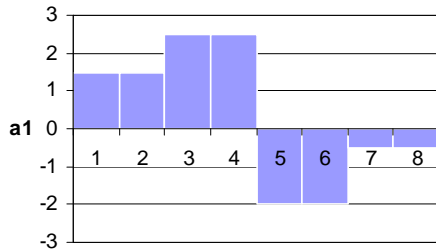
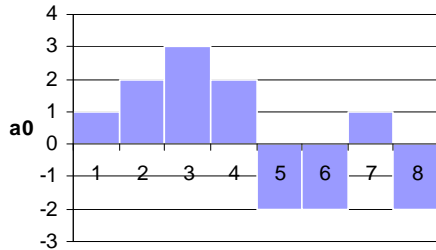


Im letzten Schritt führt die erneute Mittelung des Signalvektors a_2 auf den Signalvektor $a_3 = \{1,625 \ 1,625 \ 1,625 \ 1,625 \ 1,625 \ 1,625 \ 1,625 \ 1,625\}$, der den Mittelwert des Originalsignals darstellt und damit eine Approximation auf niedrigster Ebene bildet. Das Detailsignal berechnet sich zu $d_3 = \{1,625 \ 1,625 \ 1,625 \ 1,625 \ -1,625 \ -1,625 \ -1,625 \ -1,625\}$.

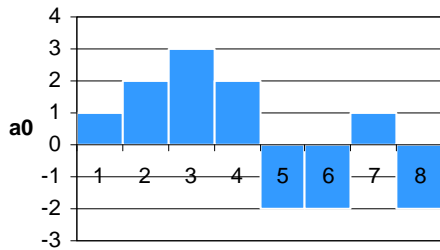




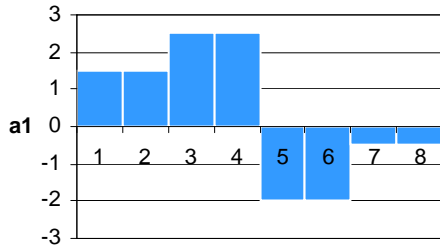
Das Originalsignal kann berechnet werden als Summe der niedrigsten Approximation und aller Detailsignale: $a_0 = a_3 + d_3 + d_2 + d_1$. Die Detailsignale werden Waveletfunktionen und Approximationssignale werden Skalierungsfunktionen genannt. Mit jeder Addition eines Teilsignals verbessert sich die Auflösung des Bildes, das Verfahren ist daher für progressive Bildübertragung bestens geeignet.



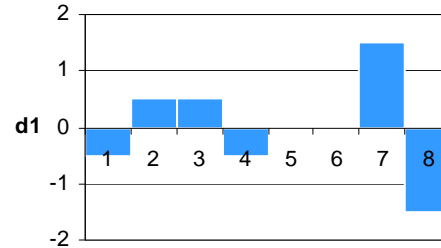
Eine nähere Analyse der Detailsignale lässt erkennen, dass sie aus gegeneinander verschobenen Funktionen bestehen. Zwei benachbarte Signalwerte haben den gleichen Betrag, aber unterschiedliches Vorzeichen. Auch für die Approximationssignale gelten ähnliche Zusammenhänge (gleiche Amplitude benachbarter Werte).



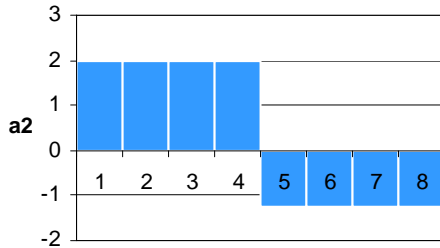
$$a_0 = \{1 \ 2 \ 3 \ 2 \ -2 \ -2 \ 1 \ -2\}$$



