

## 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:

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

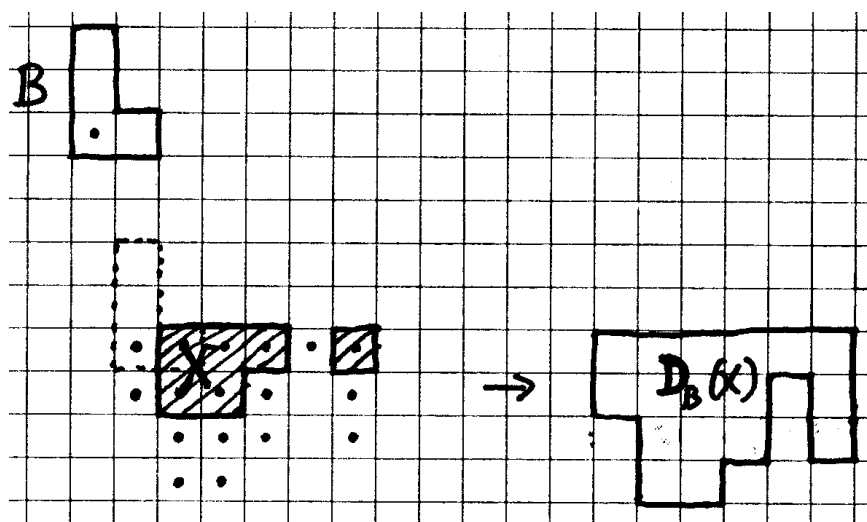
$$= X - b_1 \cup X - b_2 \cup \dots \quad (b_i \in B)$$

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

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

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

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")

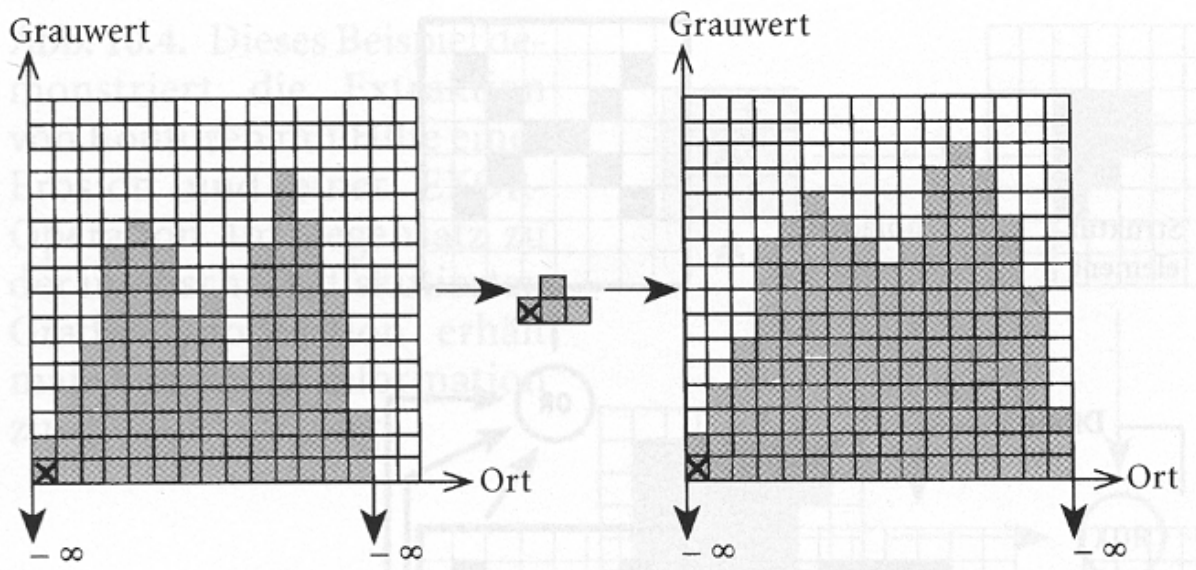
Def. für Funktionen  $f$  und  $B$  ( $D(B)$  Definitionsbereich von  $B$ ):

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

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

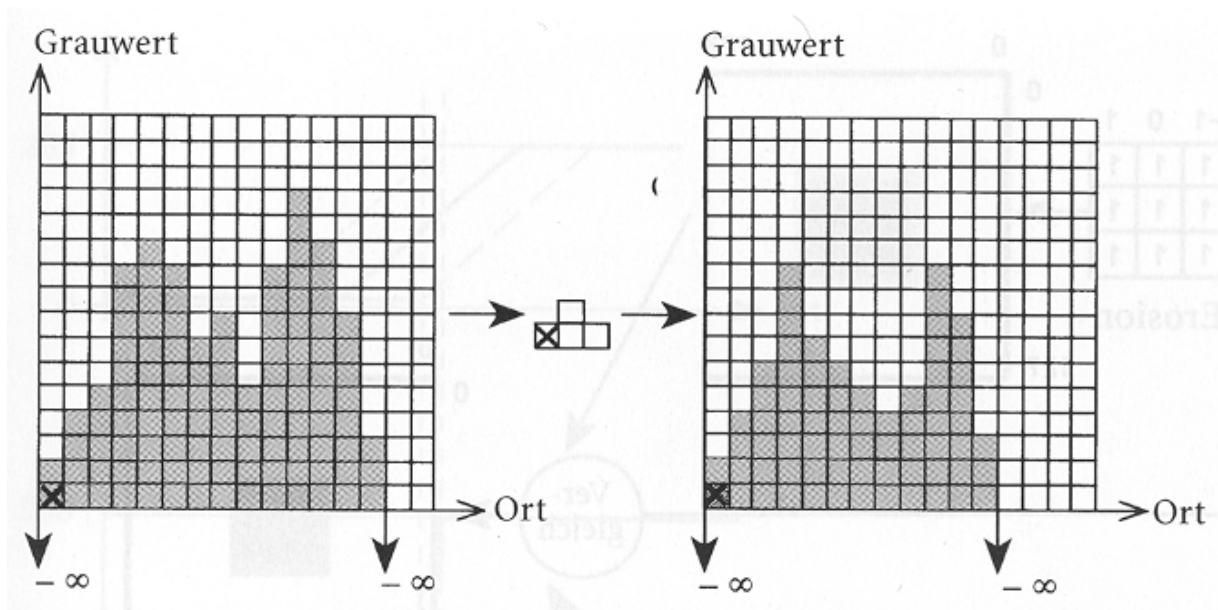
- Beachte: den Spezialfall ( $B$  Menge; man sagt: "B flach") erhält man hieraus durch eine Funktion  $B$ , die auf ihrem Definitionsbereich 0 ist.
- Anschauliche Interpretation auch hier: Dilatation bzw. Erosion der Menge unter dem Graphen von  $f$ , jetzt durch ein dreidimensionales Strukturelement.

