

2. Fortsetzung zu Kap. 4: Bildoperationen / Lokale Operationen

Morphologische Operationen

Morphologie = Formenlehre

Idee:

über flexible Festlegung eines strukturierenden Elements (Maske) Einfluss auf zu extrahierende Formen nehmen

morphologische Operatoren sind *nichtlinear*

Spezialgebiet: "mathematische Morphologie" (Begr. v. G. Matheron & J. Serra in Frankreich, Analyse poröser Materialien)

- Kantenextraktion
- Texturanalyse
- Analyse von Größenverteilungen
- Skelettierung von Objekten
- gezielte Vergrößerungen von Bildern
- pixelbasierte Klassifikationsverfahren
- ...

Hilfsmittel:

$A \cup B$ Vereinigungsmenge

$A \cap B$ Schnittmenge

$\mathbf{C}A$ Komplement

IA identischer Mengenoperator: $IA = A$

für Teilmengen von \mathbb{R}^n :

$A \boxplus B = \{ x \mid x = a+b, a \in A, b \in B \}$ Mengen-Summe (Minkowski-Summe)

$A \boxminus B = \{ x \mid \forall b \in B: x+b \in A \}$ Mengen-Differenz

$B_x = B \boxplus \{x\}$ = um x verschobene Kopie der Menge B

"Strukturelement" B : Menge von Pixeln in der Gitterebene, mit Festlegung des Nullpunktes

o o o
o • o
o o o

Definition der beiden morphologischen Basis-Operatoren "Dilatation" und "Erosion"

(Achtung: Großes Definitions-Wirrwarr in der Lit.! – Hier weitgehend nach Soille 1998)

- Definitionen zunächst für *Binärbilder*.

Gegeben: ein strukturierendes Element B und eine Menge X.

Dilatation (auch: *Dilation*)

$D_B(X)$ = alle diejenigen Punkte, für die das strukturierende Element, dorthin verschoben, die Menge X trifft

$$= \{ x \mid B_x \cap X \neq \emptyset \}$$

gleichwertig:

$$= \{ x \mid \exists b \in B: x+b \in X \}$$

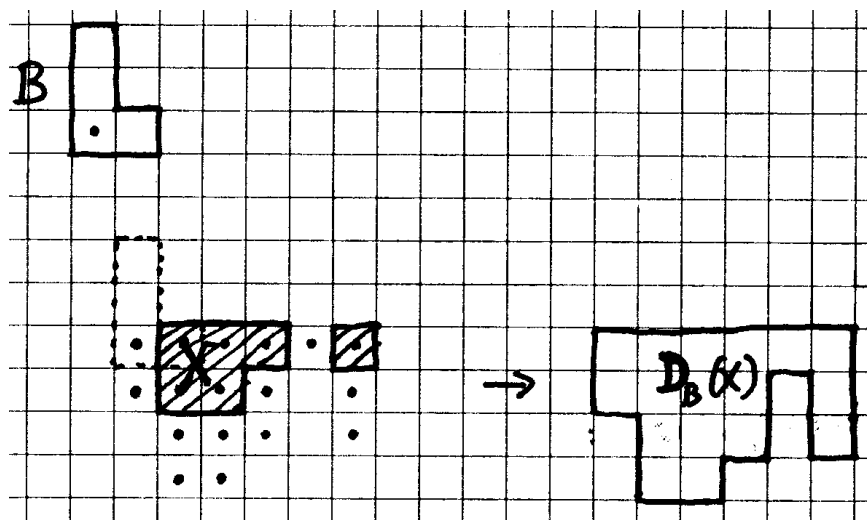
$$= X - b_1 \cup X - b_2 \cup \dots \quad (b_i \in B)$$

$$= \bigcup_{b \in B} X_{-b}$$

$$= \{ x \mid x = a - b, a \in X, b \in B \}$$

$$= X \boxplus (-B) \quad (-B: \text{die am Nullpunkt gespiegelte Menge } B)$$

Beispiel:



aus der schraffierten Menge X wird durch Dilatation mit dem Strukturelement B die rechts gezeigte Menge.

duale Operation:

Erosion

$E_B(X)$ = alle diejenigen Punkte, für die das strukturierende Element, dorthin verschoben, ganz in X liegt
 $= \{ x \mid B_x \subseteq X \}$

gleichwertig:

$$\begin{aligned} &= \{ x \mid \forall b \in B: x+b \in X \} \\ &= X - b_1 \cap X - b_2 \cap \dots \quad (b_i \in B) \\ &= \bigcap_{b \in B} X_{-b} \\ &= X \boxminus B \end{aligned}$$

Verallgemeinerung der Basisfunktionen auf *Grauwertbilder*:

Sei f_{-b} die um $(-b)$ verschobene Intensitätsfunktion f , also $f_{-b}(x) = f(x+b)$.

Ersetze in obiger Def. die Vereinigungsmenge durch das Maximum und die Schnittmenge durch das Minimum.

$$\begin{aligned} D_B(f)(x) &= \max \{ f_{-b}(x) \mid b \in B \} = \max \{ f(x+b) \mid b \in B \} \\ E_B(f)(x) &= \min \{ f_{-b}(x) \mid b \in B \} = \min \{ f(x+b) \mid b \in B \} \end{aligned}$$

- für die charakteristische Funktion f_X einer Menge X läuft dies auf die alte Def. heraus
- **D** und **E** erweisen sich damit als Verallgemeinerung des Max- bzw. Min-Operators
- anschauliche Interpretation: Dilatation bzw. Erosion der (3D-) Menge *unter dem Graphen* von f .

weitere Verallgemeinerung: auch statt B kann eine Intensitätsfunktion genommen werden ("B nicht flach")

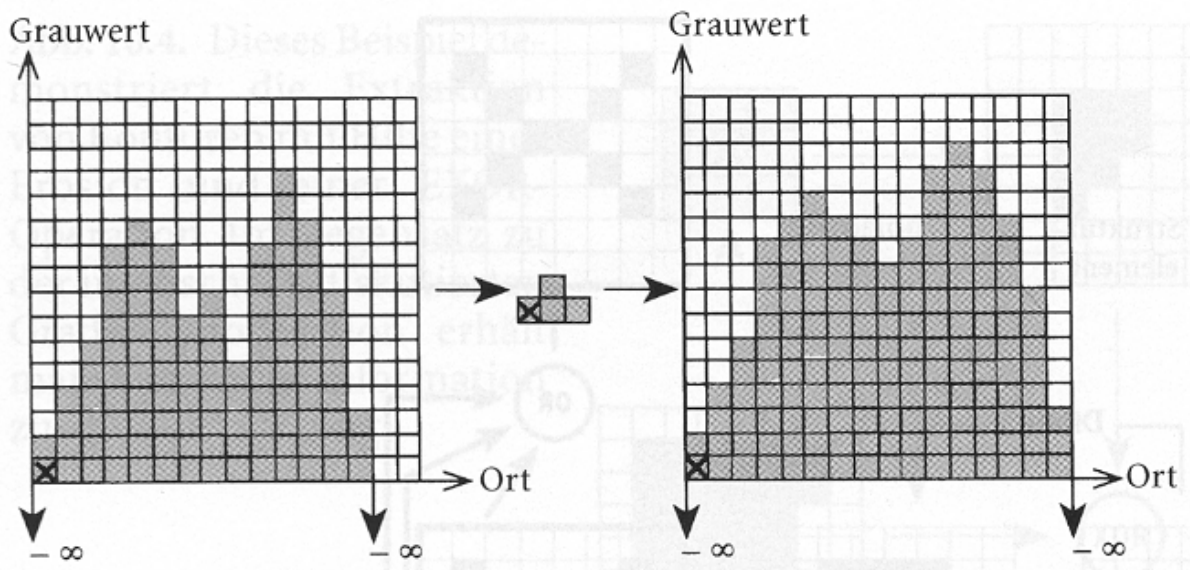
Def. für Funktionen f und B ($D(B)$ Definitionsbereich von B):

$$D_B(f)(x) = \max \{ f(x+b) + B(b) \mid b \in D(B) \}$$

$$E_B(f)(x) = \min \{ f(x+b) - B(b) \mid b \in D(B) \}$$

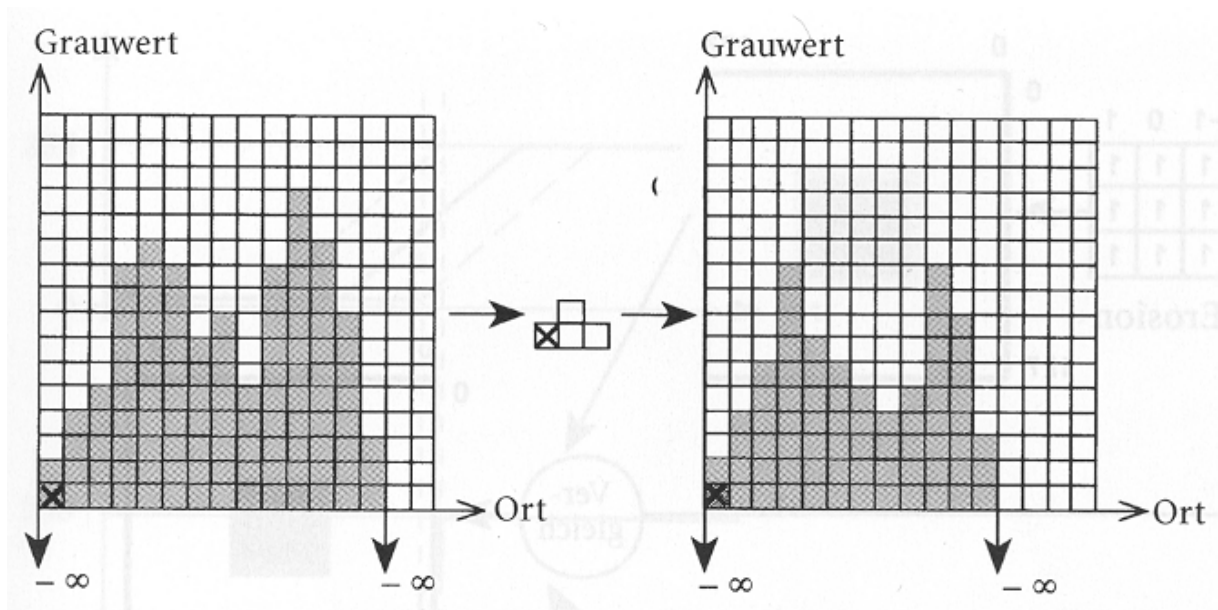
- Beachte: den Spezialfall (B Menge; man sagt: "B flach") erhält man hieraus durch eine Funktion B , die auf ihrem Definitionsbereich 0 ist.
- Anschauliche Interpretation auch hier: Dilatation bzw. Erosion der Menge unter dem Graphen von f , jetzt durch ein dreidimensionales Strukturelement.

Beispiel für Grauwert-Dilatation mit nichtebenen Strukturelement (vereinfacht auf nur eine 1D-Ortskoordinate):



(aus Bäsmann & Kreyss 1998)

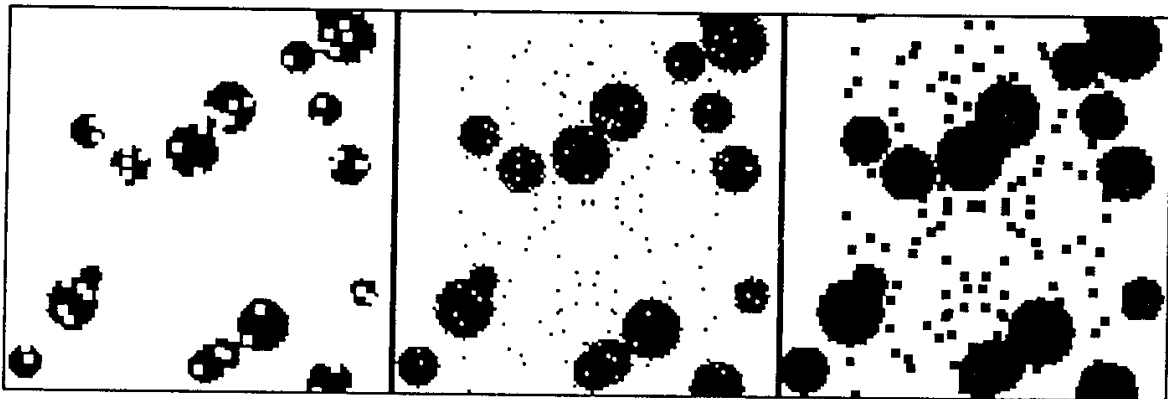
Analog für die Erosion:



Eigenschaften von Dilatation und Erosion:

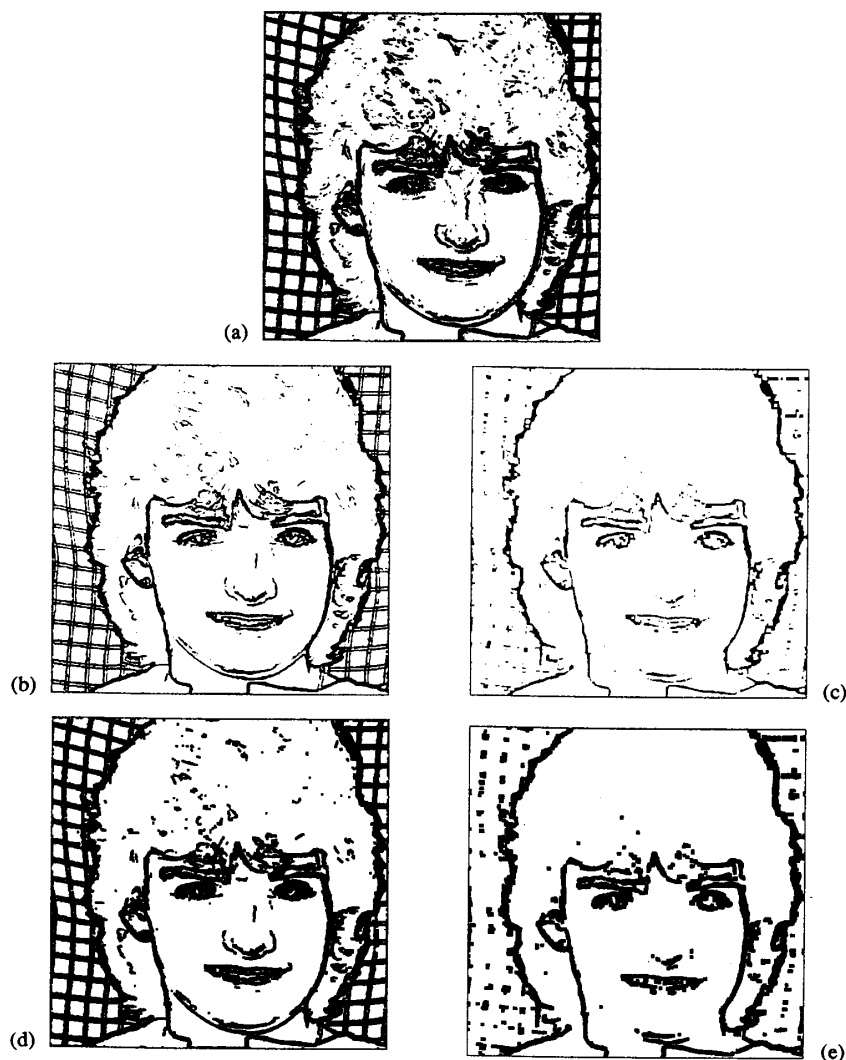
- Dualität:
 $\mathbf{E} = \mathbf{C} \mathbf{D} \mathbf{C}$
 $\mathbf{D} = \mathbf{C} \mathbf{E} \mathbf{C}$
- Verschiebungsinvarianz:
 $(\mathbf{D}(f_c)) = (\mathbf{D}(f))_c$
 $(\mathbf{E}(f_c)) = (\mathbf{E}(f))_c$
- Monotonie:
 $f \leq g \Rightarrow \mathbf{D}(f) \leq \mathbf{D}(g), \mathbf{E}(f) \leq \mathbf{E}(g)$
- Distributivgesetze bzgl. punktwisem Minimum und Maximum von Funktionen
- Zusammensetzung von Dilatationen und Erosionen:
 $\mathbf{D}_B \mathbf{D}_A = \mathbf{D}(\mathbf{D}_{-B} A)$
 $\mathbf{E}_B \mathbf{E}_A = \mathbf{E}(\mathbf{D}_{-B} A)$

