

# Bildverarbeitung Zusammenfassung

## 1. Vorlesung: Einführung

### **Wieso Bildanalyse?**

- Viele Entscheidungen sind schneller, präziser und/oder zuverlässiger zu treffen, wenn sie automatisiert werden
- Problem: was ist die Information, die in den Bildern steckt, und wie kann sie gefunden werden?

### **Lösungen**

- Objektdetektion
- Tracking
- Objekt Identifikation
- Objekt Extraktion

### **Bildanalyse ≠ Bildbearbeitung**

### **Information in Bildern**

Information in Bildern ist nicht leicht zu finden

#### **Beispiel Videokompression**

- Datenmenge DVD Video pro Bild = 720x576 RGB Pixel = ca. 1,2 MB pro Bild
- 24 Bilder/sec, DVD Speicherkapazität ca. 8.5 GB
- unkomprimiert können ca. 7000 Bilder, d.h. 5 Minuten Film gespeichert werden
- um einen Film auf DVD zu speichern, muss ca. 97% der Daten weggeworfen werden (nur wesentliche Information bleibt)

### **Bildinformation: Merkmale**

Merkmale, die den abgebildeten Objekten gehören (Ecken und Kanten, Regionen) → Tracking-Aufgaben

## **Bildinformation: Kontext**

Bilder lassen sich nur im Kontext interpretieren (A-Priori-Wissen)

- Bedeutungs-, Anfrage-, räumlich- und zeitlicher Kontext

Bild- und Kontextinformation müssen redundant sein, um auch bei Störungen oder fehlenden Objektdetails interpretierbar zu sein.

## **Informationsträger**

Es sind Pixelgruppen, die die Informationen tragen (Finden der Pixelgruppen, nicht-Informative Gruppen rauslassen(Rauschen,...))

## **Bildverarbeitung ≠ Informatik**

### **Bildverarbeitungspipeline**

1. Bilderfassung (Kamera, Satelliten)
2. Vorverarbeitung (Rauschunterdrückung, Kontrastverstärkung,...)
3. Segmentierung (Konturen, Flächen,...)
4. Merkmalsextraktion (Kettencode, Umfang, Schwerpunkt, Farbe,...)
5. Klassifikation (Objekterkennung, Einteilung in Klassen,...)

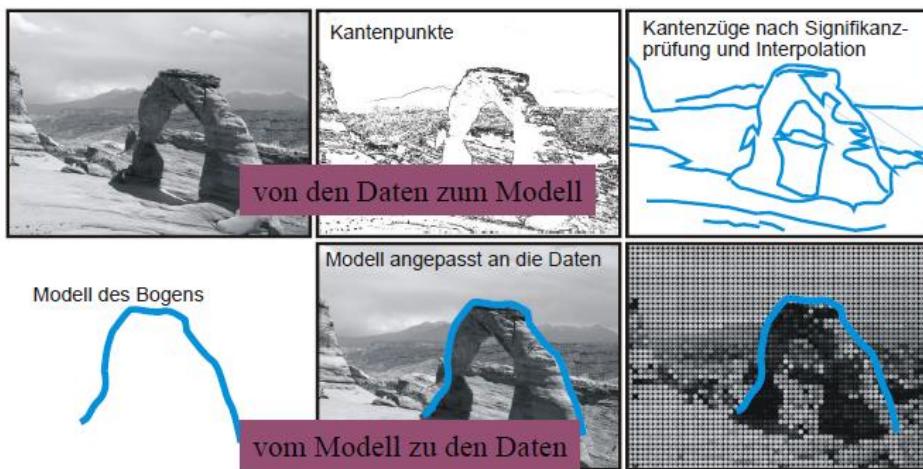
### **Bottom-Up vs. Top-Down**

Bottom-Up (obere Zeile):

- größere, semantische interessante Einheiten (im Beispiel sind dies Kantenpunkte und anschließend die aus signifikanten Kantenpunkten zusammengesetzten Kantenzüge) erzeugen

Top-Down (untere Zeile):

- es existiert zunächst eine Beschreibung des erwarteten Bildinhalts (ein Modell, das im Beispiel die Silhouette des Bogens ist).
- Das Bild wird danach daraufhin untersucht, inwieweit ein solches Modell durch die Daten bestätigt wird.



## Modelle

Modellierung beschreibt Umwandlungen, Störungen und Verluste bei der Abbildung eines Objekts

Prozesse der Bildverarbeitung: Bildrestauration, Bildverbesserung, Segmentierung, Klassifikation

## 2. Vorlesung: Digitale Bilder

### Digitales Foto

Bild besteht aus einer Menge von Bildelementen

Definitionsbereich ist dreidimensional (zwei Ortskoordinaten  $x, y$  sowie die Wellenlänge  $\lambda$ , welche die Farbe spezifiziert).

$$xmin \leq x < xmax, ymin \leq y < ymax, \lambda_{min} \leq \lambda < \lambda_{max}$$

$x_{min}, y_{min}$  untere linke und  $x_{max}, y_{max}$  obere rechte Ecke

$\lambda_{min}$  und  $\lambda_{max}$  wird der wahrnehmbare Farbbereich  
repräsentiert.