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



(aus Bäsmann & Kreyss 1998)

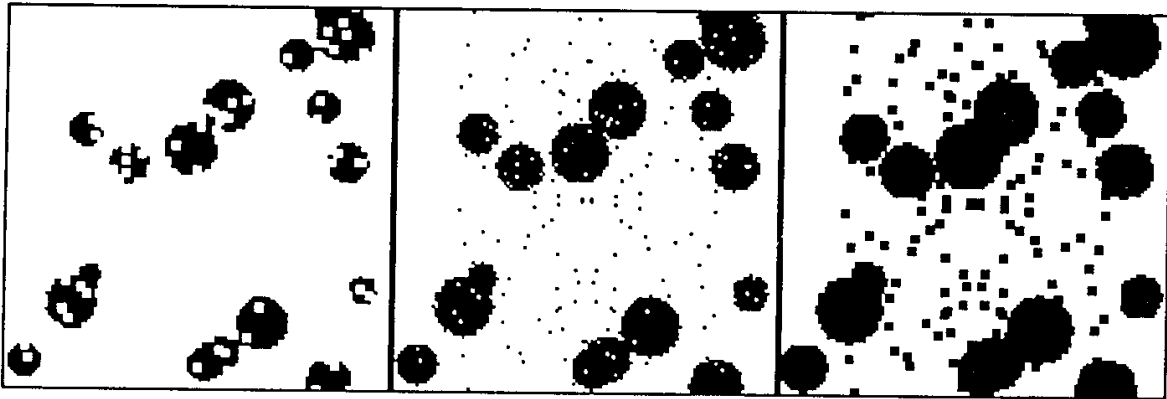
Analog für die Erosion:



Eigenschaften von Dilatation und Erosion:

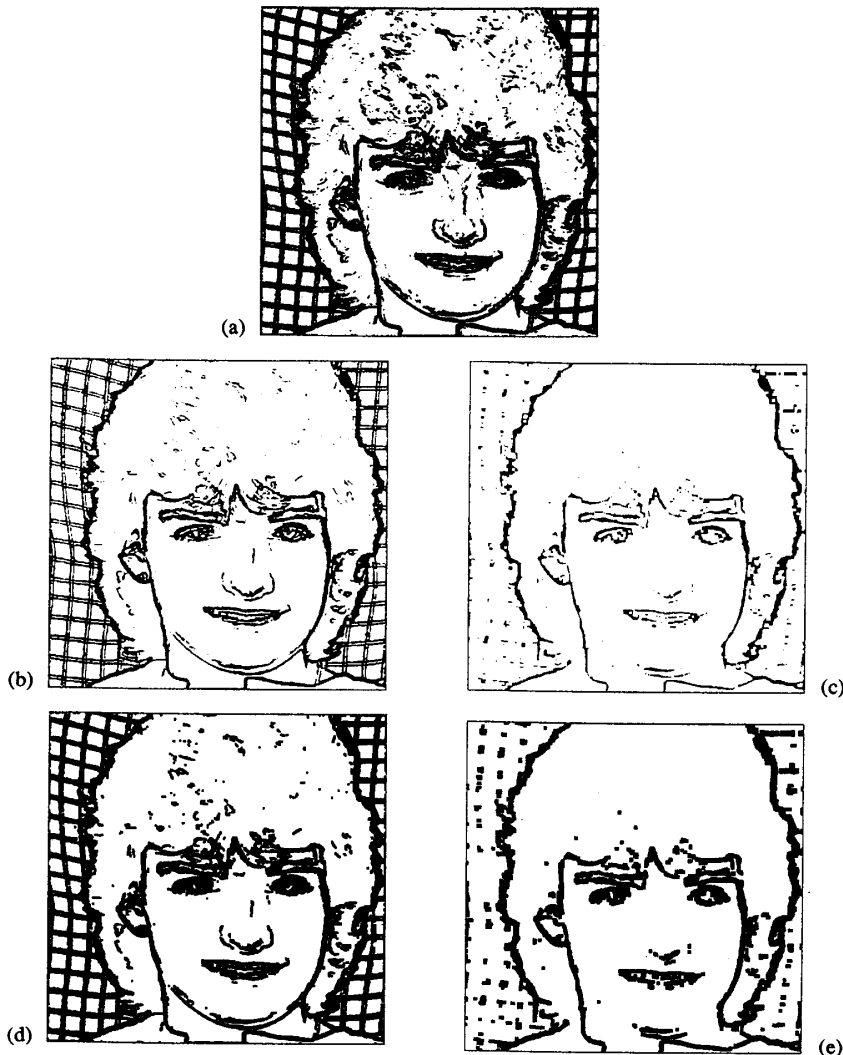
- Dualität:  
 $\mathbf{E} = \mathbf{C} \mathbf{D} \mathbf{C}$   
 $\mathbf{D} = \mathbf{C} \mathbf{E} \mathbf{C}$
- Verschiebungsinvarianz:  
 $(\mathbf{D}(f_c)) = (\mathbf{D}(f))_c$   
 $(\mathbf{E}(f_c)) = (\mathbf{E}(f))_c$
- Monotonie:  
 $f \leq g \Rightarrow \mathbf{D}(f) \leq \mathbf{D}(g), \mathbf{E}(f) \leq \mathbf{E}(g)$
- 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**):



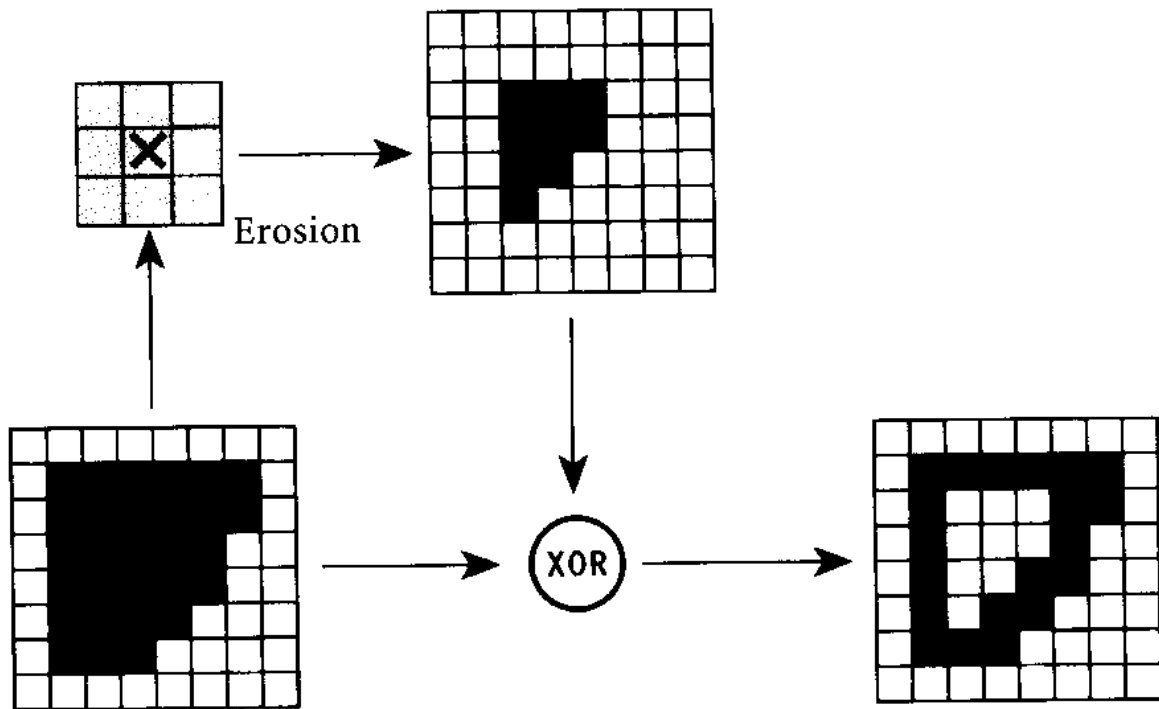
(aus Voss & Süße 1991)

Einsatz zur Manipulation von Kantenbildern aus der klassischen  
Filterung:

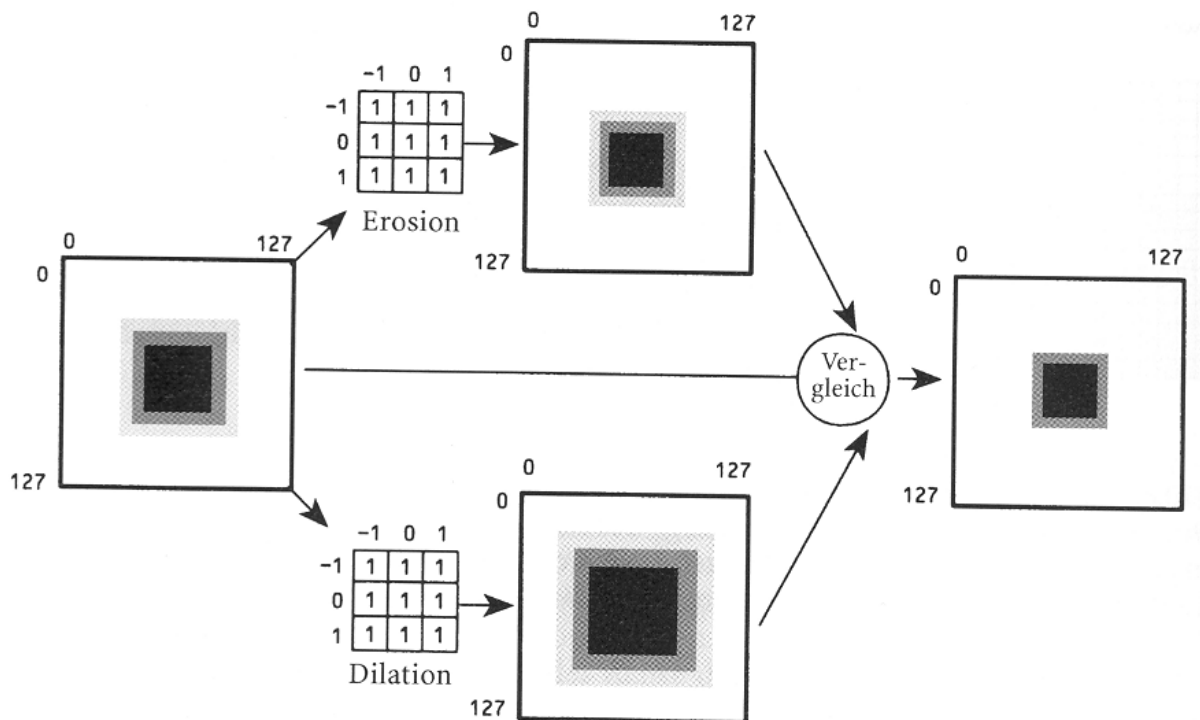


(b), (c) Erosionen von (a); (d), (e) Dilatationen (aus Haberäcker 1991)

### Extraktion von Konturen:



### Adaptive Kontrastverbesserung:



Zusammengesetzte morphologische Operationen:

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

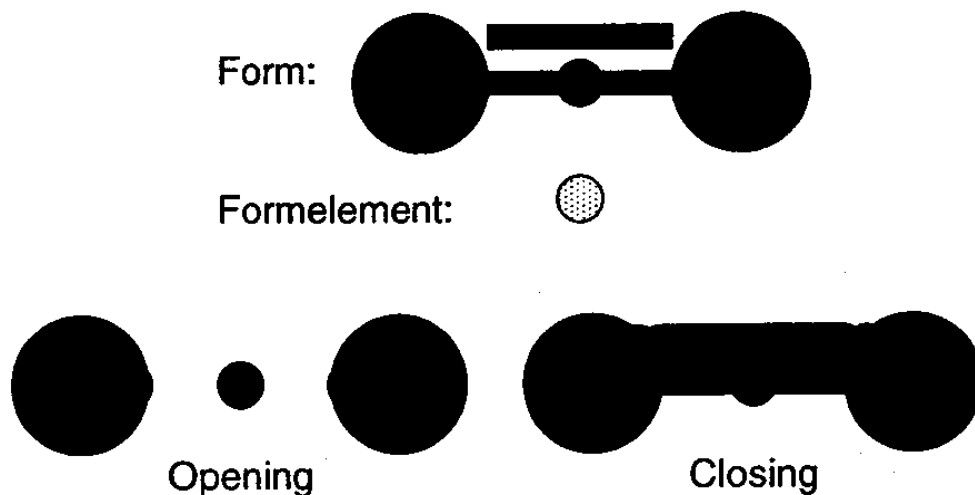
bewirkt Aufbrechen von Regionen an Stellen, wo nur schwacher Zusammenhang bestand

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

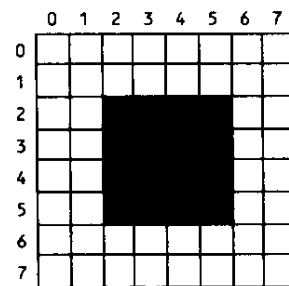
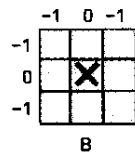
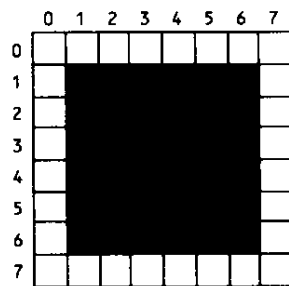
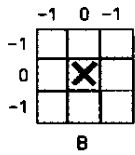
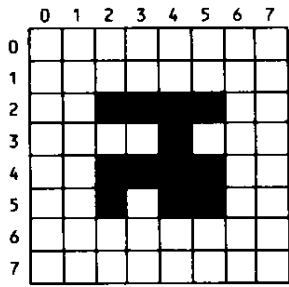
schließt kleine Zwischenräume

Eigenschaften:

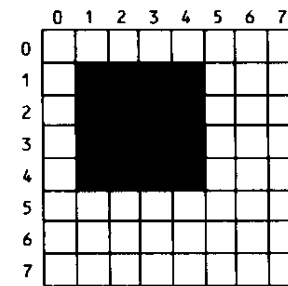
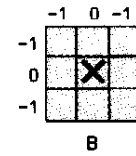
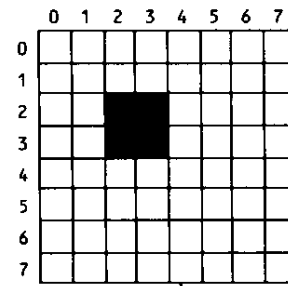
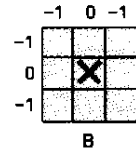
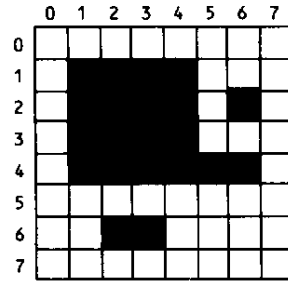
- Idempotenz  
 $\mathbf{OO} = \mathbf{O}, \mathbf{SS} = \mathbf{S}$
- Dualität  
 $\mathbf{S} = \mathbf{COC}, \mathbf{O} = \mathbf{CSC}$
- Schließung extensiv, Öffnung anti-extensiv:  
 $\mathbf{O} \leq \mathbf{I} \leq \mathbf{S}$
- Monotonie:  
 $f \leq g \Rightarrow \mathbf{O}(f) \leq \mathbf{O}(g),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