Wirkung von Erosion und Dilatation bei Binärbildern
(Erosion **EB**, Originalbild **B**, Dilatation **DB**):



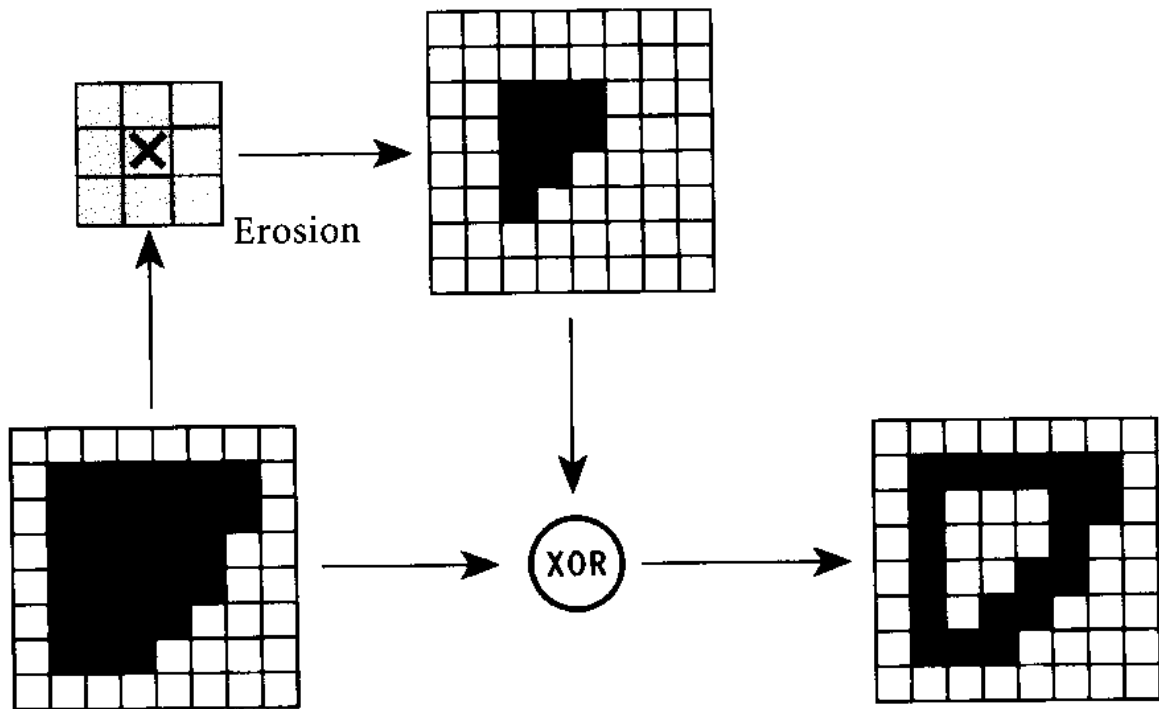
(aus Voss & Süße 1991)

Einsatz zur Manipulation von Kantenbildern aus der klassischen Filterung:

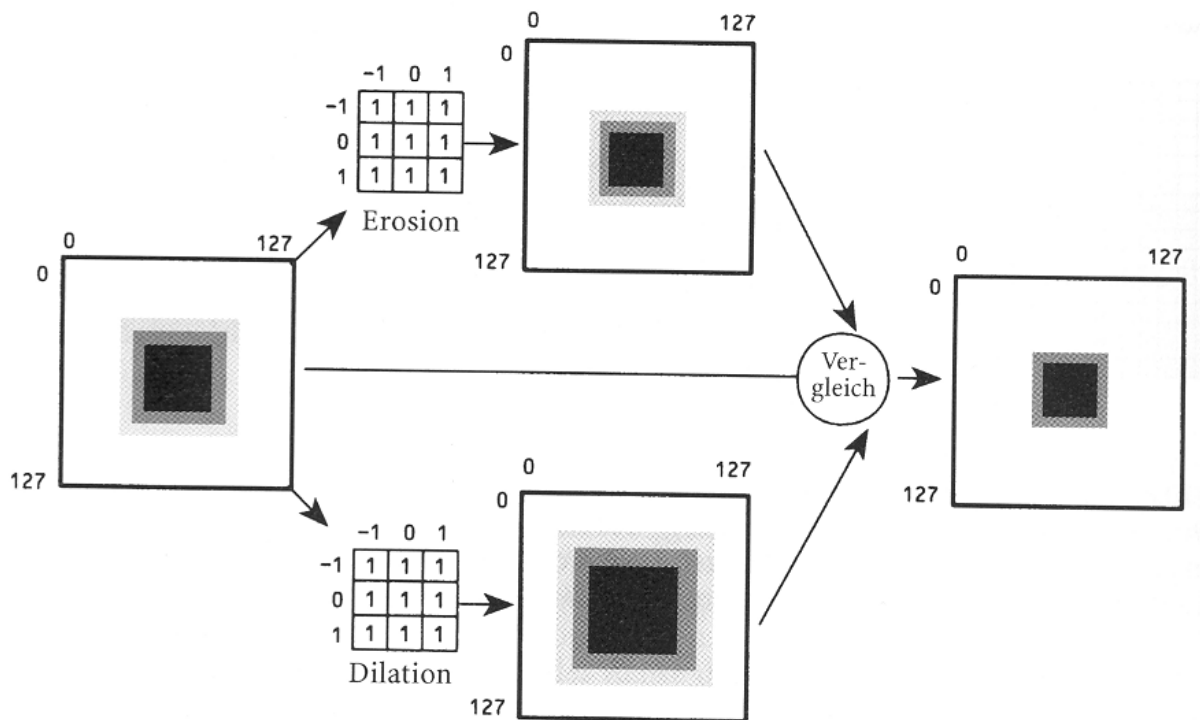


(b), (c) Erosionen von (a); (d), (e) Dilatationen (aus Haberäcker 1991)

Extraktion von Konturen:



Adaptive Kontrastverbesserung:



Zusammengesetzte morphologische Operationen:

Öffnung (Opening): $\mathbf{O}_B = \mathbf{D}_{-B}\mathbf{E}_B$
(für symmetrisches B ist $\mathbf{O} = \mathbf{DE}$)

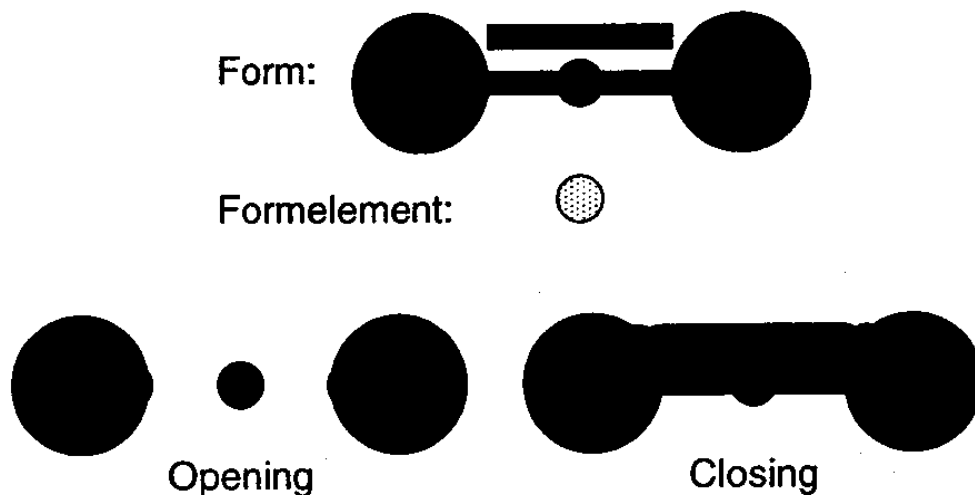
bewirkt Aufbrechen von Regionen an Stellen, wo nur schwacher Zusammenhang bestand

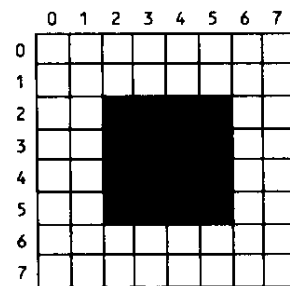
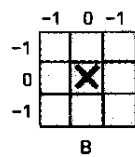
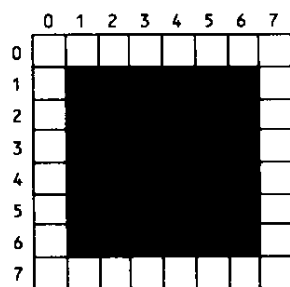
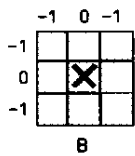
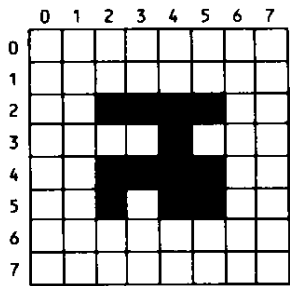
Schließung (Closing): $\mathbf{S}_B = \mathbf{E}_{-B}\mathbf{D}_B$
(für symmetrisches B ist $\mathbf{S} = \mathbf{ED}$)

schließt kleine Zwischenräume

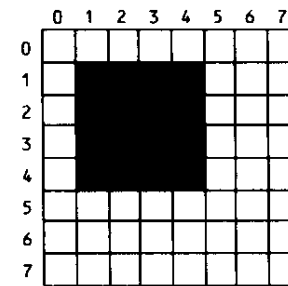
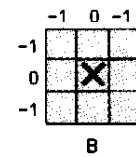
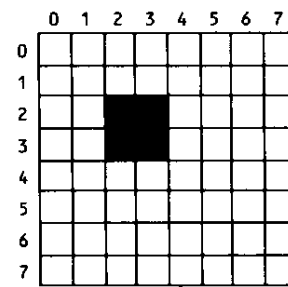
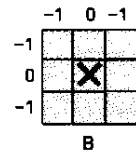
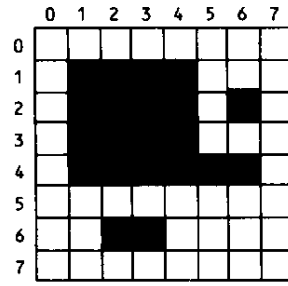
Eigenschaften:

- Idempotenz
 $\mathbf{OO} = \mathbf{O}, \mathbf{SS} = \mathbf{S}$
- Dualität
 $\mathbf{S} = \mathbf{COC}, \mathbf{O} = \mathbf{CSC}$
- Schließung extensiv, Öffnung anti-extensiv:
 $\mathbf{O} \leq \mathbf{I} \leq \mathbf{S}$
- Monotonie:
 $f \leq g \Rightarrow \mathbf{O}(f) \leq \mathbf{O}(g), \mathbf{S}(f) \leq \mathbf{S}(g)$





a



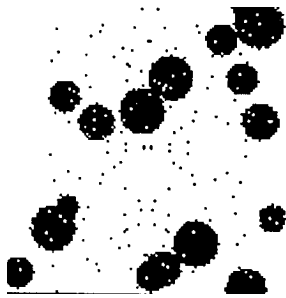
b

- a: Schließung
- b: Öffnung

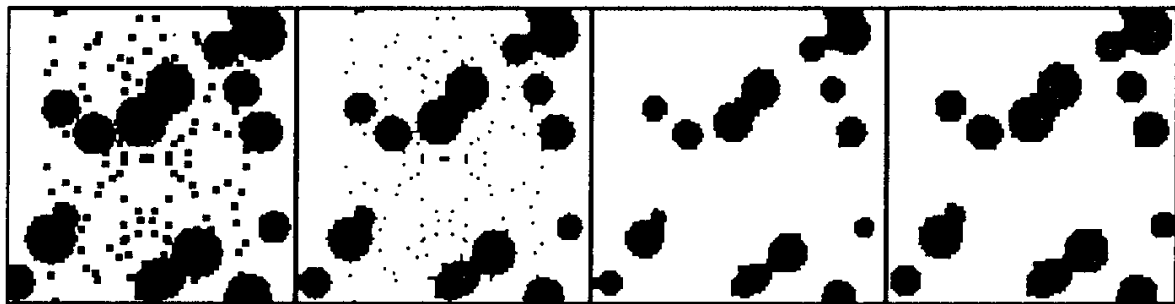
Öffnung: Abtragen "ausgefranster" Bereichsränder, Eliminieren kleiner Partikel

Schließung: schließt Lücken zwischen "Fransen"

Anwendung zur Rauschunterdrückung (erst schließen, dann öffnen):



B



DB

SB=EEDB

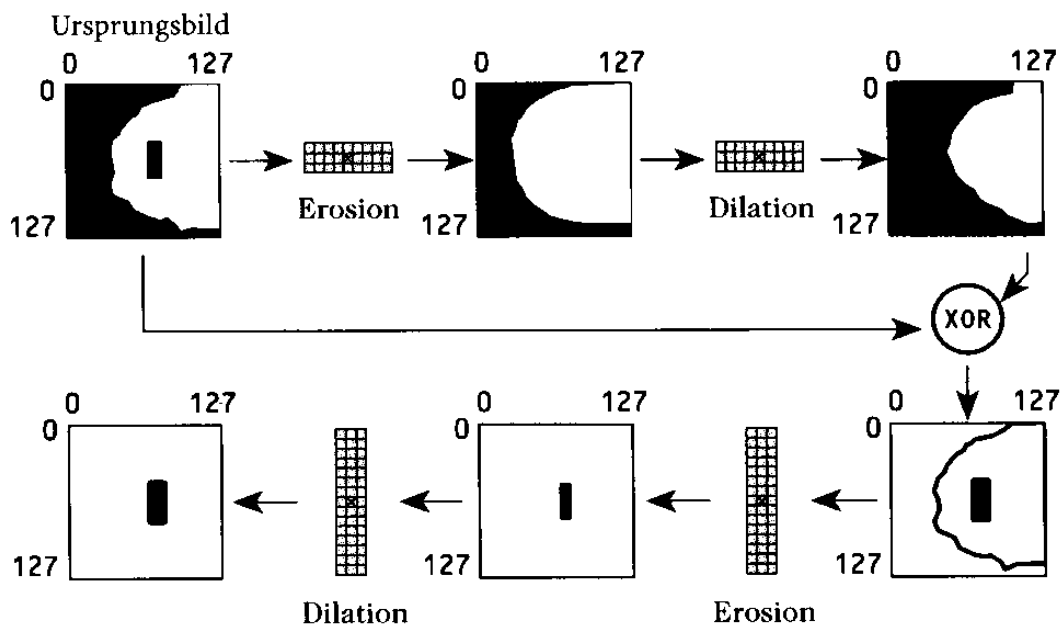
EEDB

OSB=DEEDB

(aus Voss & Süße 1991)