Wertebereich: Intensität  $R_{min} \leq R(x, y, \lambda) < R_{max}$

Definition- und Wertebereich sind beschränkt

### Ortsauflösung

Durch die Anzahl der Pixel je Zeile bzw. Spalte ist die Ortsauflösung eines digitalen Bildes festgelegt

$$m = \left\lfloor M \cdot \frac{x - x_{\min}}{x_{\max} - x_{\min}} \right\rfloor \text{ und } n = \left\lfloor N \cdot \frac{y - y_{\min}}{y_{\max} - y_{\min}} \right\rfloor \text{ für } x_{\min} \leq x < x_{\max}, y_{\min} \leq y < y_{\max}$$

## RGB-Farbmodell

Jede Farbe wird durch ein Tripel (rot, grün, blau) dargestellt

Durch die Verwendung von nur drei Farbkanälen kann ein visueller Eindruck erzeugt werden, welcher der Aufnahme eines kontinuierlichen Wellenlängenspektrums durch die farbempfindlichen Rezeptoren (die Zäpfchen) im Auge ähnlich ist.

## Repräsentation eines Bildes

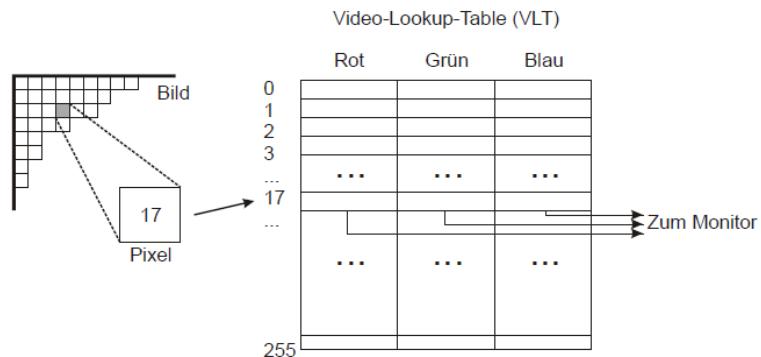
Unterscheidung von (meist) 256 Intensitätsstufen (=1 Byte)

Farbe als

1. Tripel RGB
2. einen Index in einer Hashtabelle  
Video-Lookup-Table (VLT)

## Kontrastauflösung

Quantisierung auf 256 Werte → Anzahl visuell wahrnehmbarer Helligkeitsabstufungen eng begrenzt



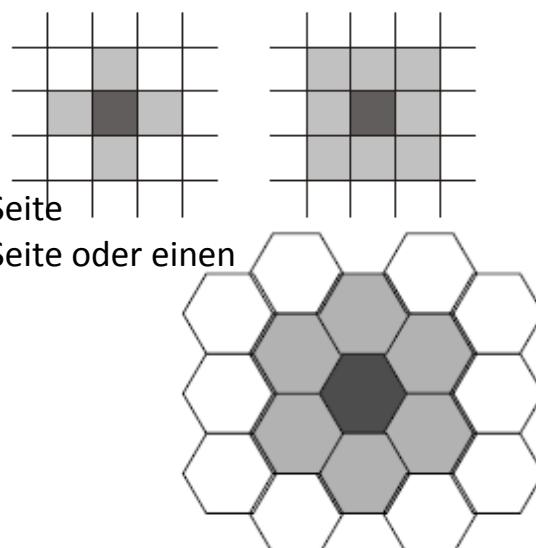
Es gibt Bereiche in den anderen Beschränkungen üblich sind z.B.: Digitalisierung von medizinischen Bildern wird zwar nur ein Spektralkanal genutzt, die Anzahl der Intensitätsstufen ist jedoch höher ( $l_{\max} = 4.096$  bei Röntgen-CT-Bildern).

## Nachbarschaft

Quadratische (rechteckige) Gitter:

- 4-Nachbarschaft: Pixel haben eine gemeinsame Seite
- 8-Nachbarschaft: Pixel haben eine gemeinsame Seite oder einen gemeinsamen Eckpunkt

Hexagonale Gitter:



- 6-Nachbarschaft: Pixel haben eine gemeinsame Seite und zwei gemeinsame Eckpunkte

## Pfad und zusammenhängendes Gebiet

Pfad: Folge von benachbarten Pixeln, die eine Homogenitätsbedingung erfüllen.

- 4-benachbarte Pfade unterscheiden sich von 8-benachbarten Pfaden.

Zusammenhängendes Gebiet: Menge aller Pixel zwischen denen Pfade existieren.

- Die Nachbarschaftsdefinitionen in Vorder- und Hintergrund müssen unterschiedlich sein.

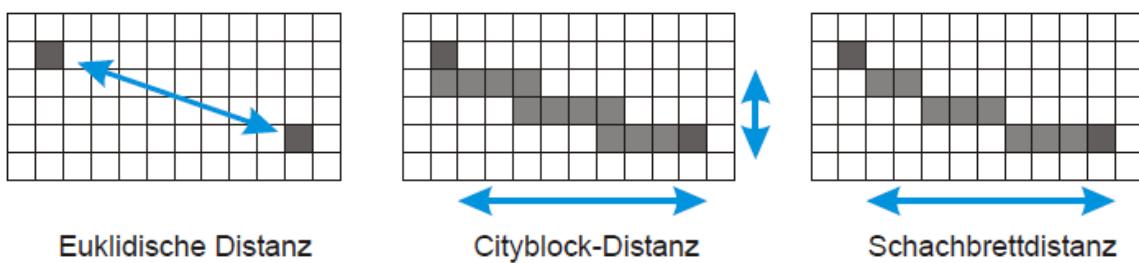
Rand eines zusammenhängenden Gebiets ist eine Folge von Pixeln, die zum Gebiet gehören und zu Pixeln benachbart sind, die nicht dazu gehören.