a



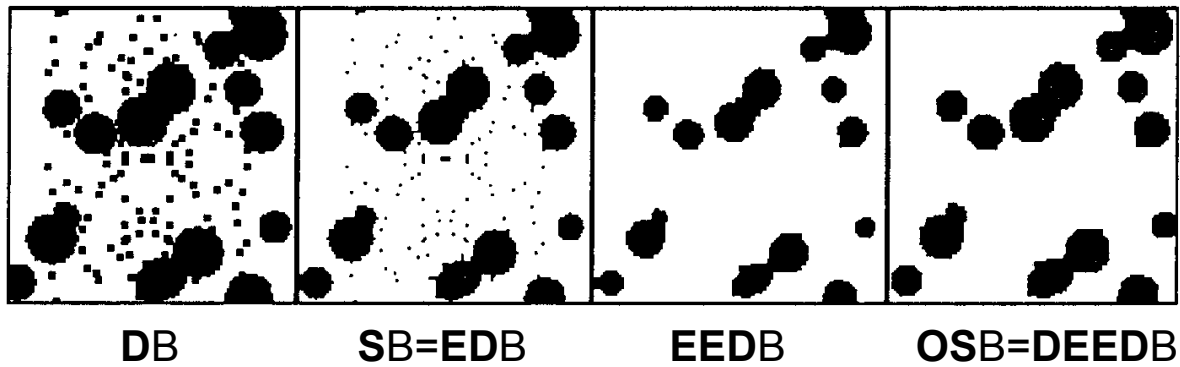
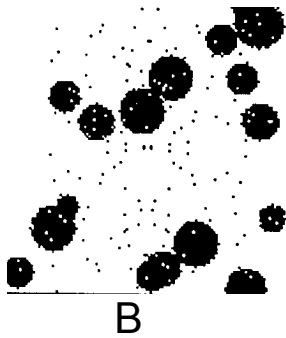
b

- a: Schließung
- b: Öffnung

Öffnung: Abtragen "ausgefranster" Bereichsränder, Eliminieren kleiner Partikel

Schließung: schließt Lücken zwischen "Fransen"

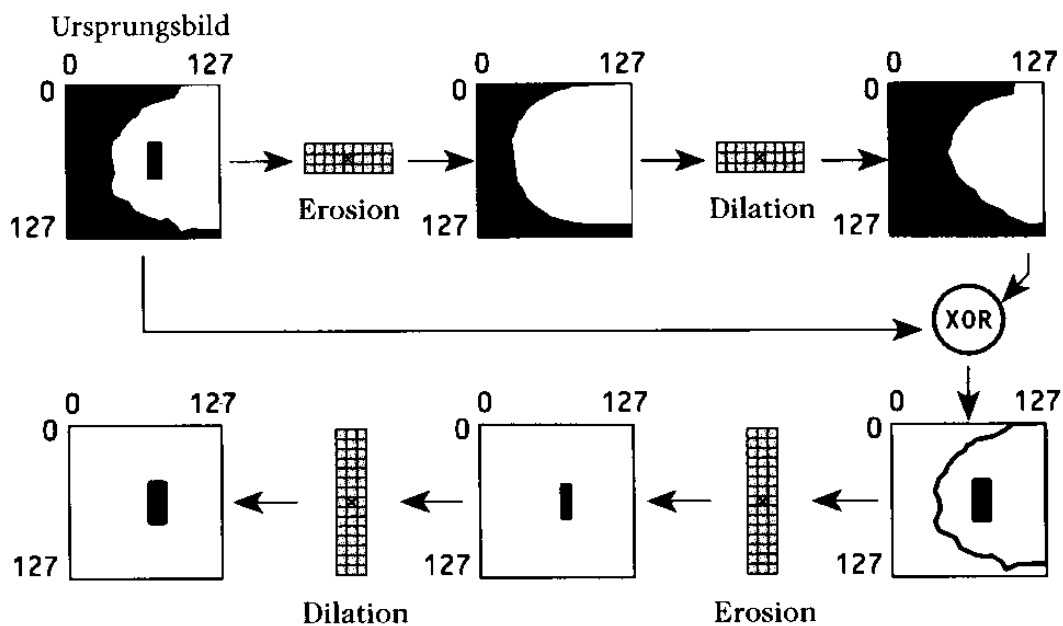
Anwendung zur Rauschunterdrückung (erst schließen, dann öffnen):



(aus Voss & Süße 1991)

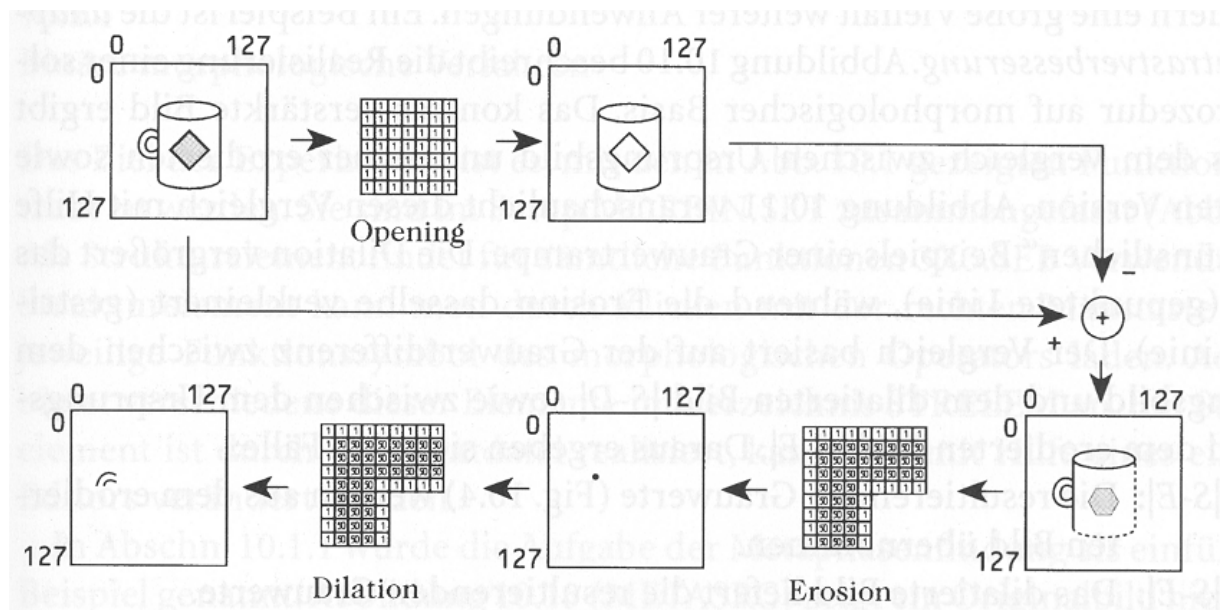
Einsatz spezieller strukturierender Elemente zur *Detektion von Binärbildanteilen mit bekannter Form*

Beispiel: Rechteck in der Mitte des Originalbildes:

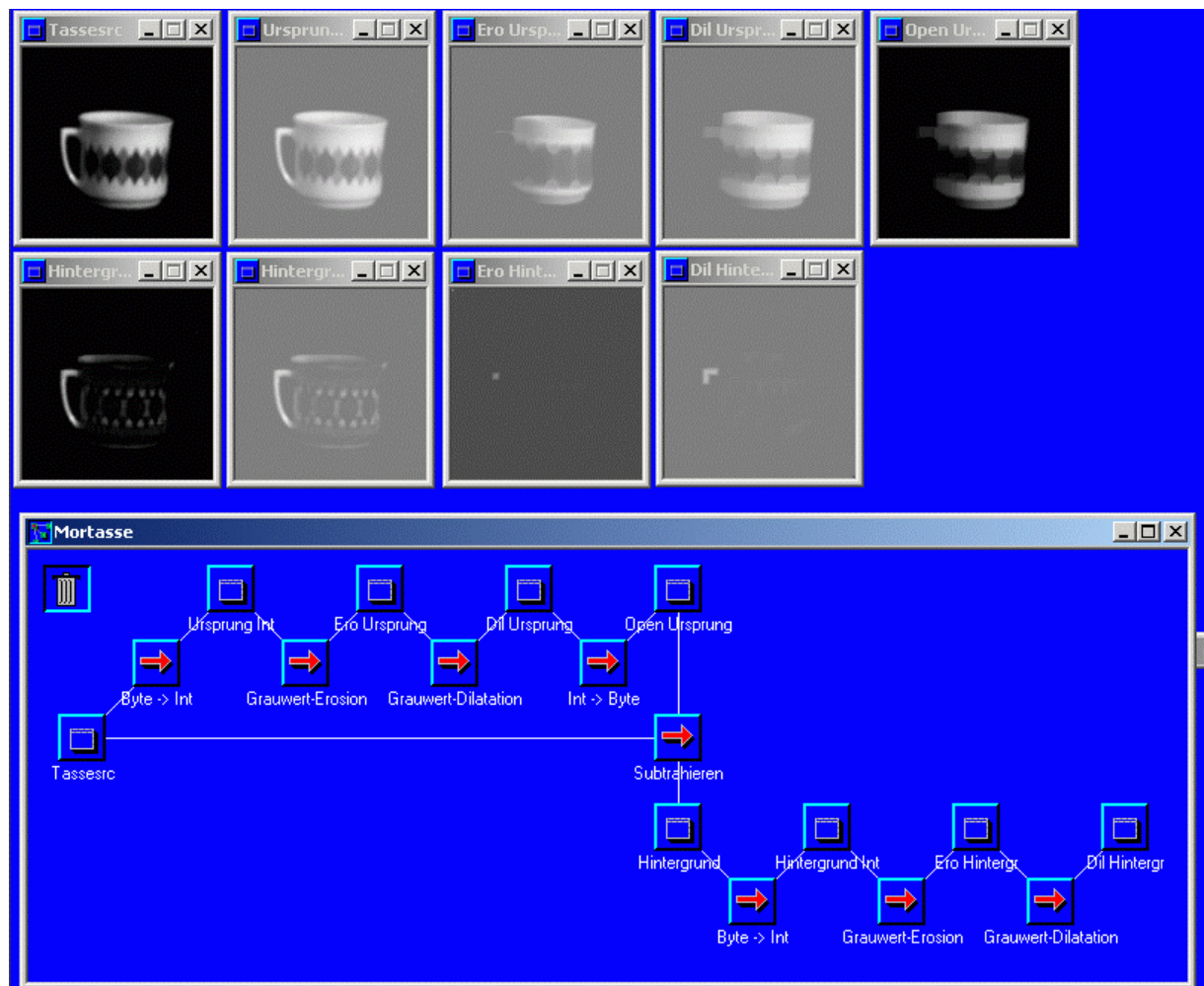


Öffnung zur Hintergrundschätzung, XOR liefert interessierende Bereiche, zweites Opening filtert gewünschten Bereich heraus (aus Bässmann & Kreyss 1998)

## Übertragung der Vorgehensweise auf Grauwertbilder:



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



Dilatationsresiduum:

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

Erosionsresiduum:

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

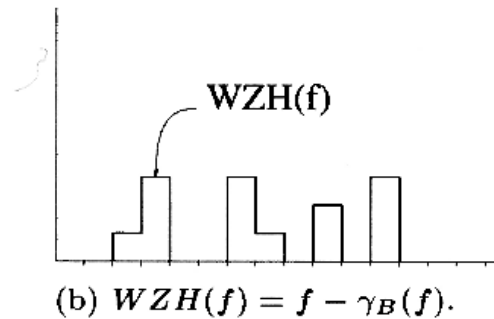
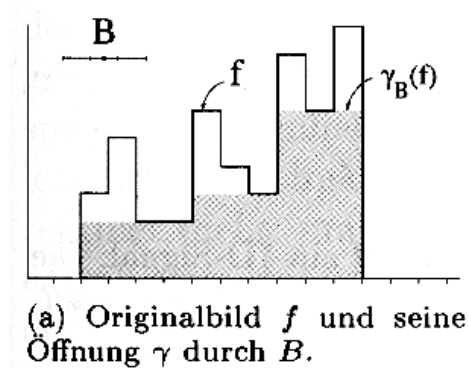
Zylinderhut-Transformationen:

Weißer Zylinderhut:

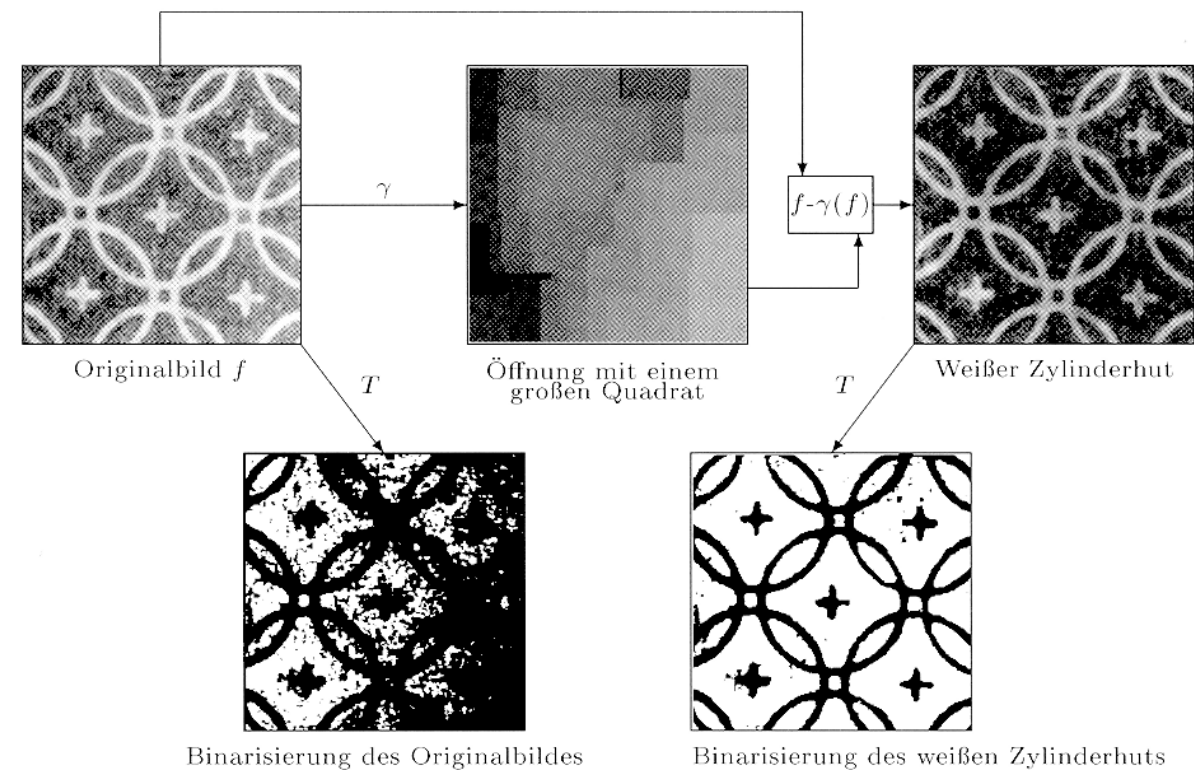
$$\mathbf{WZH}(f) = f - \mathbf{O}f$$

Schwarzer Zylinderhut:

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



Einsatz des Zylinderhuts zur Shadingkorrektur:



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

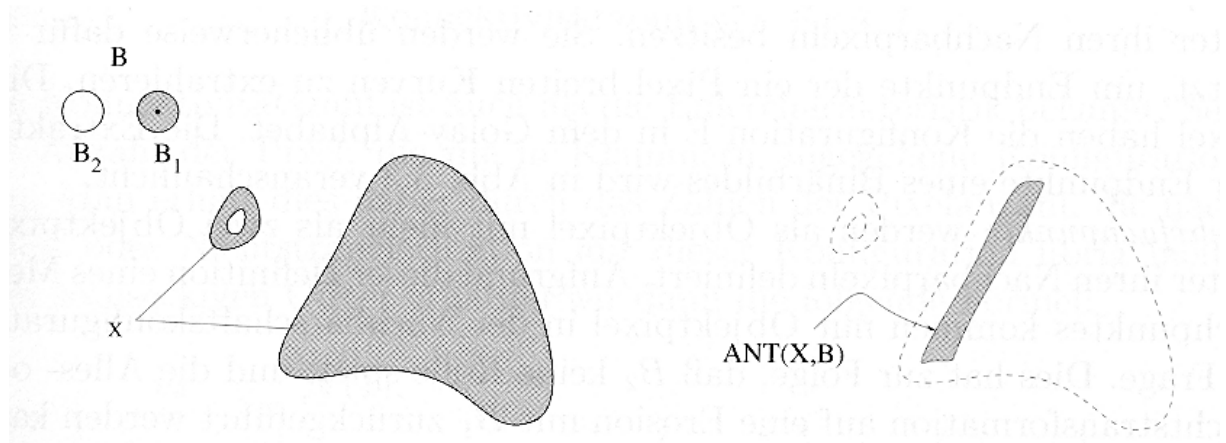
- $B_1$  und  $B_2$  sind disjunkt
- besitzen denselben Bezugspunkt (Nullpunkt)

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*:

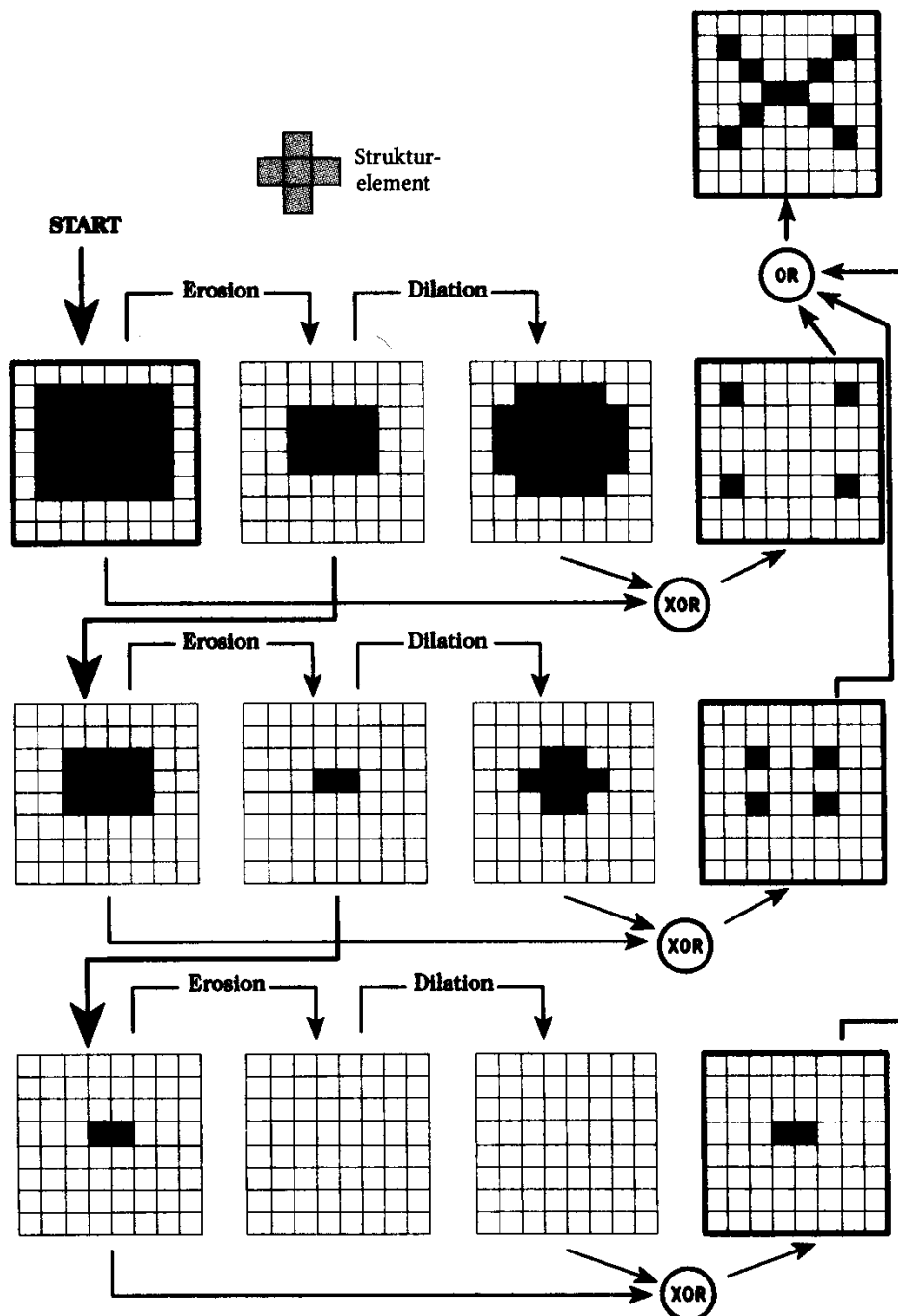
analog mit Vereinigungsmenge statt Mengendifferenz

## 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)

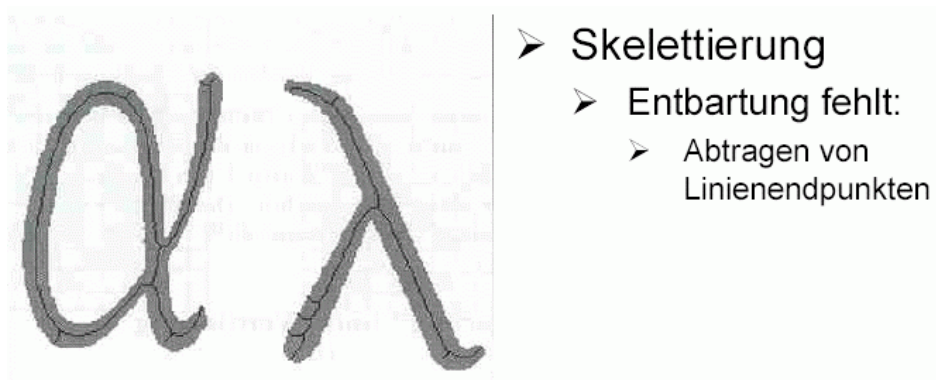
bessere Verfahren als das oben gezeigte:

Erosion + bestimmte Bedingungen

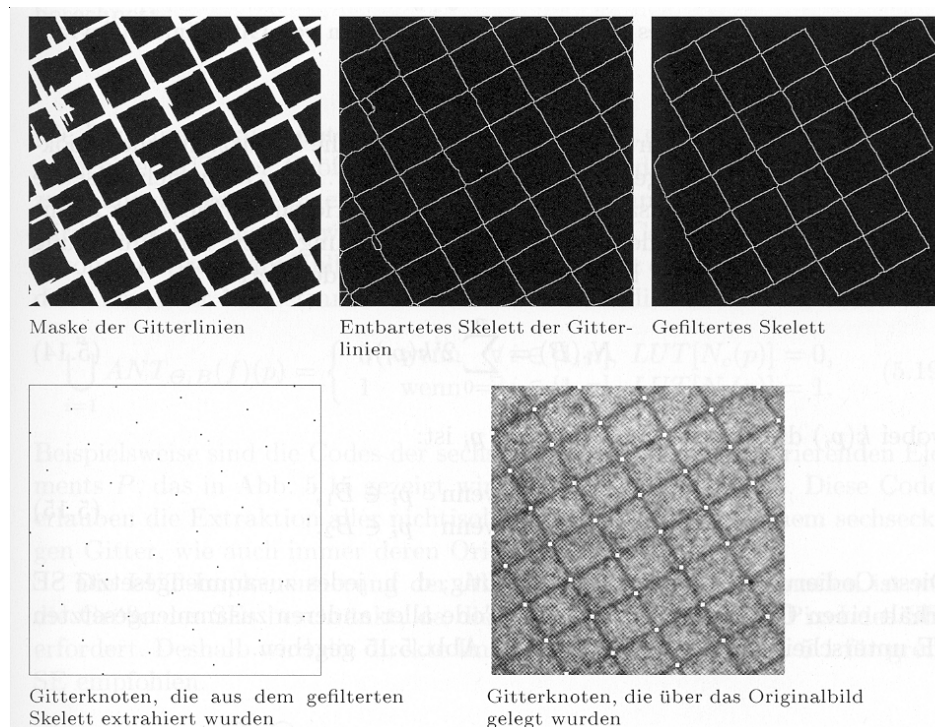
- Skelettlinie darf nur 1 Pixel breit sein
- muss in der Mitte des Objekts verlaufen
- darf sich bei weiteren Schritten nicht weiter verändern

⇒ 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

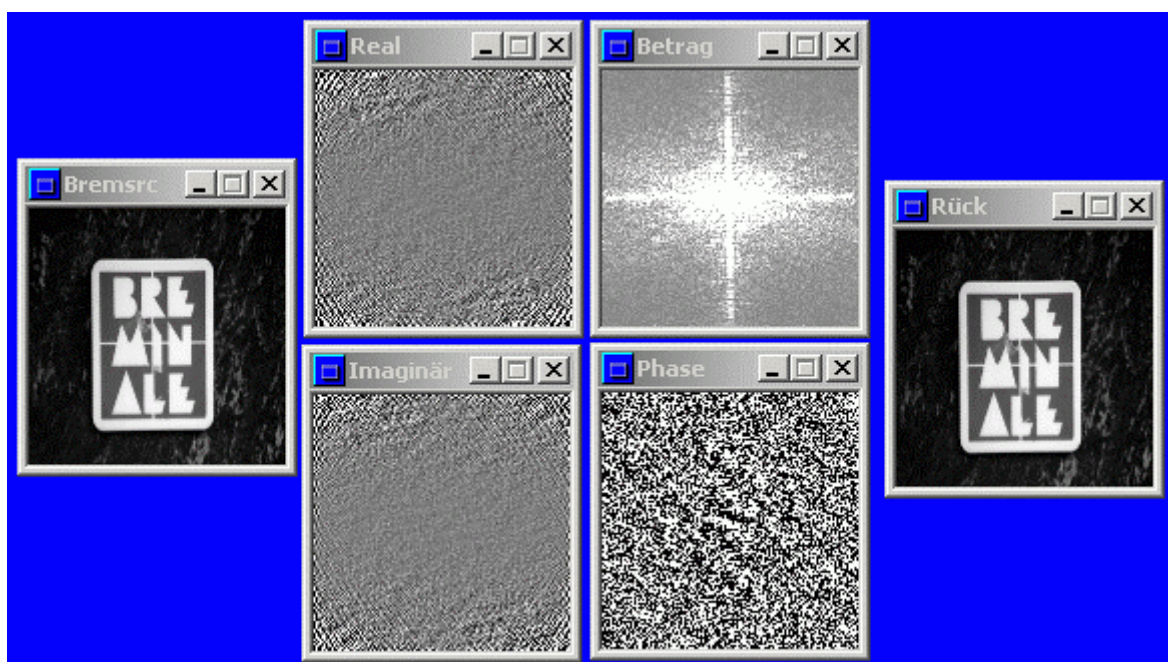
Ergebnis einer globalen Op. im Punkt  $(x, y)$ : abhängig vom gesamten Eingabebild

⇒ 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

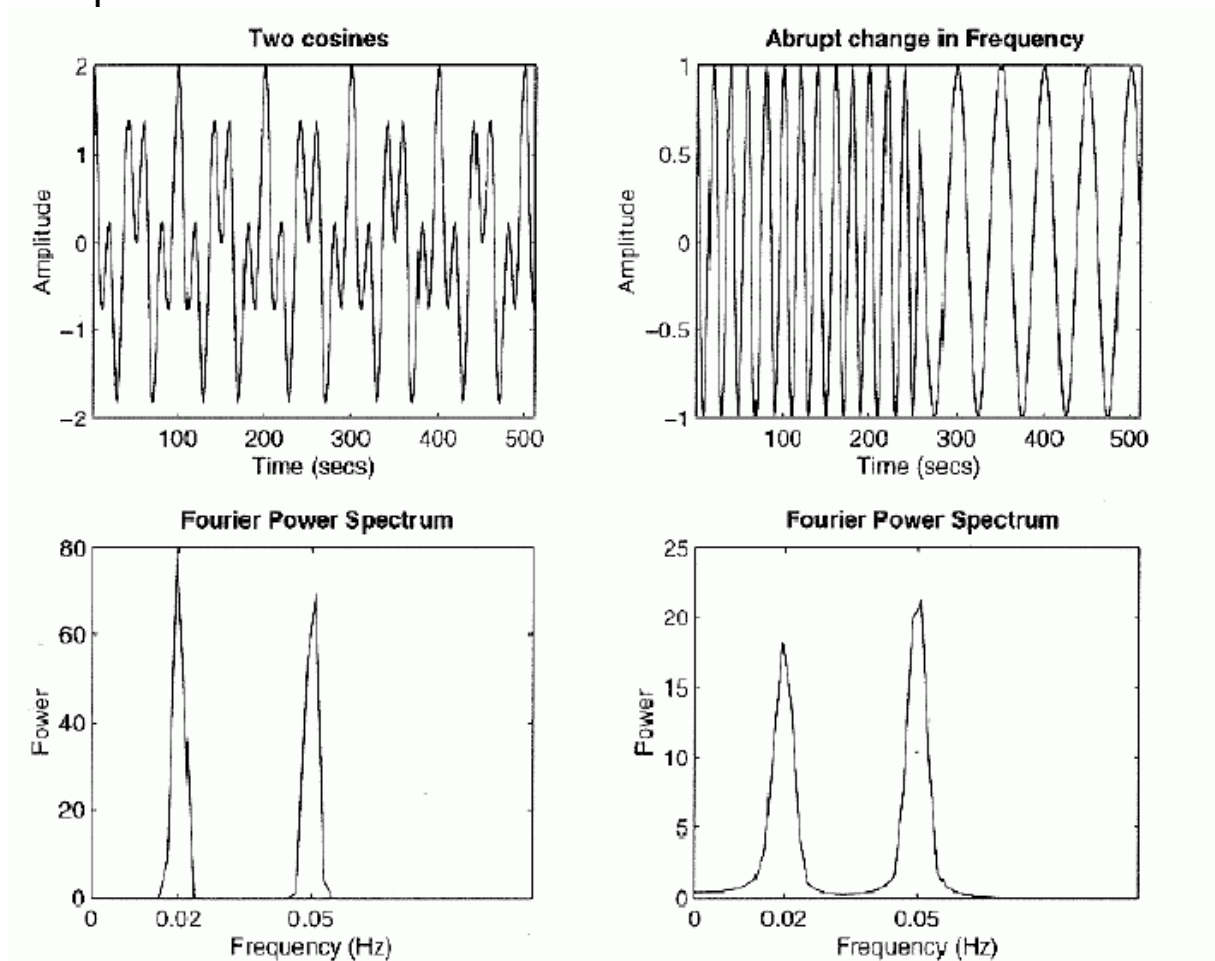
Variante: *Diskrete Kosinus-Transformation* (DCT). Vermeidung des Rechnens im Komplexen durch punktsymmetrische Fortsetzung des Bildes (⇒ Verschwinden der Sinus-Terme), dafür aber größere Matrizen involviert.