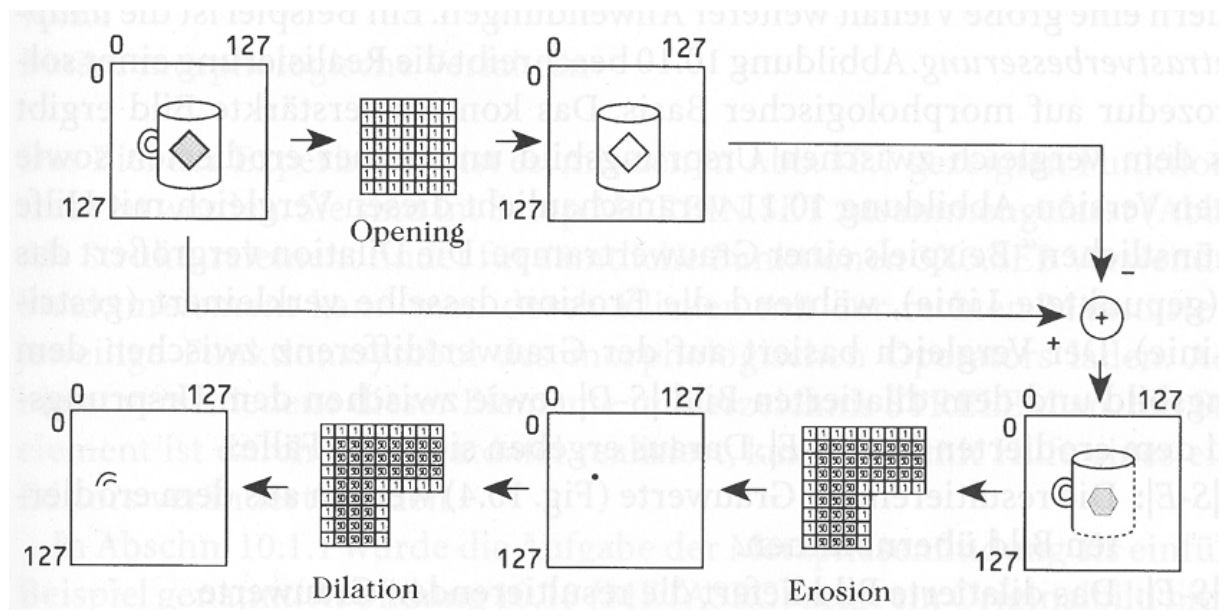
Einsatz spezieller strukturierender Elemente zur *Detektion von Binärbildanteilen mit bekannter Form*

Beispiel: Rechteck in der Mitte des Originalbildes:

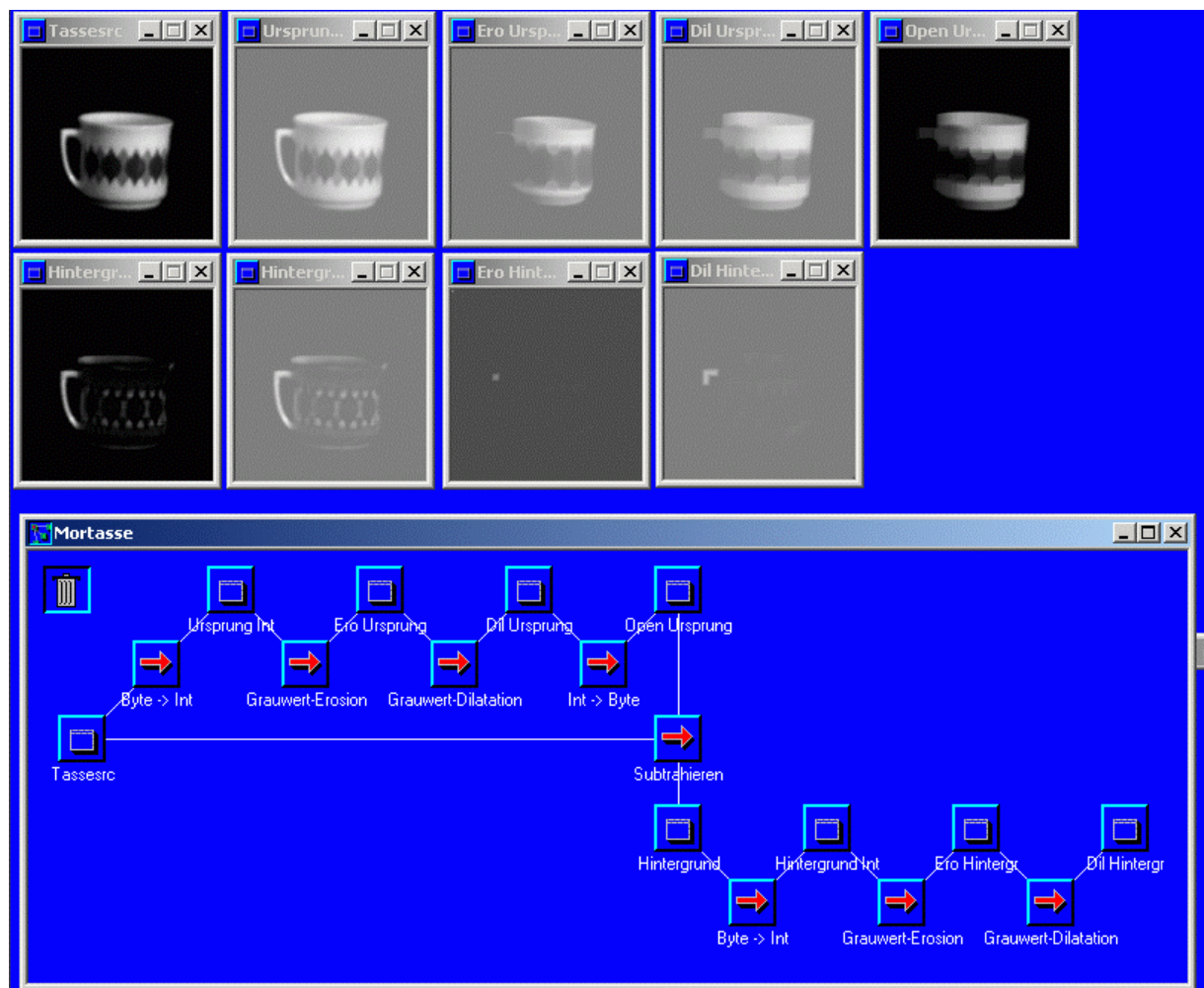


Öffnung zur Hintergrundschätzung, XOR liefert interessierende Bereiche, zweites Opening filtert gewünschten Bereich heraus (aus Bässmann & Kreyss 1998)

Übertragung der Vorgehensweise auf Grauwertbilder:



Beispiel mit AdOculos (aus Bässmann & Kreys 1998, S. 263ff.):



Dilatationsresiduum:

$$\mathbf{D}rf = \mathbf{D}f - f$$

Erosionsresiduum:

$$\mathbf{E}rf = f - \mathbf{E}f$$

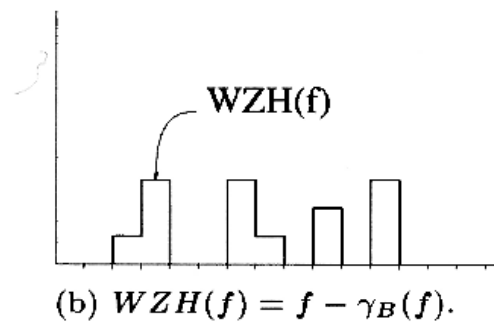
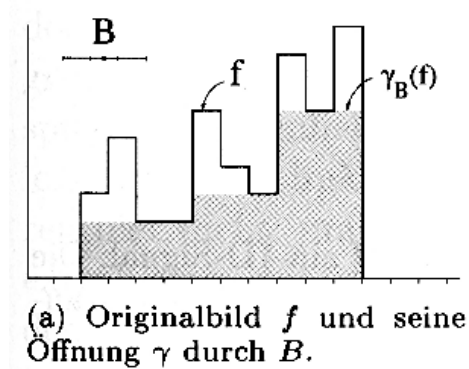
Zylinderhut-Transformationen:

Weißer Zylinderhut:

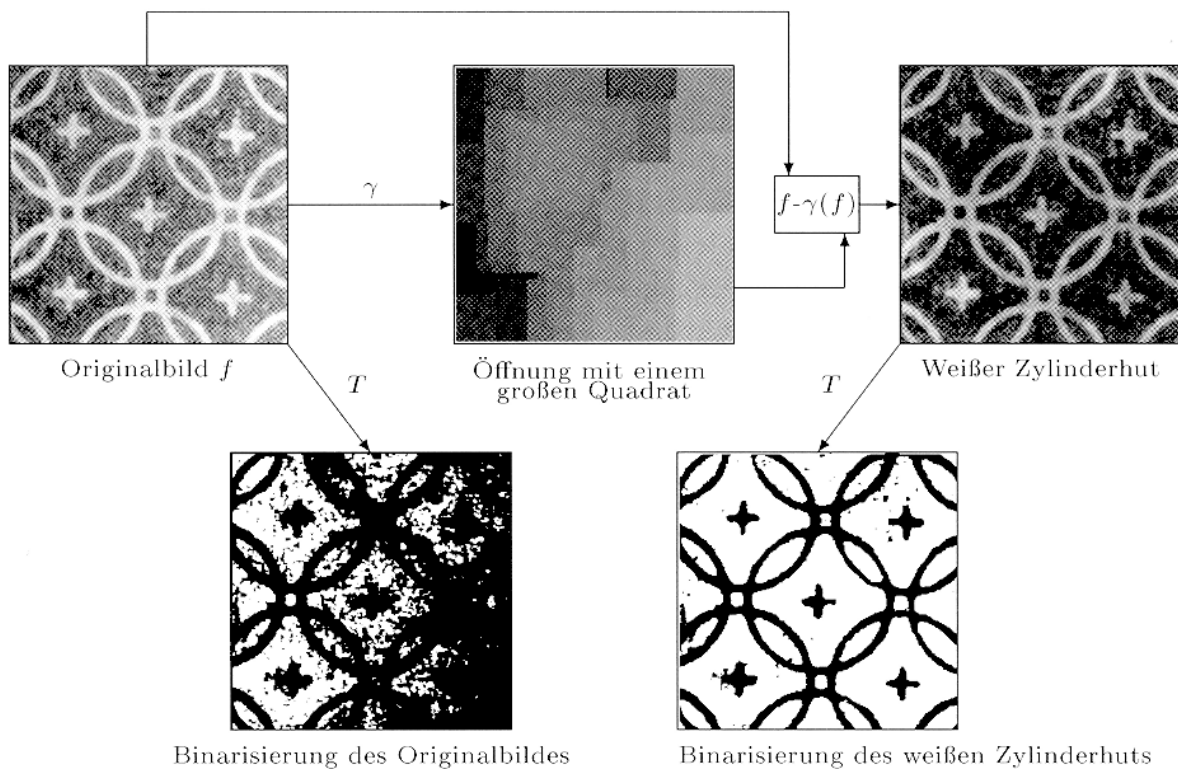
$$\mathbf{WZH}(f) = f - \mathbf{O}f$$

Schwarzer Zylinderhut:

$$\mathbf{SZH}(f) = \mathbf{S}f - f$$



Einsatz des Zylinderhuts zur Shadingkorrektur:



(aus Soille 1998)

Zusammengesetztes strukturierendes Element $B = (B_1, B_2)$:

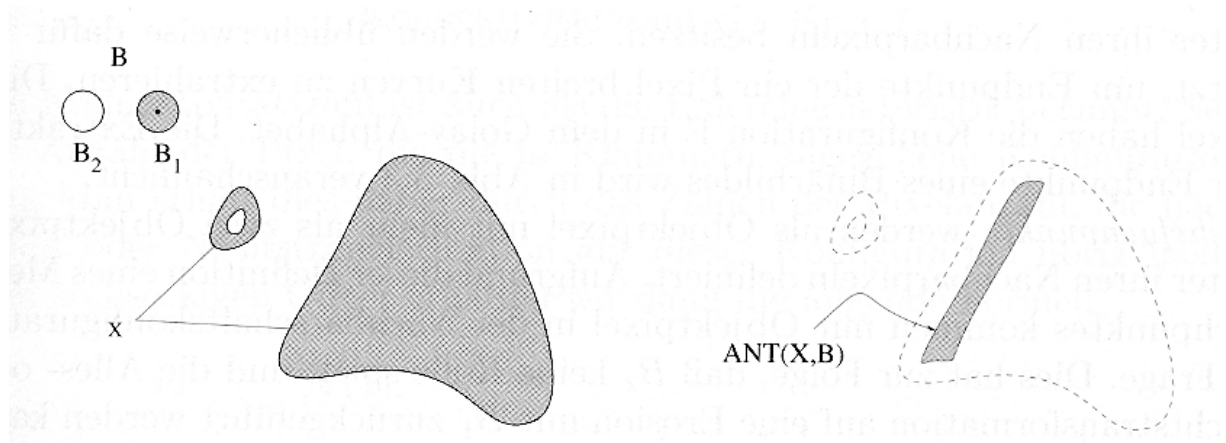
- B_1 und B_2 sind disjunkt
- besitzen denselben Bezugspunkt (Nullpunkt)

Alles-oder-Nichts-Transformation

Idee: Extraktion von Pixeln, die eine durch (B_1, B_2) gegebene Nachbarschaftskonfiguration besitzen

"Passt die erste Menge in die Bildobjekte hinein, während gleichzeitig die zweite Menge die Bildobjekte nicht berührt (d.h. in den Hintergrund hineinpasst)?"

$$\begin{aligned} \mathbf{ANT}_B(X) &= \{ x \mid (B_1)_x \subseteq X \text{ und } (B_2)_x \subseteq \mathbf{CX} \} \\ &= \mathbf{E}_{B_1} X \cap \mathbf{E}_{B_2} \mathbf{CX} \end{aligned}$$



Verdünnung eines Binärbildes bzgl. eines zusammengesetzten strukturierenden Elementes B :

$$X \circ B = X \setminus \mathbf{ANT}_B(X) \quad (\text{Mengendifferenz})$$

(auf Grauwertbilder verallgemeinerbar, s. Soille 1998, S. 140ff.)