## Abstandsmaße

$$D_e(x_1, y_1; x_2, y_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad \text{euklidische Distanz}$$

$$D_4(x_1, y_1; x_2, y_2) = |x_1 - x_2| + |y_1 - y_2| \quad \text{Cityblock-Distanz}$$

$$D_8(x_1, y_1; x_2, y_2) = \max(|x_1 - x_2|, |y_1 - y_2|) \quad \text{Schachbrett-Distanz}$$



$D_4$  bzw.  $D_8$  ist die kürzeste Distanz zwischen zwei Pixeln unter 4-und 8-Nachbarschaft

## Gängige Bildformate

Bilder können in unterschiedlich komfortable gespeichert werden

PPM, PGM, PBM

- Keine Datenkompression

- ASCII- oder als Binärdatei

### BMP (Windows Bitmap)

- Unterstützt unterschiedliche Bittiefen von 1,4,8,16,24, oder 32 Bit per Pixel (bpp)
- Gebräuchlich sind 8 Bit (für Grauwertbilder) und 24 (für 3-Kanal-Farbbilder)
- Bilddaten werden durch Run-Length-Encoding komprimiert (VL 7 Bildkompression)

### PNG (Portable Network Graphics)

- Ersatz für GIF, aber ohne Möglichkeit zur Speicherung von Animationen (eigenes Format MNG; wurde nie standardisiert)
- wird von Web-Browsern unterstützt
- Farbpalettenbilder mit 8-Bit-Pixeln
- 1-, 4-, 8-, 16-Bit Grauwertbilder
- 3x8 oder 3x16-Bit-Farbbilder (RGB)
- Zweistufig: Vorfilterung durch Differenzbildung, anschließend ZIP-Kompression (Kombination aus Lempel-Ziv und Huffman-Kodierung)
- Kompressionsrate besser als BMP, auch – durch Vorfilterung – besser als bei GIF

### TIFF (Tagged Image File Format)

- Sehr variables (und komplexes) Format
- Verwendung
- Überall da, wo hochwertige Bilder benötigt werden (z.B. Druckindustrie)

### JPEG (Joint Photographers Expert Group)

- Standardisiertes Format zur effizienten, aber verlustbehafteten Speicherung von Bildern
- Speicherung von Grauwert- und Farbbildern
- Farbbildrepräsentation als YCbCr-Repräsentation (PAL)
- Y: gewichtete Kombination von R,G,B (entspricht etwa dem Grauwertbild)
- Cb, Cr: um Y-Wert korrigierter Blau- bzw. Rot-Anteil im Bild (Farbabweichungen)

### **JPEG Blockkompression**

- Bild wird Block (8x8) für Block komprimiert gespeicher

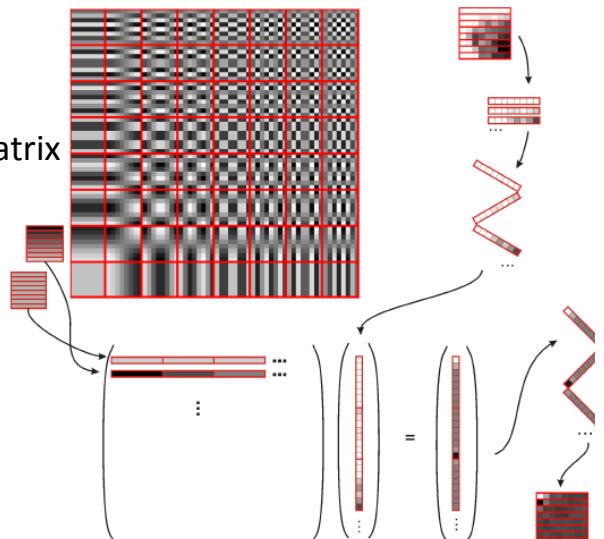
- Kompression des Blocks durch Transformation in effiziente Repräsentation
  - Bildung gewichteter Summen mit sogenannten Basisfunktionen
  - es gibt so viele Basisfunktionen, wie es Pixel in einem Block gibt
  - jede Gewichtung mit einer Basisfunktion ergibt genau einen Wert

## Berechnung

Bild als Vektor

Basisfunktion als Sequenz von Vektoren in Matrix

Operation als Matrixmultiplikation  $A^*x = b$



Eulersche Formel:  $e^{i\varphi} = \cos(\varphi) + i \sin(\varphi)$

## 3.Vorlesung: Fouriertransformation

### Fouriertransformation

- Beschreibung einer beliebigen Funktion als Summe von gewichteten periodischen Funktionen (Basisfunktionen) mit unterschiedlicher Frequenz (Frequenzraumrepräsentation)
- Anwendungen
  - Beschreibung des Informationsverlusts bei Digitalisierung
  - Restauration von linearen Störungen
  - Rekonstruktion von Bildern aus Projektionen
  - Schnelle Filterung