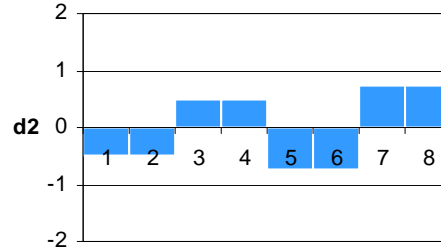
$$a_1 = \{1,5 \ 2,5 \ -2 \ -0,5\}$$



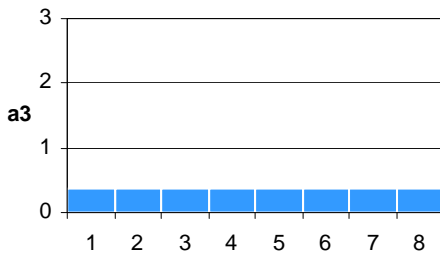
$$d_1 = \{-0,5 \ 0,5 \ 0 \ 1,5\}$$



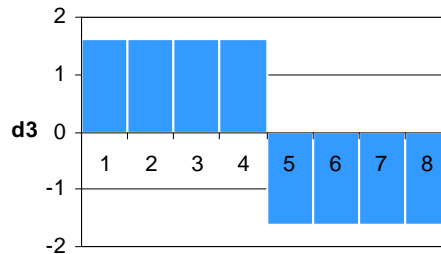
$$a_2 = \{2 \ -1,25\}$$



$$d_2 = \{-0,5 \ -0,75\}$$



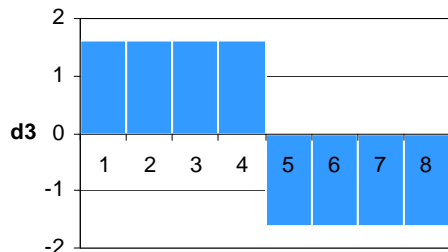
$$a_3 = \{0,375\}$$



$$d_3 = \{1,625\}$$

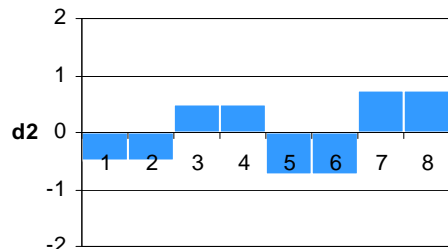
$$w = \{0,375 \ 1,625 \ -0,5 \ -0,75 \ -0,5 \ 0,5 \ 0 \ 1,5\}$$