Verdickung:

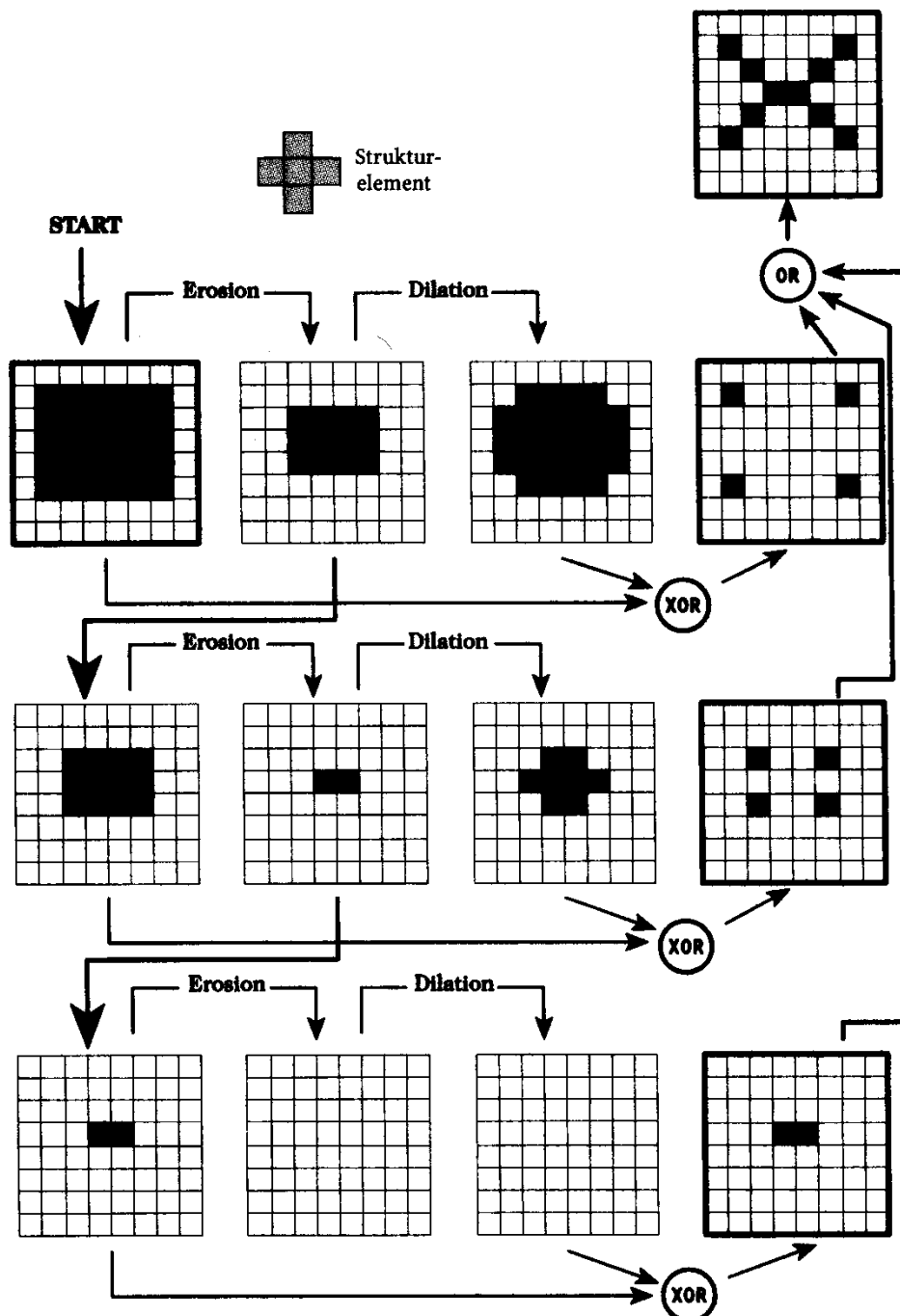
analog mit Vereinigungsmenge statt Mengendifferenz

Skelettierung

Ziele:

- pixelbreite Konturen
- aus Flächen Formbeschreibungen generieren
- Strichzeichnungen durch Scanner oft "zu dick" - Bildverbesserung

Beispiel:



(aus Bäsmann & Kreyss 1998)

Prinzip:

Erosion vom Rand bis zur Mitte des Objekts (Skelettlinie)

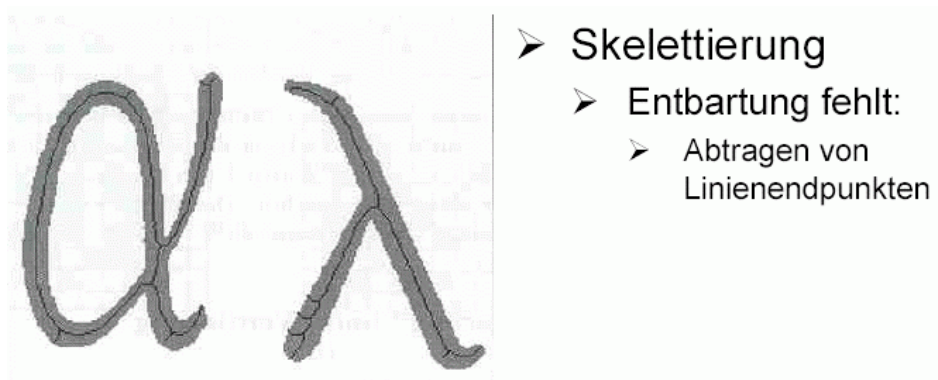
bessere Verfahren als das oben gezeigte:

Erosion + bestimmte Bedingungen

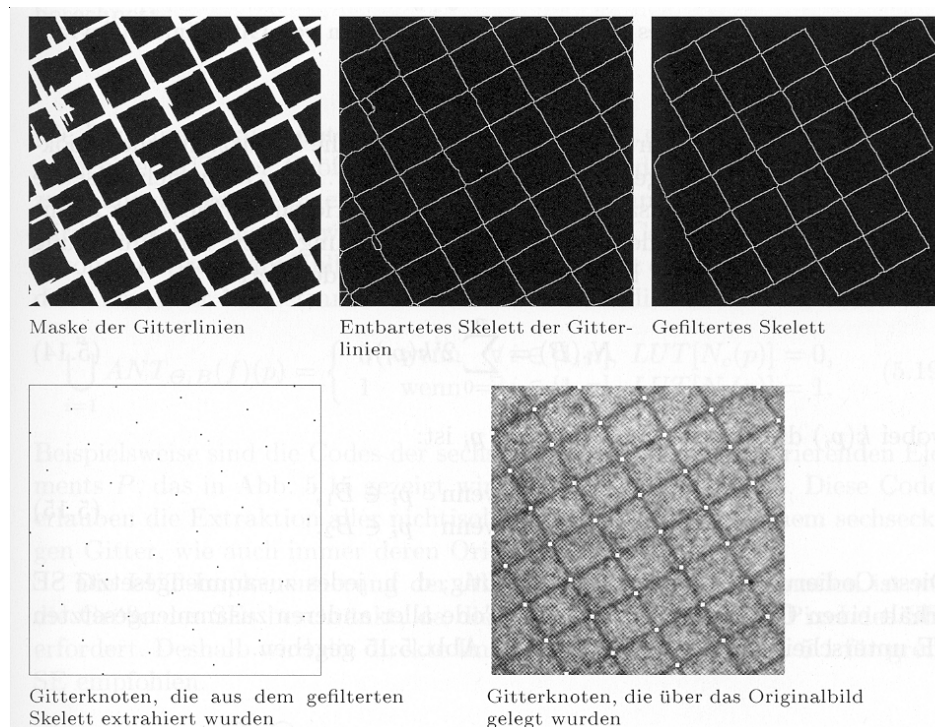
- Skelettlinie darf nur 1 Pixel breit sein
- muss in der Mitte des Objekts verlaufen
- darf sich bei weiteren Schritten nicht weiter verändern

⇒ Verwendung bestimmter strukturierender Elemente, die nur dann abtragen, wenn Punkte *keine* Skelettpunkte sind

Verwendung von **ANT** und Verdünnung



Anwendung: z.B. Extraktion der Schnittpunkte von Gitterlinien



(aus Soille 1998)

Vorteile der morphologischen Bildoperationen:

- komplexe Bildverarbeitungsaufgaben können auf einfache Basisoperationen auf Pixelebene zurückgeführt werden
- diese Grundoperationen basieren letztlich auf Boolescher Algebra
- die Verfahren sind hochgradig parallelisierbar

⇒ Hardware-freundlich

4.3. Globale Operationen

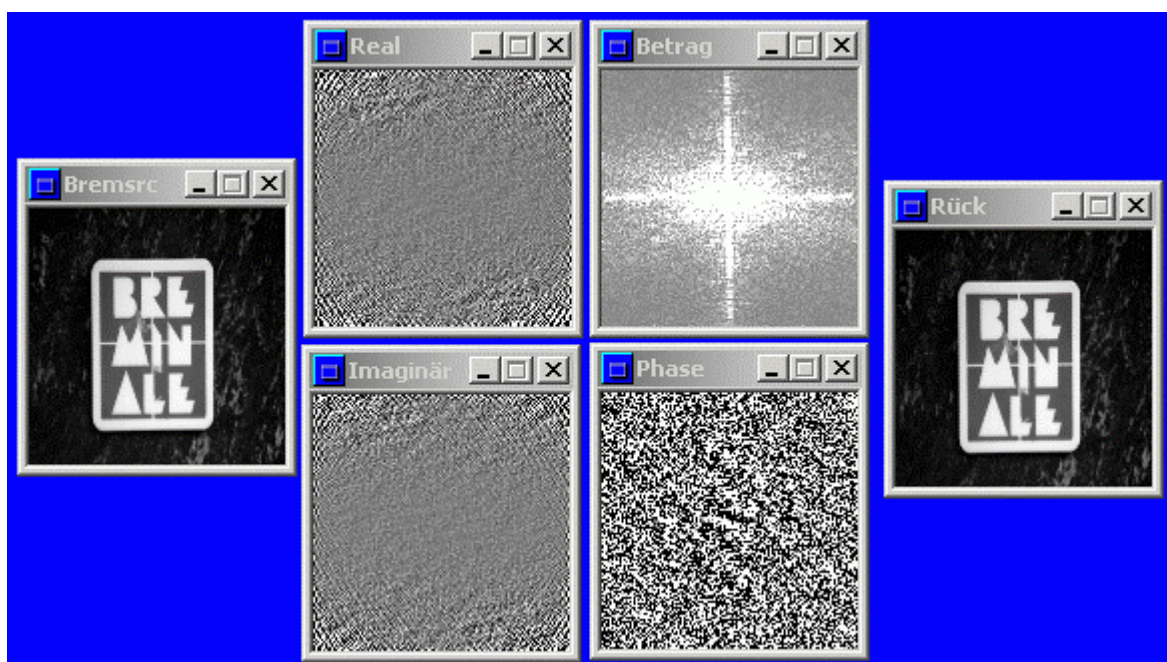
Ergebnis einer globalen Op. im Punkt (x, y) : abhängig vom gesamten Eingabebild

⇒ hoher Rechenaufwand (bis zu $O(n^4)$!)

- globale Op. spielen wichtige Rolle in techn. Signalverarbeitung
- weniger zentral bei Bildverstehen
- oft unterschiedliche Verarbeitung in verschiedenen Bildteilen nötig ⇒ lokale Op. und Punktoperationen

Fouriertransformation

- Zerlegung in Sinus- und Kosinusanteile
- knappere Darstellung mit der komplexen e-Funktion
- kontinuierliche FT: siehe Kap. 2 (Digitalisierung, Abtasttheorem)
- diskrete Fouriertransformation (DFT): siehe Übung
- mathematisch: Basiswechsel im Funktionenraum
- Darstellung der (komplexwertigen) Koeffizienten oft in Polarkoordinaten: Betrag (Fourierspektrum), Phase (Winkel).



Implementation:

- "naive" Berechnung der DFT nach der Transformationsformel: $O(n^2)$ Rechenoperationen
- "Schnelle Fouriertransformation" (*Fast Fourier Transformation*, FFT):
 - Idee: geschickte Faktorisierung der Transformationsmatrix in dünn besetzte Matrizen
 - ⇒ Aufwand $O(n \log n)$

Anwendung:

- Hochpass-, Tiefpass-, Bandpass-Filterung (bzw. alle Filterungen, die einfacher im Frequenzraum durchzuführen sind)
- Bildkompression

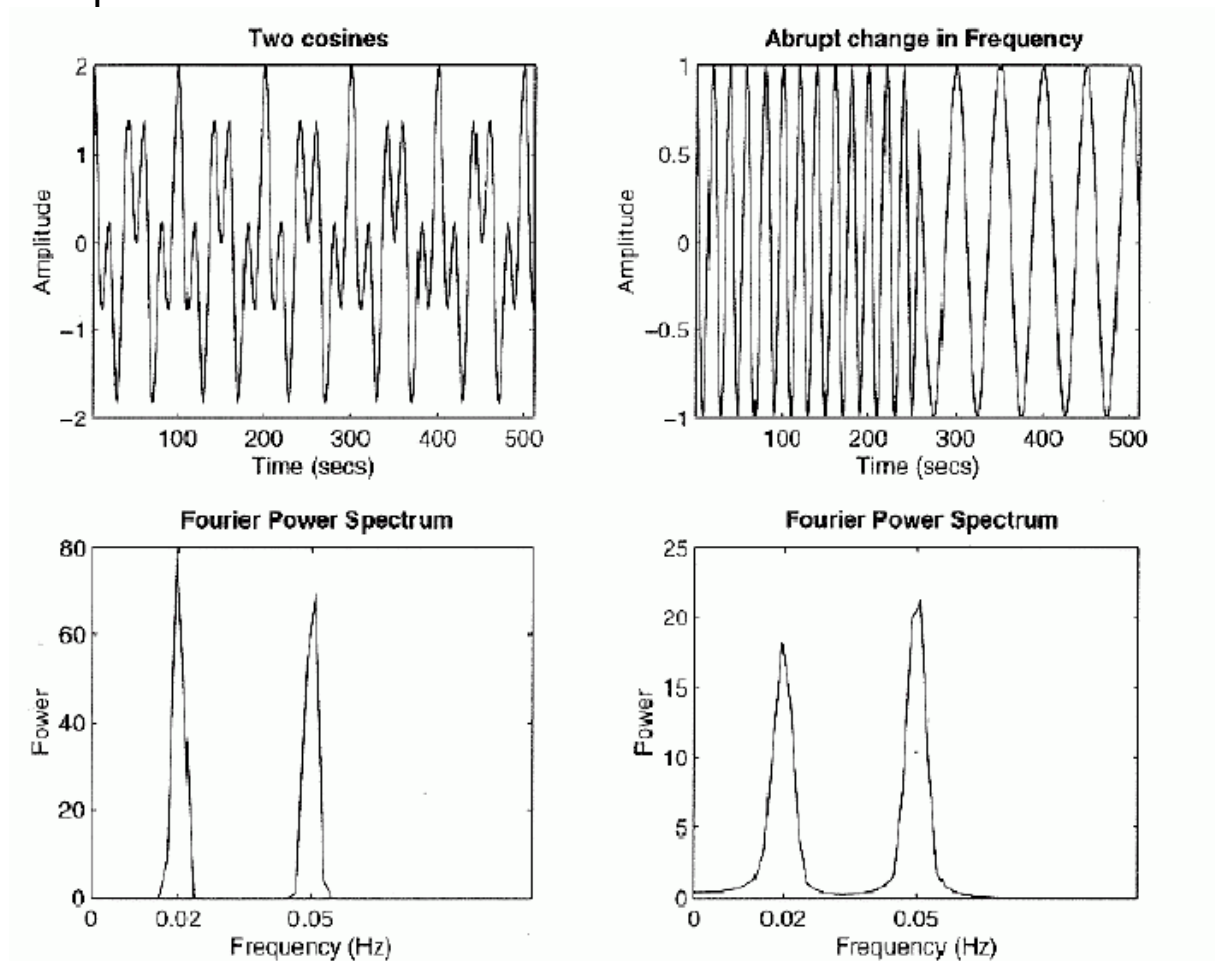
Variante: *Diskrete Kosinus-Transformation* (DCT). Vermeidung des Rechnens im Komplexen durch punktsymmetrische Fortsetzung des Bildes (⇒ Verschwinden der Sinus-Terme), dafür aber größere Matrizen involviert.