### Basisfunktionen

- Bilder können als zweidimensionale Funktion  $f(m,n)$  aufgefasst werden.
- Jede Funktion kann als Summe von gewichteten Basisfunktionen  $b_{u,v}$  aufgefasst werden:

$$f(m,n) = \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} w_{u,v} \cdot b_{u,v}(m,n)$$

- Die Wichtungen  $w_{u,v}$  bilden eine neue Funktion  $w(u,v)$ , die  $f(m,n)$  zusammen mit den Basisfunktionen genau beschreibt.
- Die Auswirkung von Veränderungen auf die Wichtungen (z.B. Filterung) hängen von den Basisfunktionen ab.
- Die Fouriertransformation ist die Transformation von einer Ortsbasis in eine Frequenzbasis.

## Ortsbasis

- Jede Basisfunktion ist ein Bild, das in einem Pixel den Wert 1 und in allen anderen Pixeln den Wert 0 hat.
- Zwei Basisfunktionen unterscheiden sich dadurch, dass sie den 1-Pixel an unterschiedlichen Stellen haben.
- Vorgeschlagene Basis ist vollständig und eindeutig.

## Invertierbarkeit

- Transformation: „Projektion“ der Funktion auf die neue Basis.
- Inverse Transformation: Projektion auf die alte Basis.

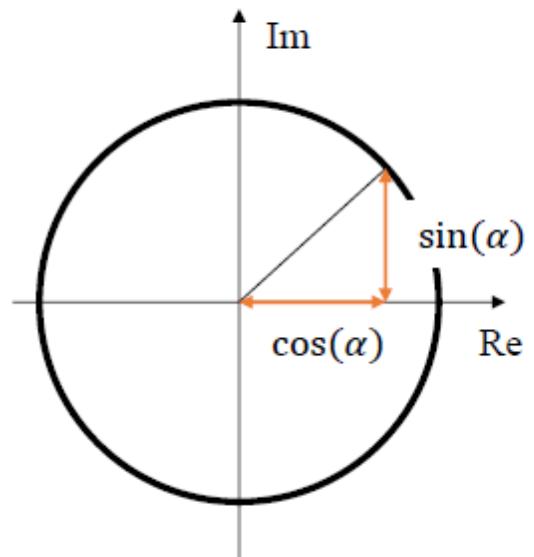
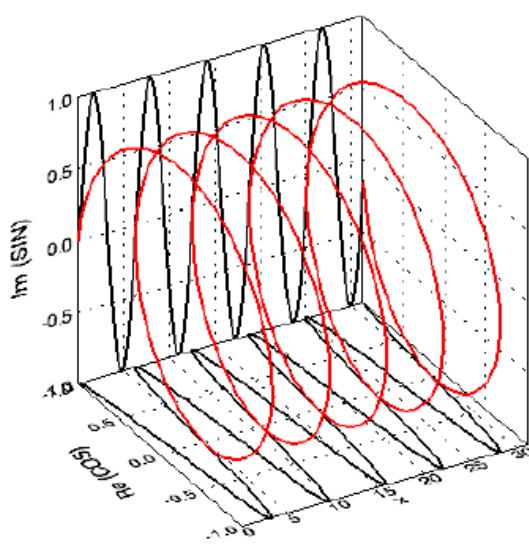
Basisfunktionen	=		Wichtung
	$\times$		$w_{u,v}$
$f(m,n)$	$+$		$x$
	$\dots$		
		$b_{u,v}(m,n)$	$w_{u,v}$

## Orthogonale Vektorbasis

- Zwei Vektoren sind orthogonal, wenn ihr Skalarprodukt Null ist.
- N Vektoren bilden eine Basis für einen N-dimensionalen Raum, wenn sie alle orthogonal zueinander sind.
- Projektion eines Vektors  $p$  auf einen Basisvektor  $b$  ist durch das normierte Skalarprodukt zwischen ihnen gegeben.
- Beispiel:
- Koordinatenvektoren eines dreidimensionalen Raums.

$\cos(\alpha)+i \cdot \sin(\alpha)$

Alle Werte für komplexe Zahlen der Form  $\cos(\alpha)+i \cdot \sin(\alpha)$  liegen auf einem Kreis mit Abstand 1 in der komplexen Ebene.



## Repräsentation als Exponentialfunktion

- Taylorreihenentwicklung für Kosinus und Sinus:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad \sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

- Taylorreihenentwicklung für  $e^{ix}$ :

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \Rightarrow e^{ix} = 1 + \frac{ix}{1!} + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \dots$$

- Es gilt daher wegen  $i^2 = -1$ :  $\cos(x) + i \cdot \sin(x) = e^{ix}$

- Phasenverschiebung  $\alpha$  kann in komplexen Funktionen als Multiplikation ausgedrückt werden:

$$\cos(x + \alpha) + i \sin(x + \alpha) = \exp(i(x + \alpha)) = \exp(i\alpha) \exp(ix)$$

## 1D-Basisfunktion

- Transformation FT :  $FT(f) = F = f \times B$  (Vektor-Matrix-Schreibweise)

$$F(u) = \sum_n f(n) \exp\left(-i \cdot \frac{2\pi}{N} nu\right), \text{ für alle } u=0, N-1$$