0,375



0,625

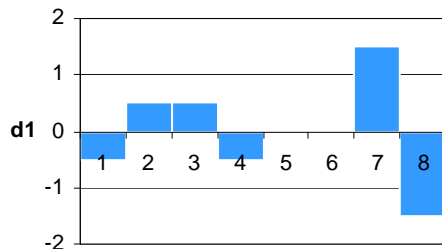
-0,5



-0,75

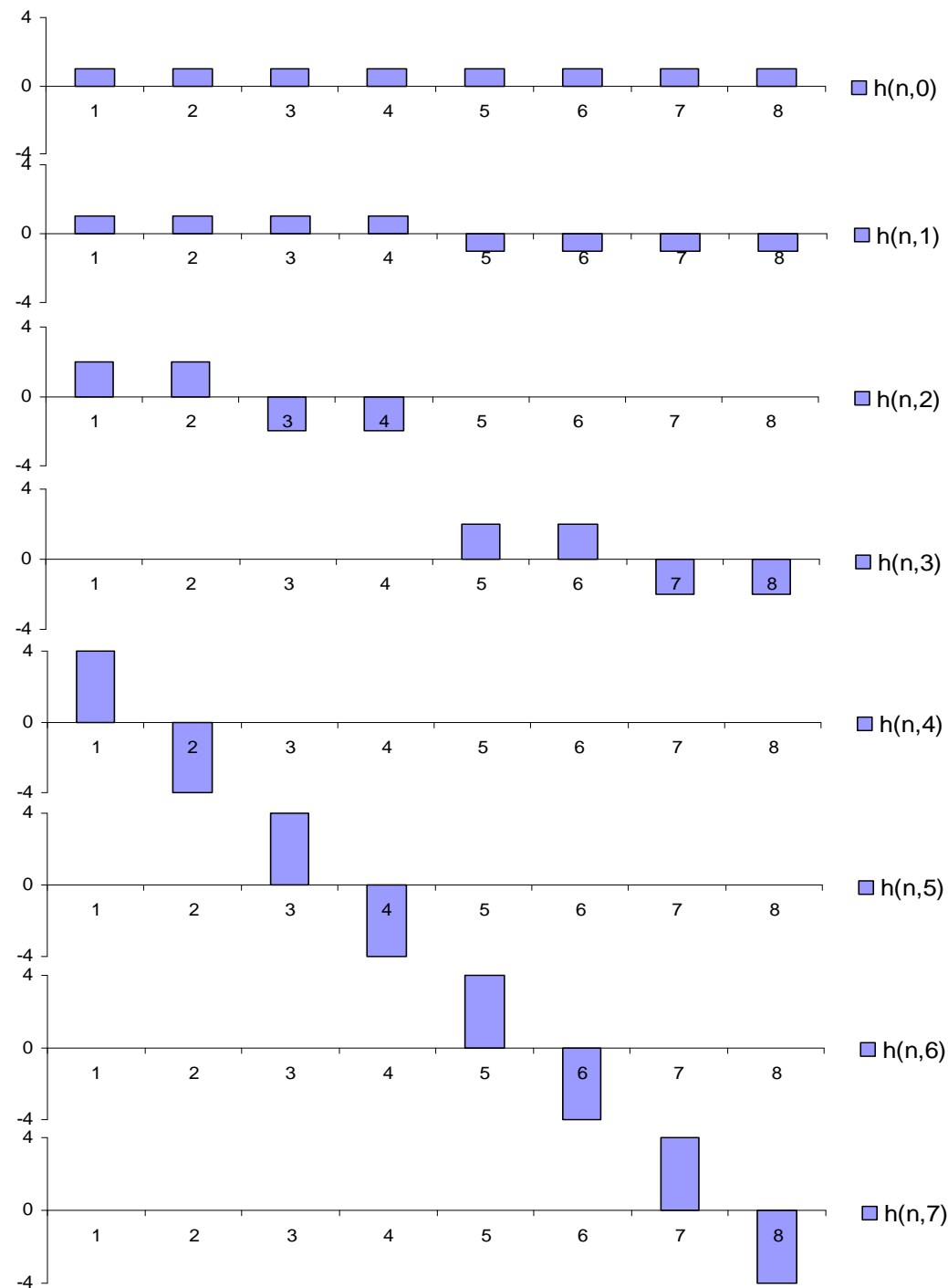
-0,5

0,5

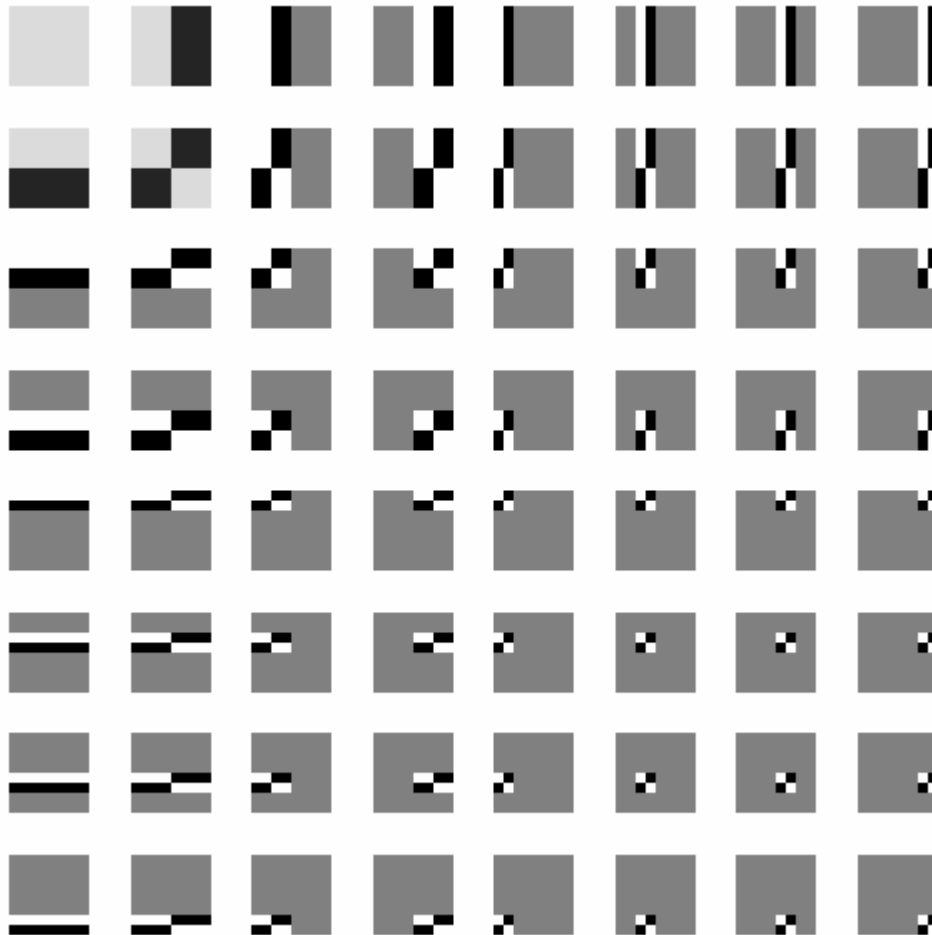


0

1,5



Basisbilder der DHT



Demo