- Anwendung in der (klassischen) JPEG-Kompression; siehe Vorlesung Computergrafik, Kap. 4, http://www-gs.informatik.tu-cottbus.de/~wwwgs/cg_v04.htm

Nachteile der Fourier-Transformation:

- Berechnungsaufwand mit $O(n \log n)$ immer noch hoch
- Basisfunktionen haben unendlichen Träger im Ortsraum
⇒ Operationen im Frequenzraum können nicht auf Gebiete im Ortsraum beschränkt werden
- Information über räumliches Auftreten der einzelnen Frequenzen einer Fourier-transformierten Funktion ist nicht direkt erkennbar, ist "versteckt" in den Phasen: *Verlust der Ortsinformation*

Beispiel:



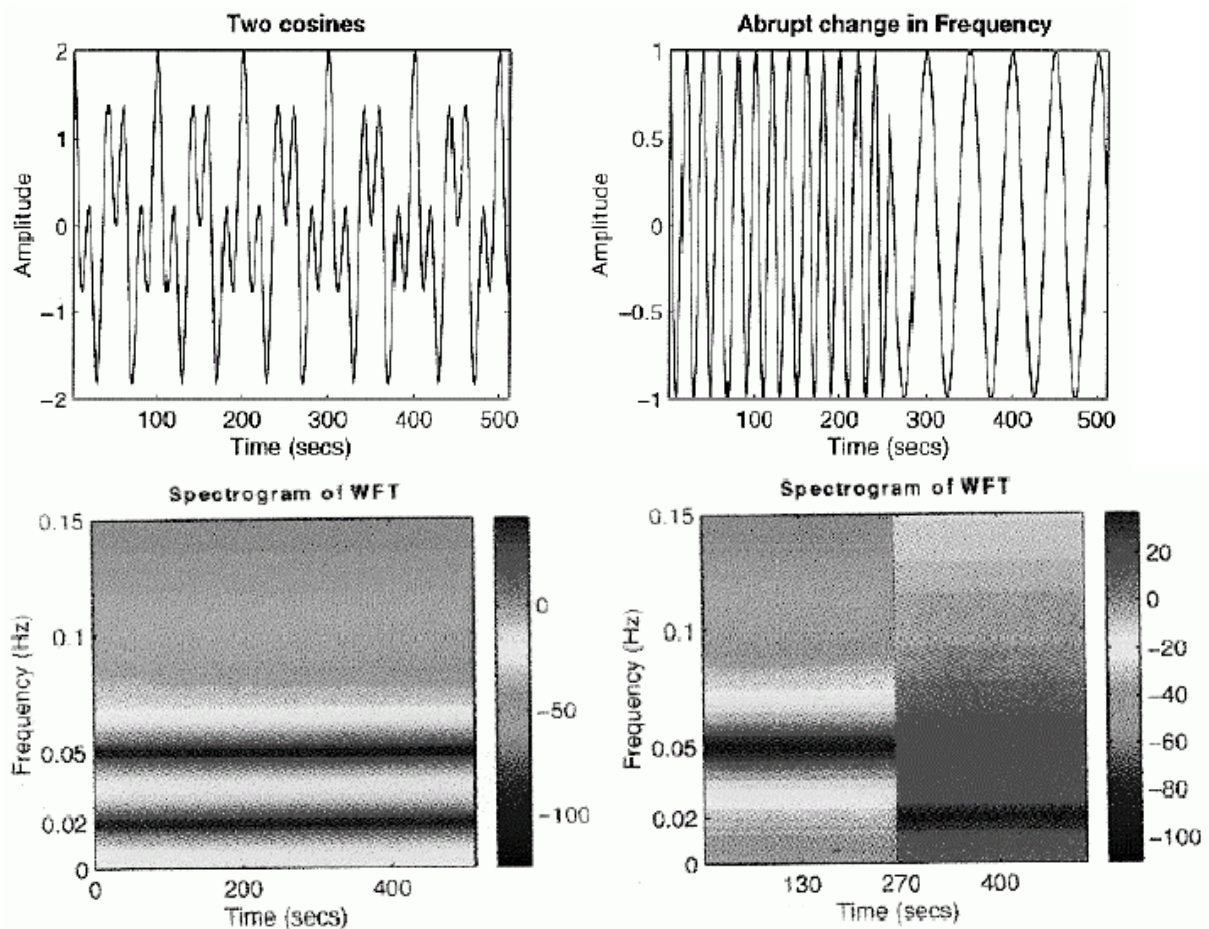
links: Überlagerung zweier Kosinusfunktionen (stationär),
rechts: Änderung der Frequenz;
beide Kurvenverläufe führen auf nahezu identische Fourierspektren (aus Esser 2001)

Ziel: bessere Lokalisierung von Frequenzbändern
⇒ Mischung von Raum- und Frequenzdarstellung

Erster Ansatz zur Abhilfe:

gefensterte Fourier-Transformation (Windowed Fourier-Transform: WFT), auch: *Short-Term Fourier Transformation* (STFT), *Short-Window Fourier Transformation* (SWFT)

- Funktion wird in gleichgroße, nicht überlappende Bereiche (Fenster) aufgeteilt und auf diese die DFT oder DCT angewandt (siehe JPEG-Kompression)



⇒ die beiden obigen Beispiele werden jetzt (bei geeignet gewähltem Fenster) unterscheidbar

allgemein: Lokalisierung durch Multiplikation der zu transformierenden Funktion mit einer Fensterfunktion g

$$\text{STFT}[m, k] = \sum_{n=0}^{N-1} f[n]g[n - m]e^{-i\frac{2\pi nk}{N}}$$

neben Rechteckfunktionen wird für g oft verwendet: Gauß-Fensterfunktion (⇒ Gabor-Transformation):

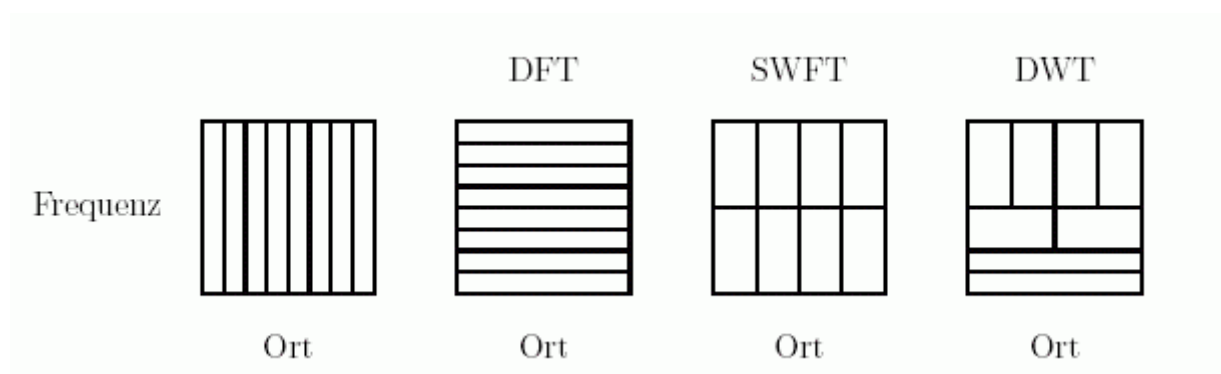
$$g_{\sigma}[n] = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{n^2}{2\sigma^2}}$$

Nachteil der STFT:

feste Fenstergröße \Rightarrow kann nicht an die Bildstruktur angepasst werden!

- schmales Analysefenster: gute Ortsauflösung, aber unempfindlich gegenüber niedrigen Bildfrequenzen (schlechte Frequenzauflösung)
- breites Analysefenster: schlechte Ortsauflösung
- dieser Tradeoff kann letztlich nicht beseitigt werden (Analogie zur Heisenbergschen Unschärferelation aus der Quantenmechanik)
- aber man kann eine optimale, variable Fenstergröße anstreben (angepasst ans Bild)

\Rightarrow Idee der *diskreten Wavelet-Transformation* (DWT)



ganz links: Beschreibung einer Funktion im Ortsbereich (z.B. klassische Bildmatrix), Mitte links: Beschreibung im Frequenzbereich mittels Fouriertransformation (DFT), alle Ortsinformationen sind in Frequenzinformationen umgewandelt; Mitte rechts: Darstellung mittels Short-Wavelet FT mit fester Fenstergröße, ganz rechts: flexible Orts- und Frequenzdarstellung mittels DWT. (aus Schütze 2002)

Bei der Wavelet-Transformation wird das Analysefenster durch dehn- und stauchbare Funktionen verkörpert, die in ihrer Skalierung und Verschiebung variabel während einer Analyse an das Signal angepasst werden können.

Wavelets

"Nothing to do with the sea or anything else"
Shinkichi Takahashi

wörtlich: "kleine Wellen"

Geschichte:

- Grundlagen 1873 bei Karl Weierstraß
- 1909 Alfred Haar: Orthonormalsystem von Funktionen mit kompaktem Träger ("Haar-Basis", auch heute noch verwendet)
- Bezeichnung "Wavelet" aus der Seismologie (Begriff geprägt von Norman Ricker)
- Wavelet-Theorie maßgeblich um 1980 von Jean Morlet und Alex Grossmann (2 frz. Physiker) entwickelt
- Yves Meyer und Stephane Mallat: Weiterentwicklung; "Multiresolution Analysis"
- breitere Anwendung ist relativ neu

Anwendungen:

- Approximationsverfahren
- Signalverarbeitung
- Bildanalyse; allg.: Datenanalyse (insbes. Zeitreihen)
- Bildkompression (JPEG2000-Standard)
- Bildklassifikation; Klassifikation von Histogrammen
- Bildbearbeitung

math. Grundlage der Haar-Wavelets im 1-dim. Fall:

Vektorraum V^j aller stückweise konstanten Funktionen, die auf $[0; 1)$ def. sind und auf jedem der 2^j gleich großen Teilintervalle von $[0; 1)$ konst. sind.

V^0 : konstante Funktionen auf $[0; 1)$

V^1 : konst. auf $[0; 0,5)$ und auf $[0,5; 1)$

usw.

$$V^0 \subseteq V^1 \subseteq V^2 \subseteq \dots$$

Skalarprodukt: wie üblich:

$$\langle f | g \rangle = \int_0^1 f(x) g(x) dx \quad f, g \in V^j$$

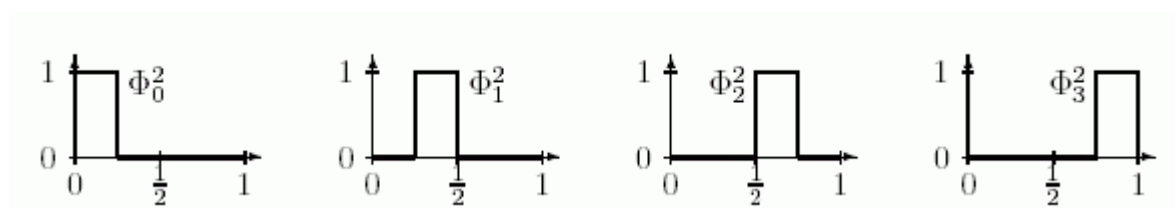
f orthogonal zu $g \Leftrightarrow \langle f | g \rangle = 0$

Eine Orthogonalbasis des V^j ist gegeben durch skalierte und verschobene Rechteckfunktionen ("Scaling-Funktionen"):

$$\Phi_i^j(x) = \Phi(2^j x - i), \quad i = 0, \dots, 2^j - 1,$$

$$\Phi(x) = \begin{cases} 1 & \text{für } 0 \leq x < 1 \\ 0 & \text{sonst.} \end{cases}$$

Beispiel: Rechteckbasis für V^2 :



Beispiel: Das 1-dim. Signal $f = (9 \ 7 \ 3 \ 5 \ 1 \ 1 \ 1 \ 5)$ kann mittels der Rechteckbasis des V^3 dargestellt werden in der Form

$$f(x) = c_0^3 \Phi_0^3 + c_1^3 \Phi_1^3 + c_2^3 \Phi_2^3 + c_3^3 \Phi_3^3 + c_4^3 \Phi_4^3 + c_5^3 \Phi_5^3 + c_6^3 \Phi_6^3 + c_7^3 \Phi_7^3$$

mit $c_0^3 = 9, c_1^3 = 7, \dots, c_7^3 = 5$. (Darstellung im Ortsbereich.)

Der Vektorraum W^j sei def. als das orthogonale Komplement von V^j in V^{j+1} .

Eine Menge von lin. unabh. Funktionen Ψ_i^j , die den Vektorraum W^j aufspannen, wird *Wavelet-Basis* genannt.

Grundlegende Eigenschaften einer solchen Basis:

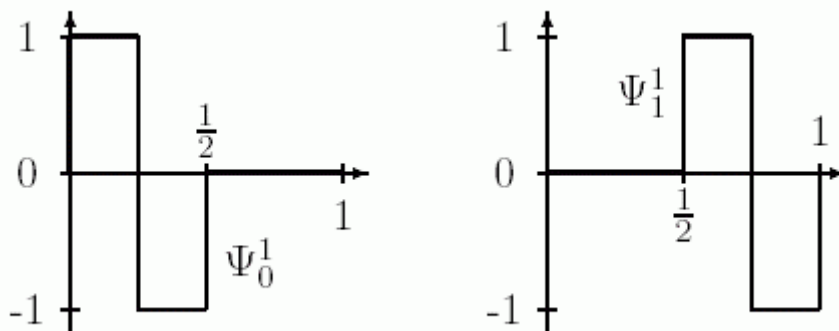
1. Die Basisfunktionen Ψ_i^j von W^j , zusammen mit den Basisfunktionen Φ_i^j von V^j formen eine Basis von V^{j+1} .
2. Jede Basisfunktion Ψ_i^j von W^j ist orthogonal zu jeder Basisfunktion Φ_k^j von V^j unter Verwendung des gewählten inneren Produktes.

Die Wavelets, die zur Rechteckbasis gehören, werden *Haar-Wavelets* genannt und haben folgenden Verlauf:

$$\Psi_i^j(x) = \Psi(2^j x - i), \quad i = 0, \dots, 2^j - 1,$$

$$\Psi(x) = \begin{cases} 1 & \text{für } 0 \leq x < 1/2 \\ -1 & \text{für } 1/2 \leq x < 1 \\ 0 & \text{sonst.} \end{cases}$$

Am Beispiel W^1 :



(Die Haar-Wavelets sind auch untereinander orthogonal.)

Eine Funktion aus V^3 kann dargestellt werden durch:

- die Basisfunktionen von V^3 (Rechteckbasis), oder
- die Basisfunktionen von V^2 und W^2 , oder
- die Basisfunktionen von V^1 , W^1 und W^2 , oder
- die Basisfunktionen von V^0 , W^0 , W^1 und W^2 .

Interpretation der Koeffizienten im letzteren Fall:

- Durchschnitt über das gesamte Signal
- und "Detailkoeffizienten" verschiedener Auflösungsebenen von grob bis fein.

Die Darstellung einer Funktion f aus V^n (oder die approximative Darstellung einer beliebigen Funktion) mittels der eindeutig bestimmten Koeffizienten bzgl. der Basisfunktionen von V^0, W^0, \dots, W^{n-1} heißt (*Haar-*) *Wavelet-Zerlegung* oder *Wavelet-Transformation* von f .