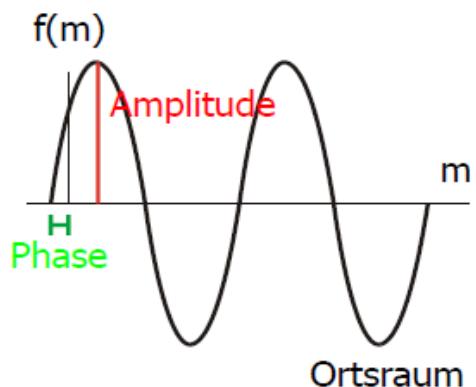
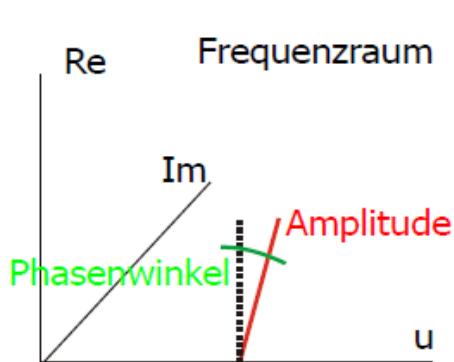
- Rücktransformation  $FT^{-1}$  :  $FT^{-1}(F) = f \times B^T$  (Vektor-Matrix-Schreibweise)

$$f(n) = \frac{1}{N} \sum_u F(u) \exp\left(i \cdot \frac{2\pi}{N} nu\right), \text{ für alle } n=0, N-1$$

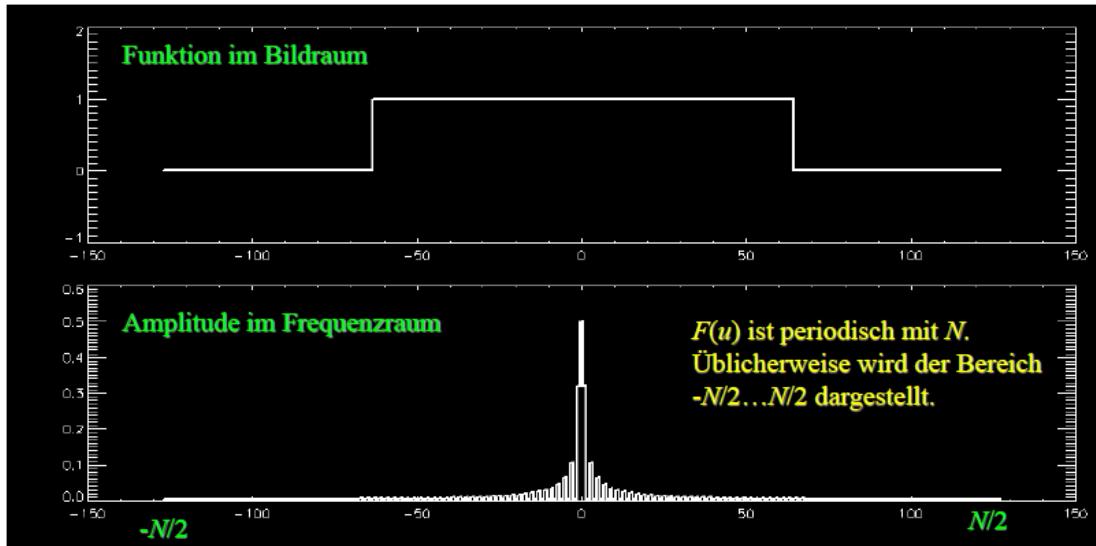
Skalierungsfaktor, weil die Basisfunktionen nicht normiert sind.

## Phase und Amplitude

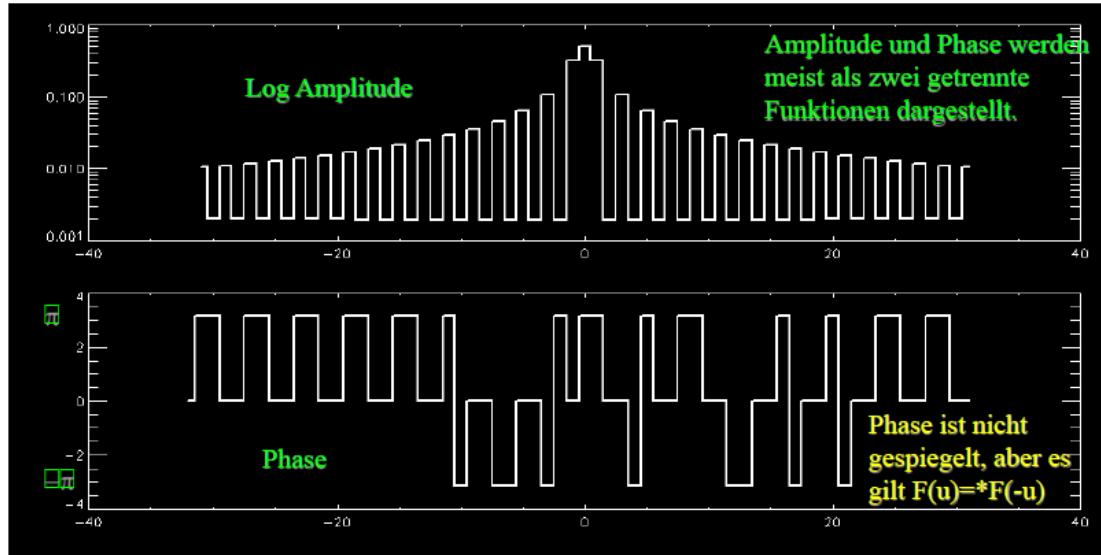
- Das Resultat der Fouriertransformation ist eine komplexe Funktion  $F(u)$ .
- Der Betrag eines Funktionswerts ist die Amplitude und der Winkel zur reellen Achse ist die Phase zur Gewichtung der betreffenden Basisfunktion