- Anwendung in der (klassischen) JPEG-Kompression; siehe Vorlesung Computergrafik, Kap. 4, [http://www-gs.informatik.tu-cottbus.de/~wwwgs/cg\\_v04.htm](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:



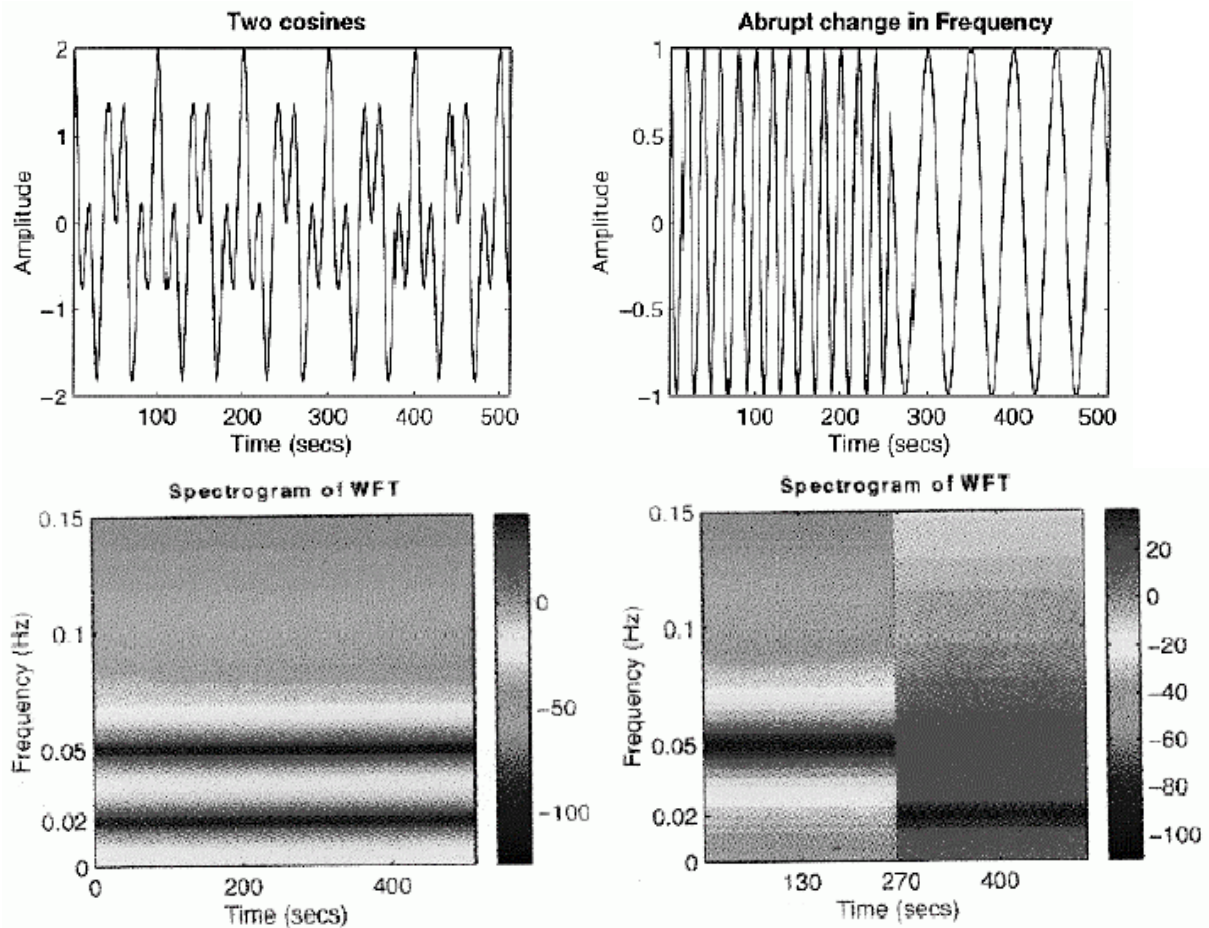
links: Überlagerung zweier Kosinusfunktionen (stationär),  
rechts: Änderung der Frequenz;  
beide Kurvenverläufe führen auf nahezu identische Fourierspektren (aus Esser 2001)

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):

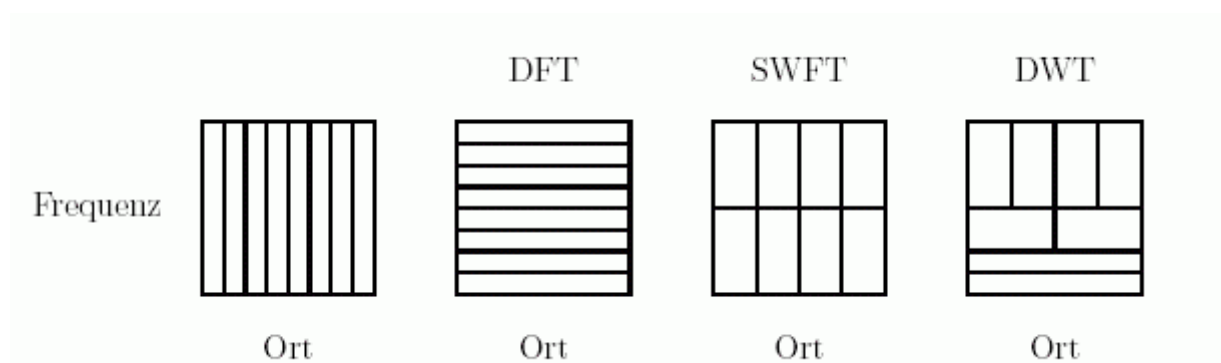
$$g_{\sigma}[n] = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{n^2}{2\sigma^2}}$$

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.