Beispiel:

Das Signal (9 7 3 5 1 1 1 5) (vgl. oben) wird in der Wavelet-Zerlegung dargestellt als

$$f(x) = c_0^0 \Phi_0^0 + d_0^0 \Psi_0^0 + d_0^1 \Psi_0^1 + d_1^1 \Psi_1^1 + d_0^2 \Psi_0^2 + d_1^2 \Psi_1^2 + d_2^2 \Psi_2^2 + d_3^2 \Psi_3^2$$

Die 8 Basisfunktionen in dieser Darstellung repräsentieren die *Haar-Basis* des Vektorraumes V^3 .

Gewöhnlich wird noch eine Normierung durchgeführt, um die Haar-Basis zu einer Orthonormalbasis (ONB) zu machen: Man ersetzt die obigen Definitionen durch

$$\Phi_i^j(x) = 2^{j/2} \Phi(2^j x - i)$$

$$\Psi_i^j(x) = 2^{j/2} \Psi(2^j x - i)$$

⇒ die Skalarprodukte dieser Basisfunktionen mit sich selbst sind 1

Im obigen Beispiel sind die Koeffizienten der (normierten) Haar-Basis:

$$\left[\begin{array}{cccccccc} 4 & 2 & \sqrt{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} & 0 & -1 \end{array} \right]$$

Vorteile der Normierung:

- ermöglicht Vergleich der einzelnen Wavelet-Koeffizienten hinsichtlich ihrer "Energie" (Quadrat der eukl. Norm des Koeffizientenvektors) – Koeffizienten mit dem größten Absolutbetrag haben größten Einfluss auf das Signal
- für die Basisvektoren b_i einer ONB gilt: die Koeffizienten lassen sich einfach als Skalarprodukte berechnen (vgl. DFT):

$$\sum_i \langle x, b_i \rangle b_i = x \quad \forall \langle b_i, b_j \rangle = 0 \wedge i \neq j$$

Praktische Gewinnung der Haar-Wavelet-Zerlegung:

Multiresolutions-Analyse

- sukzessive werden Mittelwerte jeweils zweier aufeinanderfolgender Funktionswerte gebildet
- der Informationsverlust wird ausgeglichen, indem zusätzlich die Abweichung vom Durchschnitt gespeichert wird ("Detailkoeffizienten") – nur für den ersten der beiden Ausgangswerte, die Abweichung des zweiten Wertes unterscheidet sich nur im Vorzeichen von der des ersten
- für die gewonnenen Durchschnittswerte wird der Prozess iteriert, bis nur noch 1 Durchschnittswert übrig bleibt

Beispiel: wieder obiges Signal (9 7 3 5 1 1 1 5)

Auflösung	Durchschnitte	Detailkoeffizienten
8	[9 7 3 5 1 1 1 5]	
4	[8 4 1 3]	[1 -1 0 -2]
2	[6 2]	[2 -1]
1	[4]	[2]

Als Wavelet-Zerlegung ergibt sich der Gesamtdurchschnitt, gefolgt von den Detailkoeffizienten beginnend mit der größten Auflösung:

(4 2 2 -1 1 -1 0 -2)

(hier fehlt noch die Normierung)

- das Verfahren ist leicht zu implementieren und effizient
- man braucht die Haar-Wavelets selber gar nicht zu kennen!

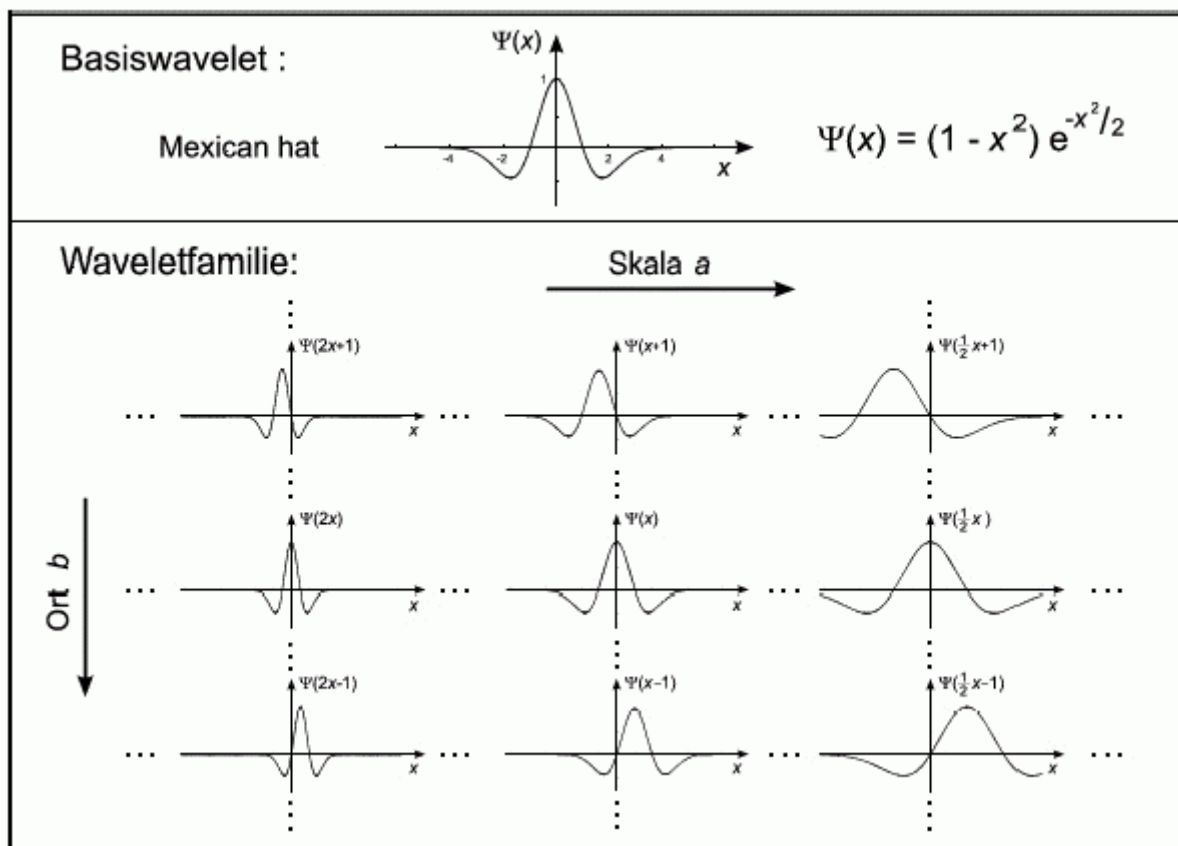
Anwendung: Da viele Detailkoeffizienten i.allg. sehr klein sind, kann man sie, ohne einen großen Informationsverlust zu erleiden, entfernen (*verlustbehaftete Wavelet-Kompression* – größere Kompressionsraten als bei der gefensterter DCT bei gleichen Qualitätsverlusten)

Allgemeine Wavelets:

Ausgangsobjekt ist ein *Basis- oder Mutter-Wavelet* Ψ .

In der Transformation kommen skalierte und örtlich verschobene Varianten dieses Basis-Wavelets zum Einsatz ("Wavelet-Familie").

Beispiel: Mexican hat - Waveletfamilie



Ein Wavelet aus der Familie wird beschrieben durch:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (a > 0, b \text{ beliebig}).$$

Ein Basiswavelet muss folgenden Bedingungen genügen:

- Zulässigkeitsbedingung:

$$0 < c_\psi := 2\pi \int \frac{|\psi(\omega)|^2}{|\omega|} d\omega < +\infty$$

d.h. das Wavelet unterscheidet sich nur auf einem begrenzten Intervall "deutlich" von 0;

- das Integral über die Waveletfunktion muss 0 sein:

$$\int \psi(x) dx = 0.$$

Für die Durchführbarkeit der Multiresolutions-Analyse benötigt man außerdem die *Dilatations-Gleichung* für die zugehörigen Skalierungsfunktionen:

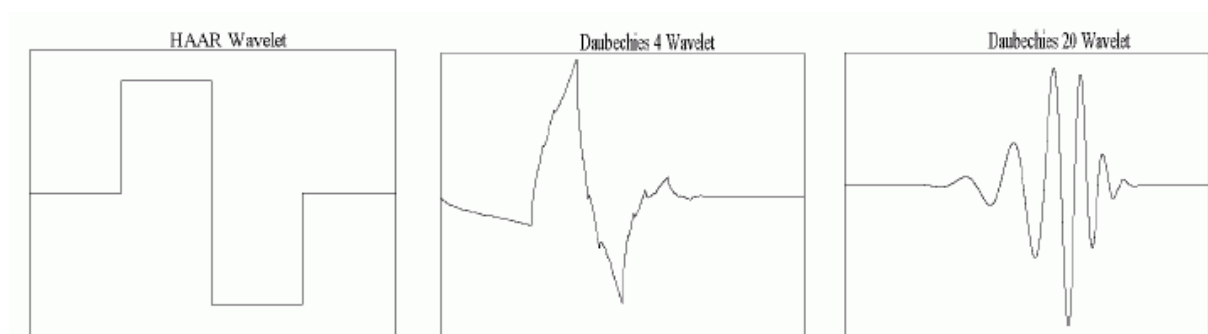
$$\Phi(x) = \sqrt{2} \sum_k c_k \Phi(2x - k)$$

$$\sum_k c_k = 1, \quad \sum_k (-1)^k c_k = 0$$

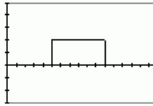
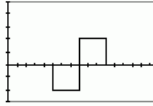
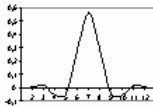
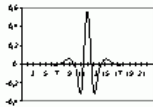
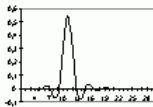
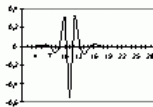
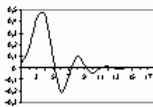
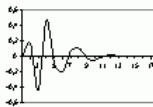
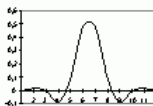
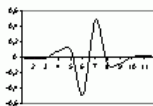
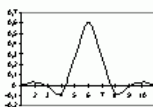
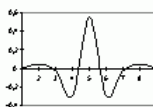
(aus Sakas 2002)

d.h. die Funktion kann als Linearkombination skaliertes, gedehnter und verschobener Versionen von sich selbst dargestellt werden (Selbstähnlichkeitseigenschaft).

Beispiele für verschiedene Basis-Wavelets:



Einige weitere Wavelets mit zugehörigen Skalierungsfunktionen:

Filterklasse	Scaling-Funktion	Wavelet	Beschreibung
Haar			Einfachste Filterklasse, Mittelwert- und Differenzfilter
CloseToCoiflet			
Coiflet			Frühe Waveletklasse, entwickelt von R.Coifman
Daubechies			nicht symmetrisch, streng orthogonal, strenger "compact support" selbstähnlich
Johnston-Barnard			
Biorthogonal-Spline			Symmetrisch, werden aus Binominalkoeffizienten berechnet

Merke:

es gibt nicht "die" Wavelet-Transformation, sondern zahlreiche verschiedene Varianten, je nach Wahl des zugrundeliegenden Basis-Wavelets!

kontinuierliche Wavelet-Transformation (1-dim.):
 durch Skalarprodukt mit den Mitgliedern der Wavelet-Familie
 (a, b : Skalierungs- bzw. Verschiebungs-Parameter)

$$CWT(a,b) = \int_{-\infty}^{+\infty} x(t) * \psi_{a,b}(t) dt$$

Interpretation:

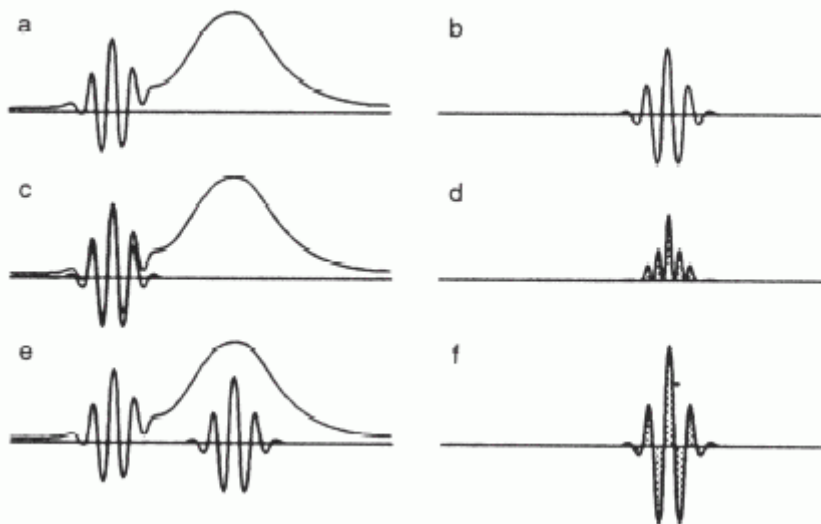
- Das Wavelet (skaliert und verschoben) tastet das Signal ab
- das Integral misst den Grad der Übereinstimmung

a: Ausgangsfunktion x

b: verwendetes Wavelet Ψ

c: Stelle mit guter Übereinstimmung, d: punktweises Produkt beider Funktionen an dieser Stelle (Wert des Integrals hoch)

e: Stelle mit schlechter Übereinstimmung, f: punktweises Produkt an dieser Stelle (Wert des Integrals niedrig)



(aus Esser 2001)

kontinuierliche WT würde Berechnung der Übereinstimmung für unendlich viele skalierte und verschobene Mitglieder der Wavelet-Familie erfordern:

- redundant wegen Überlappung von Wavelets
- extrem rechenintensiv

⇒ *diskrete Wavelet-Transformation* (DWT):

Berechnung nur für diskrete Werte von Skalierungsparameter a und Verschiebungsparameter b

oft: *dyadische DWT* mit $a = 2^m$ (m ganzzahlig).