## 1D Fouriertransformation



## Amplitude und Phase



## 4. Vorlesung: 2D Fouriertransformation

Die Basisfunktionen der 2-D Fouriertransformation sind zerlegbar:

$$\exp\left(i \cdot \frac{2\pi}{N} \cdot (mu + nv)\right) = \\ \exp\left(i \cdot \frac{2\pi}{N} \cdot mu\right) \cdot \exp\left(i \cdot \frac{2\pi}{N} \cdot nv\right)$$

### 2D Fouriertransformationspaare

Größe MxN:

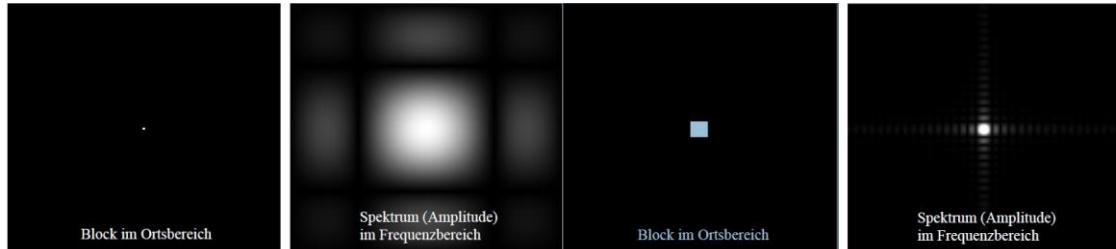
$$F(u, v) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cdot e^{-i \cdot 2\pi \cdot \left(\frac{um}{M} + \frac{vn}{N}\right)}$$

Größe NxN:

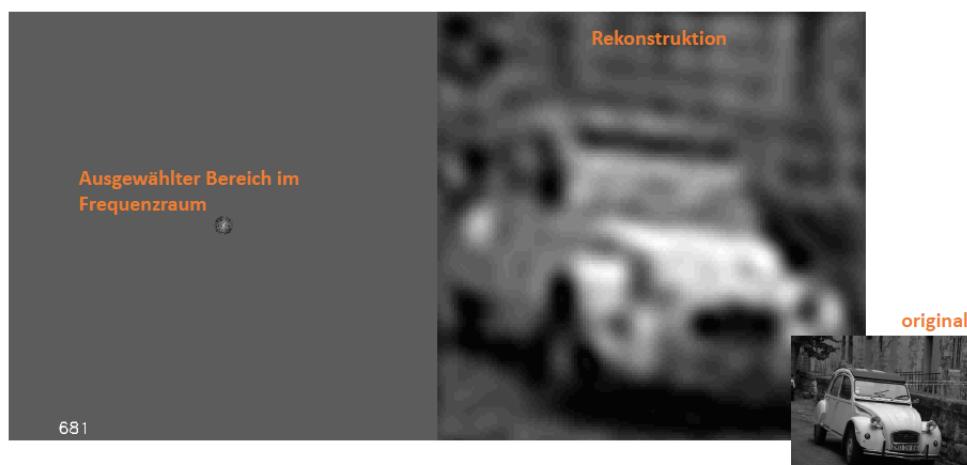
$$F(u, v) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \cdot e^{-i \cdot \frac{2\pi}{N} \cdot (um + vn)} = \sum_{m=0}^{N-1} F(u, v) \cdot e^{i \cdot 2\pi \cdot \left(\frac{um}{M} + \frac{vn}{N}\right)}$$

$$f(m, n) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \cdot e^{i \cdot \frac{2\pi}{N} \cdot (um + vn)}$$

## Bilder im Orts- und Frequenzraum:



## Schrittweise Summation der Komponenten

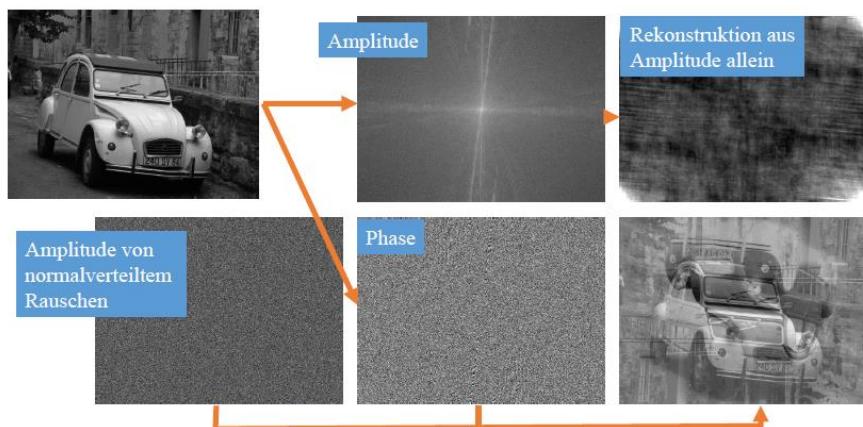
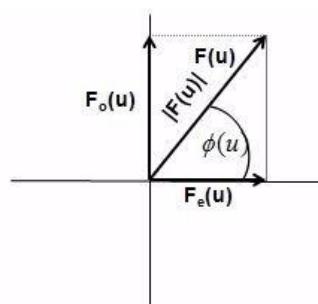