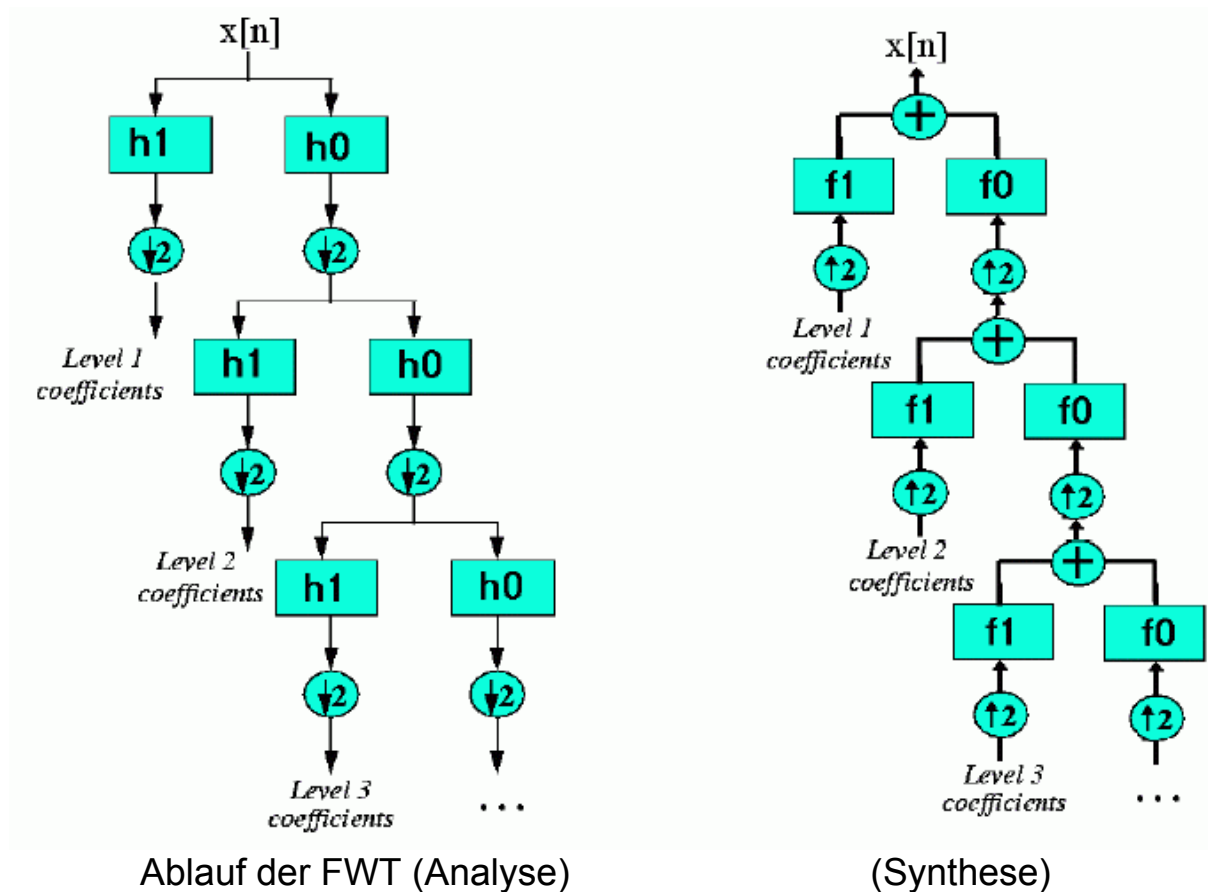
- auch hiermit (bei "naiver" Anwendung der Formel für die Wavelet-Koeffizienten) noch zu hoher Rechenaufwand

Durchbruch für die Wavelet-Transformation in der Bildverarbeitung: *Schnelle Wavelet-Transformation* (FWT) (Mallat & Meyer 1986)

- Verallgemeinerung der oben dargestellten *Multiresolutions-Analyse* (Multiresolution Analysis, MRA)
- Motivation: menschliches Sehsystem, Verarbeitung optischer Reize gleichzeitig auf mehreren Auflösungsebenen
- Kombination des Wavelet-Ansatzes mit klassischem Filter-Ansatz (Hochpass, Tiefpass) – vgl. oben die Mittelwertbildung!
- Mittelung benachbarter Pixel mittels einer Skalierungsfunktion (Tiefpass-Filter); die dabei wegfallenden Details (hohe Frequenzen) werden durch geeignete Wavelets aus einer Wavelet-Familie beschrieben – technisch durch Hochpass-Filterung
- um den Faktor 2 dezimierter Tiefpass-Anteil wird iterativ auf den nächstniedrigen Auflösungsstufe bearbeitet – bis das Signal kleiner als die Größe des Filters ist
- letzter Tiefpass liefert mittleren Grauwert (oder Farbintensitätswert) des analysierten Bildes
- kein Informationsverlust – alle Information wird aufgeteilt in Approximationen und Details

h_1, h_0 : Analyse- Hochpass- bzw. Tiefpass-Filter

Umkehrung des Verfahrens: *Synthese*, Einsatz entsprechender inverser Filter zur Rekonstruktion des Bildes aus dem mittleren Grauwert und den Detailkoeffizienten (Wavelet-Koeffizienten) – siehe Abb. (aus Esser 2001)



Zusammenhang zwischen Filtern und Wavelets bzw. Skalierungsfunktionen:

Sei h_0 ein Tiefpass-Filter (Faltungsmaske) der Länge N . Die Skalierungsfunktion ϕ errechnet sich rekursiv durch:

$$\phi(x) = 2 \sum_{k=0}^N h_0(k) \phi(2x - k)$$

Die Wavelet-Funktion Ψ kann aus dem zugehörigen Hochpassfilter h_1 und der Skalierungsfunktion folgendermaßen bestimmt werden:

$$\Psi(x) = 2 \sum_{k=0}^N h_1(k) \phi(2x - k)$$

Die Menge der in der Analyse und Synthese verwendeten, zusammengehörigen Hoch- und Tiefpassfilter bezeichnet man als *Filterbank*.

Beispiele:

Die Filterbänke, die den Haar-Wavelets und den Daubechies-4-Wavelets entsprechen (aus Esser 2001):

Haar: (Filterlänge = 2)

$$h_0: \begin{array}{|c|c|} \hline 0.70710678118655 & 0.70710678118655 \\ \hline \end{array}$$

$$h_1: \begin{array}{|c|c|} \hline 0.70710678118655 & -0.70710678118655 \\ \hline \end{array}$$

$$f_0: \begin{array}{|c|c|} \hline 0.70710678118655 & 0.70710678118655 \\ \hline \end{array}$$

$$f_1: \begin{array}{|c|c|} \hline 0.70710678118655 & 0.70710678118655 \\ \hline \end{array}$$

Daubechies 4: (Filterlänge = 8)

$$h_0: \begin{array}{|c|c|c|c|} \hline 0.23037781330886 & 0.71484657055254 & 0.63088076792959 & -0.02798376941698 \\ \hline -0.18703481171888 & 0.03084138183599 & 0.03288301166698 & -0.010597401785 \\ \hline \end{array}$$

$$h_1: \begin{array}{|c|c|c|c|} \hline -0.010597401785 & -0.03288301166698 & 0.03084138183599 & 0.18703481171888 \\ \hline -0.02798376941698 & -0.63088076792959 & 0.71484657055254 & -0.23037781330886 \\ \hline \end{array}$$

$$f_0: \begin{array}{|c|c|c|c|} \hline -0.010597401785 & 0.03288301166698 & 0.03084138183599 & -0.18703481171888 \\ \hline -0.02798376941698 & 0.63088076792959 & 0.71484657055254 & 0.23037781330886 \\ \hline \end{array}$$

$$f_1: \begin{array}{|c|c|c|c|} \hline -0.23037781330886 & 0.71484657055254 & -0.63088076792959 & -0.02798376941698 \\ \hline 0.18703481171888 & 0.03084138183599 & -0.03288301166698 & -0.010597401785 \\ \hline \end{array}$$

Zusammenfassung der Wavelet-Transformation (Analyse)
im 1-dimensionalen Fall:

gegeben: Tiefpassfilter h_1 , Hochpassfilter h_0
Signal f

Transformation:

$$f \mapsto \{h_0 f, h_0 h_1 f, h_0 h_1^2 f, \dots, h_0 h_1^{j-1} f, h_1^j f\}$$

Anwendung auf 2-dimensionale Funktionen (Bilder)

3 Varianten:

(a) Standard-Zerlegung

1-dim. WT wird auf jede Zeile angewendet. \Rightarrow jede Spalte besteht aus Durchschnitten oder Wavelet-Koeffizienten

1-dim. WT wird auf jede Spalte angewendet.

\Rightarrow bis auf 1 Wert für den Durchschnitt des gesamten Bildes erhält man nur Detailkoeffizienten

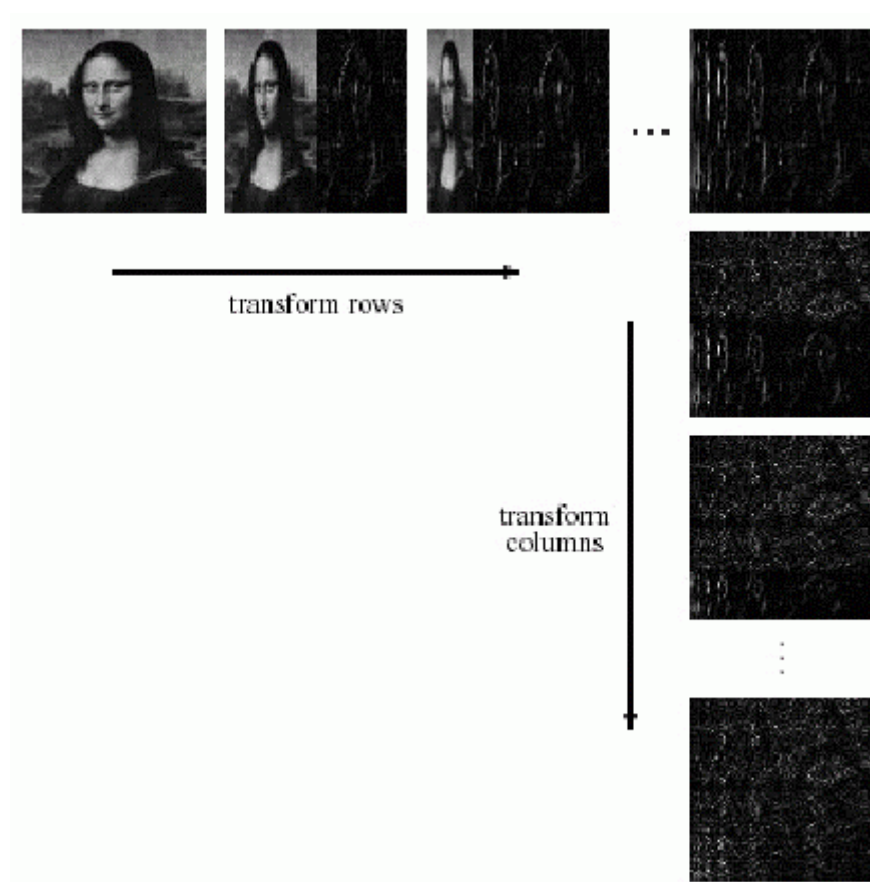
(b) Nichtstandard-Zerlegung

Operationen für Zeilen und Spalten wechseln sich ab:

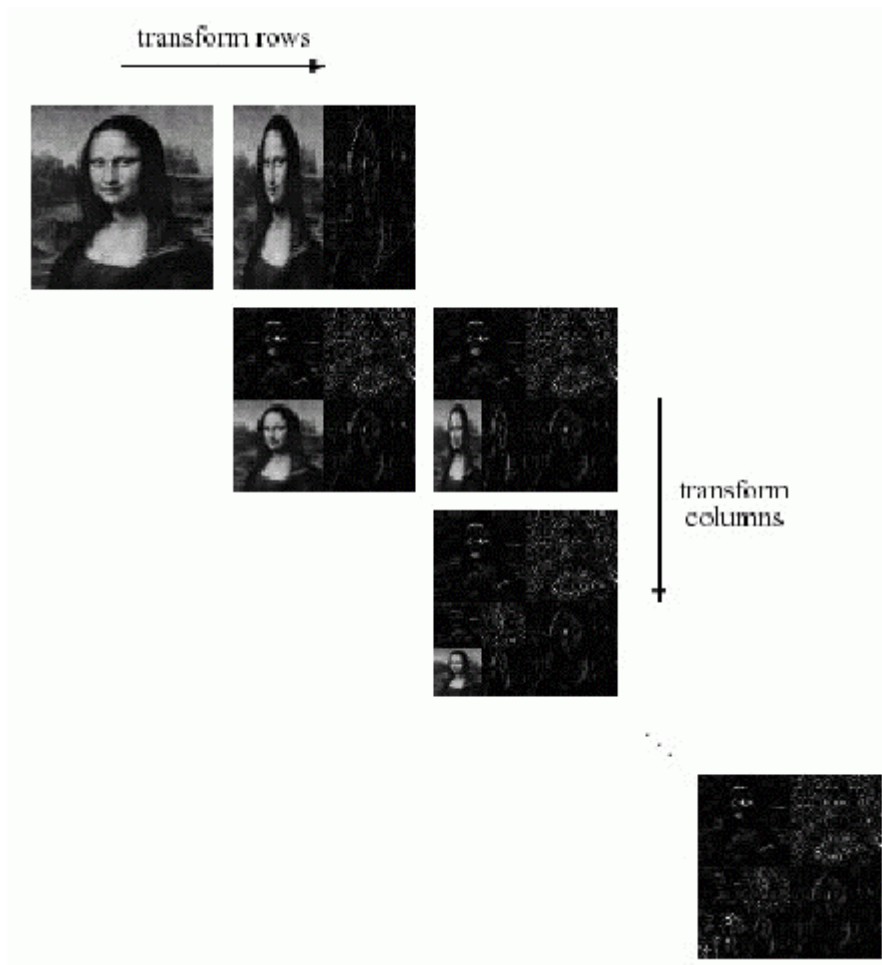
Durchschnitte und Detailkoeffizienten einer Auflösungsstufe werden für jede Zeile bestimmt

dann derselbe Vorgang für jede Spalte

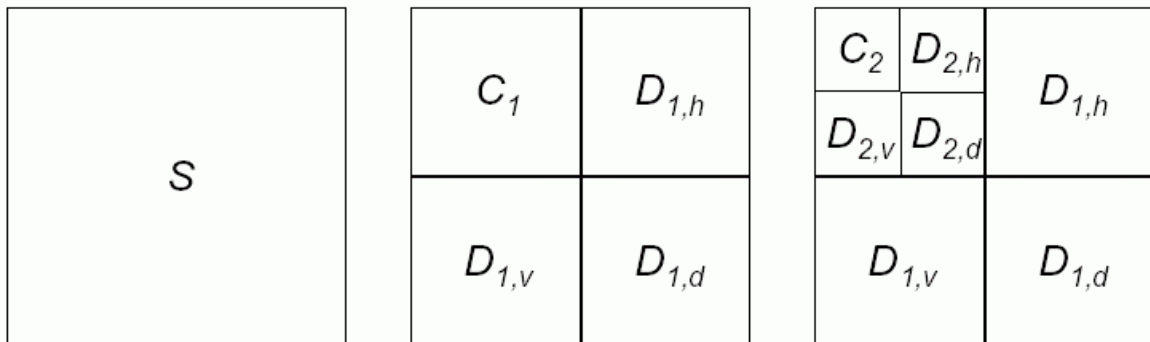
dann wieder für jede Zeile, usw.



Standard-Zerlegung

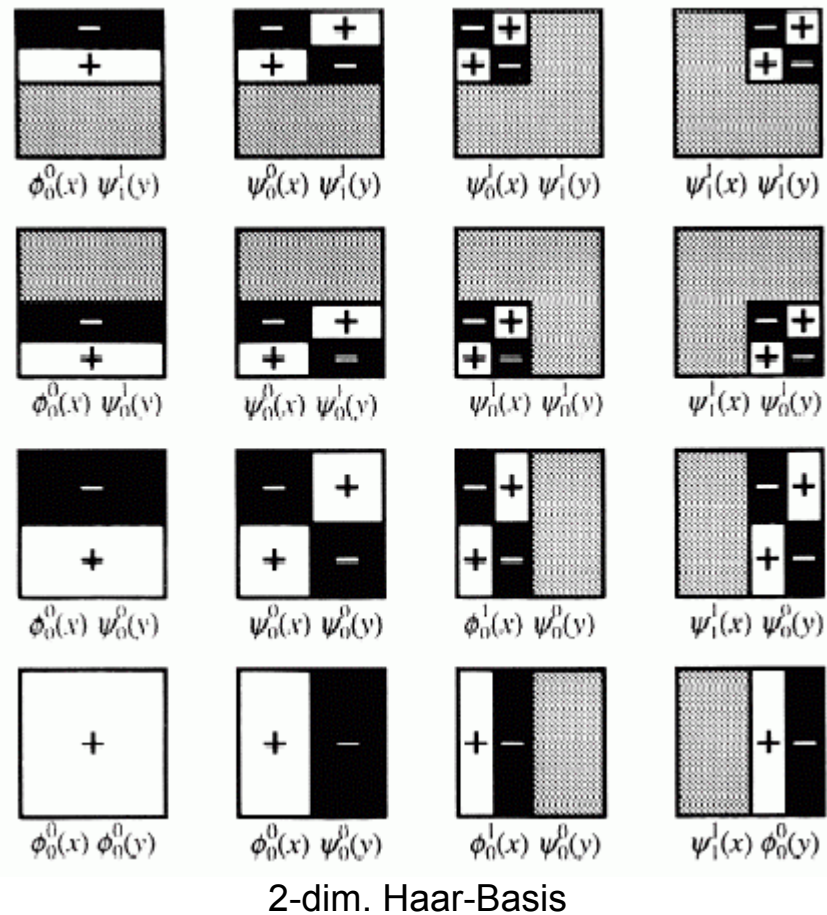


Nichtstandard-Zerlegung



Nichtstandard-Zerlegung (aus Sakas 2002)

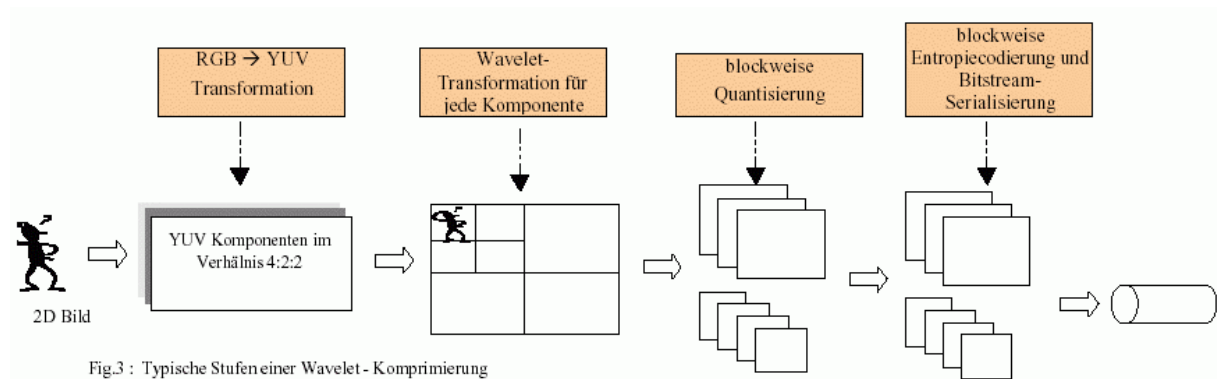
(c) Block-Zerlegung:
Verwendung von 2-dim. Filtermasken für die Multiresolutions-Analyse



Anwendung der Wavelet-Transformation für die Bildkompression:

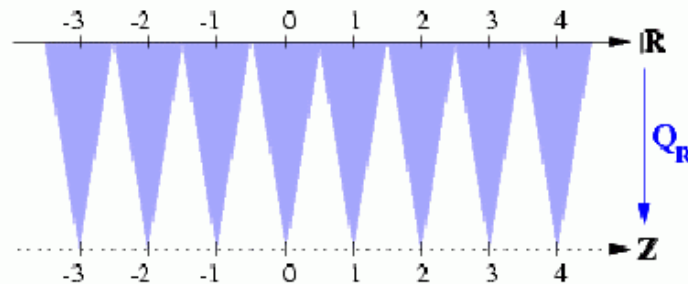
Weglassen der Detailinformationen der höheren Auflösungsstufen

Genauer:



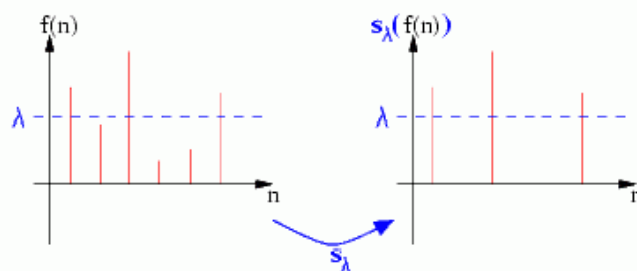
Unter Quantisierung versteht man das Ausdünnen des Wertebereichs einer Funktion, wobei ein Quantum von Werten einem einzigen neuen Wert zugeordnet wird.

Bsp: Rundungsoperation



Alle Koeffizienten, die betragsmäßig kleiner als bestimmter Wert λ sind, werden durch die Schwellwertfunktion s_λ auf Null abgebildet

$$s_\lambda(x) = \begin{cases} 0 & \text{falls } |x| < \lambda \\ x & \text{sonst} \end{cases}$$



Achtung: Informationsverlust !

(aus Sevenich 2000)

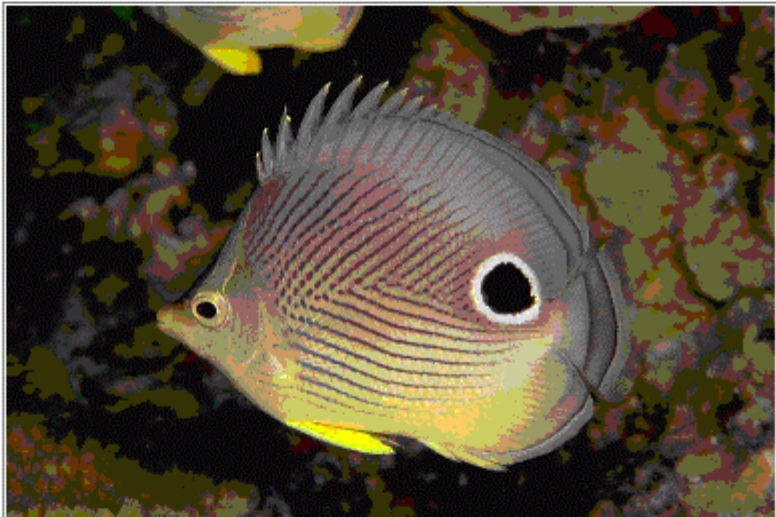
Beispiel: Lurawave-Kompression (Fa. LuraTec, Berlin-Adlershof),
Format LWF (aus Kroll & Zerbe 2002)

Originalbild
GIF = 35160 Bytes
BMP=294 KBytes
TIF=289 KBytes

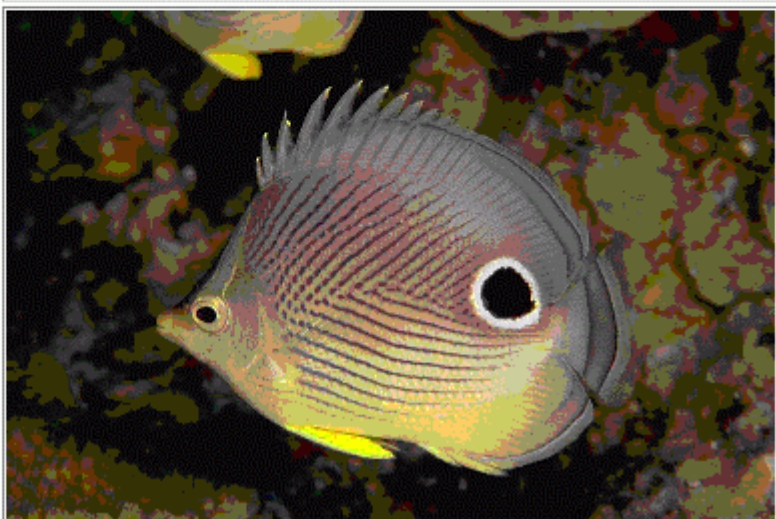
(c=1 bezügl. TIF)



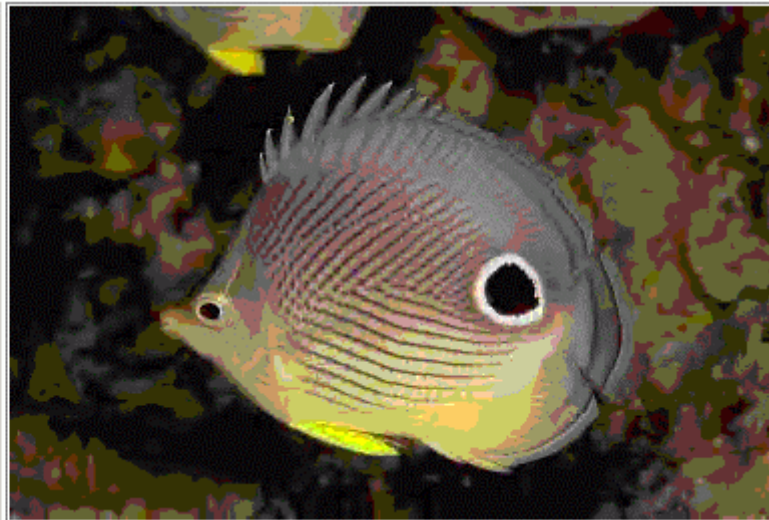
LWF = 58980 Bytes
c = 1:5



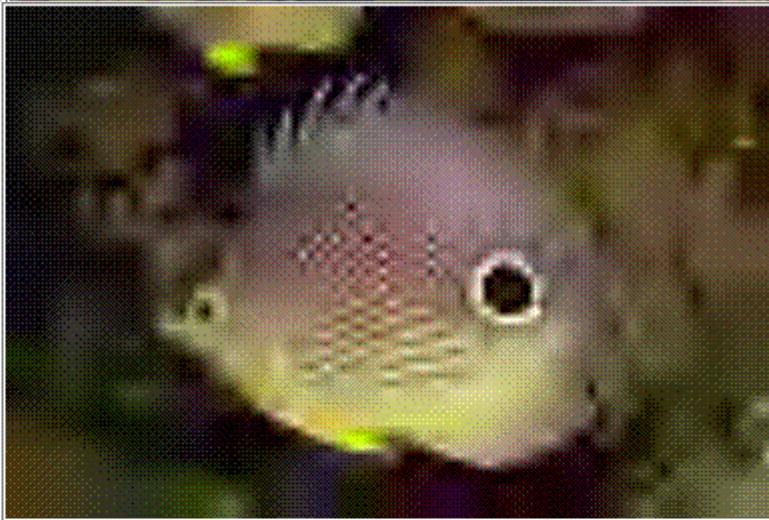
LWF = 29488 Bytes
c = 1:10



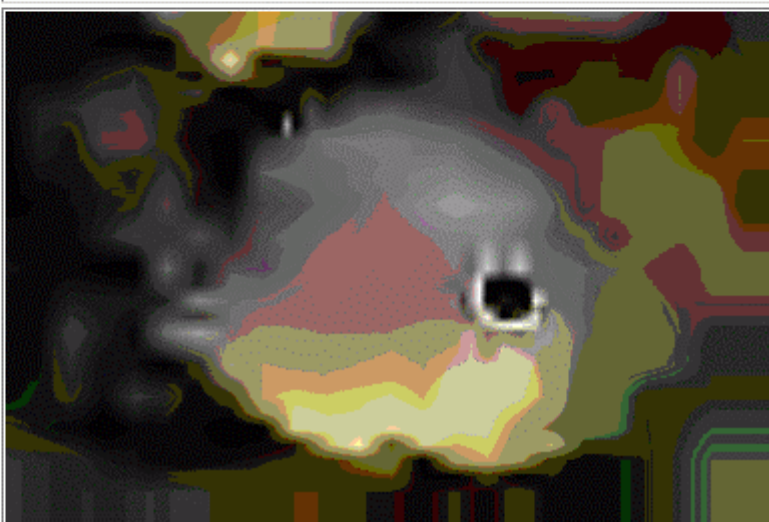
LWF = 5896 Bytes
c = 1:50



LWF = 588 Bytes
c = 1:500



LWF = 292 Bytes
c = 1:1000



weitere Anwendung der Wavelet-Transformation:

Bildanfrage

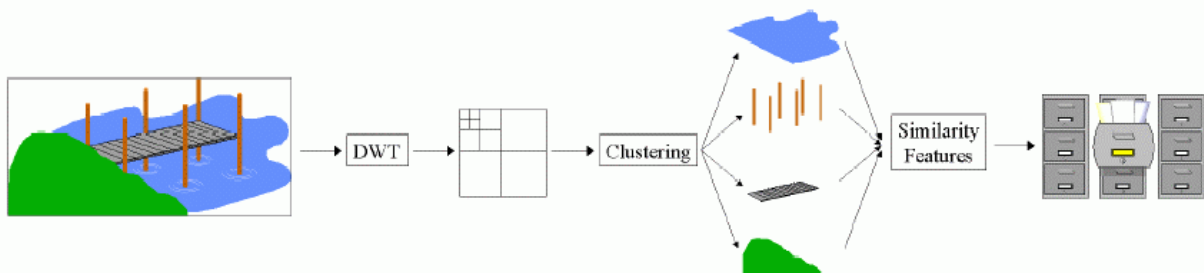
→ finde in Datenbank ähnliche Bilder

Vorgehen:

1. bestimme Waveletkoeffizienten aller Bilder (Vorverarbeitung)
2. bestimme Waveletkoeffizienten im Anfragebild
3. bestimme Übereinstimmung anhand übereinstimmender (ähnlicher) Koeffizienten

besser mit nachgeschalteter Merkmalsextraktion und Klassifikation.

Verfahren nach Ardizzoni et al. (1999):



Beispiel für Suchanfrage und Ergebnisse (aus großer Bilddatenbank):



(query)



(WS1)



(WS2)



(WS3)



(WS4)



(WS5)

2. Beispiel:



(query)



(WS1)



(WS2)



(WS3)



(WS4)



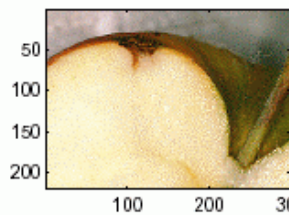
(WS5)

anderes Beispiel (aus Woodford et al. 1999):

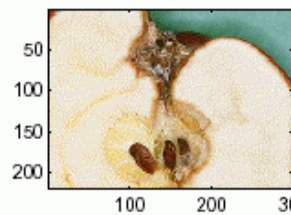
Query Image



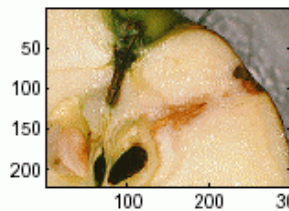
Distance of cmftbor4.bmp is 44.3394



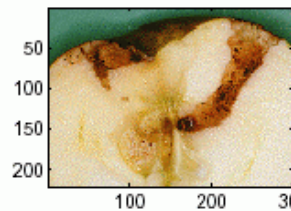
Distance of lrftbor1.bmp is 56.0545



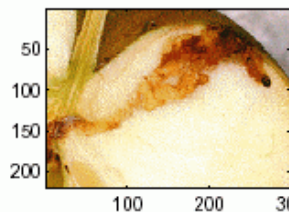
Distance of cmftbor6.bmp is 73.7659



Distance of cmftbor5.bmp is 74.807



Distance of cmftbor1.bmp is 76.2394



Vorteile der Wavelet-Transformation:

- Zeitkomplexität $O(n)$
- Koeffizientenmatrix oft dünn besetzt \Rightarrow starke Kompression möglich
- Flexibilität durch Wahl des Basis-Wavelets bzw. der Filterbank
- ausgewogenes Verhältnis von Orts- und Frequenzinformation
- Änderungen im Bild bleiben auch im Wavelet-transformierten Bild in ihren Auswirkungen beschränkt