## Phaseninformation wichtig!

Es ist

$$|F(u)| = \sqrt{F_s(u)^2 + F_o(u)^2} \quad \text{Amplitude (Magnitude)}$$

$$\phi(u) = \arctan\left(\frac{F_o(u)}{F_s(u)}\right) \quad \text{Phase}$$



## Translation

Verschiebung einer Funktion  $f$  um  $(dx, dy)$  im Ortsraum führt im Frequenzraum zu einer Veränderung der Phasen, ohne die Amplituden zu verändern. Das ist leicht einzusehen, denn eine Phasenverschiebung entspricht der Verschiebung der durch den Funktionswert repräsentierten Welle.

## Rotation

Eine Rotation der Funktion im Ortsbereich um den Koordinatenursprung mit dem Winkel  $\alpha$  führt zur gleichen Rotation der fouriertransformierten Werte im Frequenzbereich. Auch das ist einleuchtend, weil die das Bild ergebenden Wellen um den gleichen Winkel gedreht werden müssen wie das Bild selbst.

## Mittelwert

Der Mittelwert  $\mu(f)$  einer Funktion ist leicht aus der fouriertransformierten Funktion ableitbar. Wegen

$$F(0,0) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) \exp\left(-i2\pi\left(\frac{0m}{M} + \frac{0n}{N}\right)\right) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)$$

ist  $\mu(f) = F(0,0)/MN$

## Separabilität

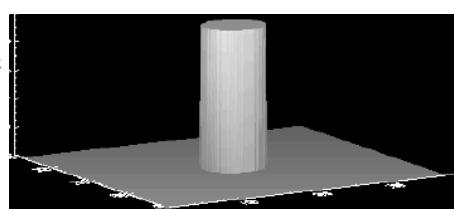
Die Fouriertransformation ist separabel, d.h., sie kann zunächst in M-Richtung und anschließend auf diesen Zwischenergebnissen in N-Richtung ausgeführt werden.

## Filterung

- Filterung im Frequenzraum: Veränderung der Funktionswerte vor der Rücktransformation.
- Motivation: Semantische Trennung von Bildern im Frequenzraum, z.B., niedrige Frequenzen zeigen Intensitäten, hohe zeigen Kanten
- Idealer Tiefpassfilter

$$H_{F_{\max}}(u, v) = \begin{cases} 1 & , \text{falls } u^2 + v^2 \leq F_{\max}^2 \\ 0 & , \text{sonst.} \end{cases}$$

$F_{\max}$  – Cut-Off-Frequenz



## Ringing Artefakte

treten bei Rauschunterdrückung durch Tiefpassfilter vorwiegend an Kanten auf

An Kanten zwischen hellen und dunklen Bereichen werden die hohen Werte mit abnehmender Amplitude weit in den dunklen Bereich und die niedrigen Werte weit in den hellen Bereich hinein verteilt.

## Butterworth-Filter

Frequenzen werden nicht gelöscht, sondern nur abgeschwächt.

$$\bullet \text{ Tiefpass-Filter: } H(u,v) = \frac{1}{1 + (D(u,v)/D_0)^{2n}}$$

$$\bullet \text{ Hochpass-Filter: } H(u,v) = \frac{1}{1 + (D_0 / D(u,v))^{2n}}$$

- $D_0$ : Cutoff-Frequenz
- $D(u,v)$ : Frequenz, d.h. Abstand vom Ursprung

## Tiefpassfilter

Kanten und starke Grauwertänderungen im Bild entsprechen den hochfrequenten Anteilen (mit großem  $u$  und  $v$ ) der Fourier-Transformierten. Bei einem Tiefpass-Filter werden hohe Frequenzanteile abgeschwächt oder ganz eliminiert. Dadurch erscheint das gefilterte Bild im Vergleich zum Originalbild unschärfer, Rauschen und Störungen ähnlicher Art werden dabei aber unterdrückt.

## Hochpassfilter

Bei einem Hochpass-Filter werden tiefe Frequenzanteile abgeschwächt oder ganz eliminiert. Dadurch bleiben Stellen mit abrupten Grauwertänderungen wie Kanten deutlicher erhalten. Wird das Ergebnis eines Hochpass-Filters zum Originalbild hinzugefügt, so erhält man eine Kantenverstärkung und Verschärfung des Bildes.

## Filtern im Ortsraum

Wenn die Fouriertransformation invertierbar ist, dann sollte die Filterung auch im Ortsraum durchführbar sein.

$$\sum_k f(k)h(m-k)$$

## Konvolution (Faltung)

Konvolution erzeugt ein neues Bild  $h$  durch eine gewichtete Summe von Bildelementen in  $g$ :

$$h(m, n) = (g * f)(m, n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} g(i, j) \cdot f(m-i, n-j)$$

Die Konvolution entspricht der Filterung durch Multiplikation im Frequenzraum.

Eigenschaften:

- Operatorzeichen: „\*“ (bedeutet nicht Multiplikation!)
- Linear
- Verschiebungsinvariant
- Kommutativ:  $[g1 * g2](m, n) = [g2 * g1](m, n)$ .
- Assoziativ:  $g1 * ([g2 * g3](m, n)) = [g1 * g2](m, n) * g3(m, n)$

Problem Rand vom Bild:

- Lösung 1: periodische Fortsetzung des Bildes (diese Lösung wird automatisch gewählt, wenn im Frequenzraum multipliziert wird)
- Lösung 2: Rand ist undefiniert
  - Beibehaltung des ursprünglichen Resultats
  - Löschen der Randwerte

## 5. Vorlesung: Abtastung

### Abtastung

Helligkeiten werden nur an einer begrenzten Anzahl von Orten registriert (abgetastet) und Intensitätsunterschiede werden nur erfasst, wenn sie einen bestimmten Mindestunterschied überschreiten.

### CCD-Kamera

Vorteil:

- hohe Lichtempfindlichkeit → geringes Rauschen

Nachteil:

- aufwändige Herstellung, Integration zusätzlicher Elektronik kompliziert
- langsam und viel Energie

### CMOS-Kamera

Vorteile:

- schnell, wenig Energie nötig

Nachteil:

- geringe Lichtempfindlichkeit → hohe Rauschanfälligkeit
- schlechtes Dynamikverhalten

## Zusammenfassung

- Semantik eines Bilds wird zum Teil durch das Aufnahmeverfahren bestimmt
- Ein-Kanal- oder Mehr-Kanal-Bilder
- Bilder sind immer eine Abtastung der (bezüglich des Aufnahmesystems sichtbaren) Szene
- Oberfläche
- Projektion von transparenten/transluzenten Flächen
- Schichtbild

Die Abtastung bildet einen reellen auf einen ganzzahligen Definitionsbereich ab (=Informationsverlust)

## Abtastung (mathematisch Beschrieben)

- Multiplikation der Bildfunktion mit einer Impulsfolge (Folge von Dirac Impulsen  $\delta(x)$ )
- $\delta(x)=0$  für  $x \neq 0$  und  $\int \delta(x)=1$
- Näherungsweise Bestimmung durch Rechteckimpuls:
- Je kleiner die Impulsbreite desto mehr nähert sich das Ergebnis dem tatsächlichen Funktionswert am Abtastort an.
- Die Transformationen Translation, Rotation und Skalierung sind auf reellen Zahlen definiert

## Transformation

- Translation, Rotation und Skalierung auf reelle Zahlen
- Bilder haben ganzzahligen Definitionsbereich
- nach Transformation eines bereits abgetasteten Bildes Interpolation notwendig

## Interpolation

- konstante Interpolation
- lineare Interpolation
- Interpolation durch Polynome höheren Grades
  - Die Bildfunktion wird besser angenähert, wenn mehr Terme der Taylor-Approximation berücksichtigt werden.

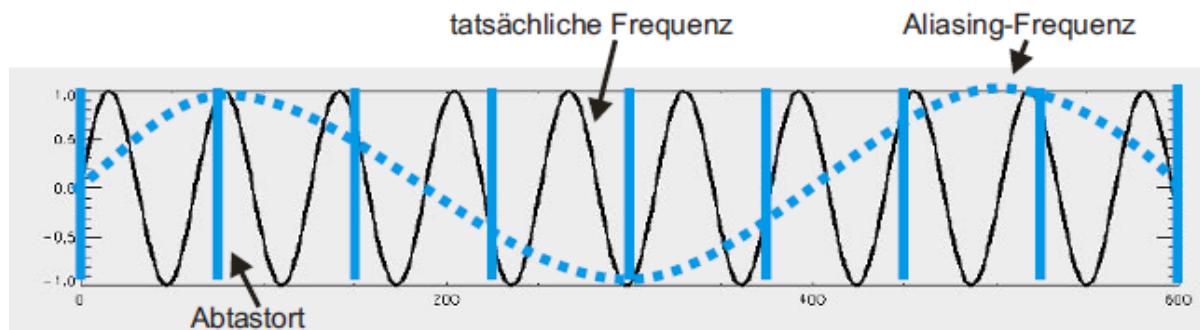
- Ableitungen für die Taylor-Reihe werden durch Differenzen angenähert.
- Grad des Polynoms ist ein Kompromiss zwischen
  - steigender Anzahl berücksichtiger Terme der Taylor-Reihe.
  - steigender Ungenauigkeit der geschätzten Ableitungen.
- Interpolation im Frequenzraum
  - Die Basisfunktionen der Fouriertransformation haben einen reellen Definitionsbereich.

## Moiré-Effekt

- wenn im abgetastetem Bild Information reduziert und verfälscht wurden
- zu große Abtastintervalle in Abhängigkeit zu der Frequenz des aufgenommenen Musters

## Aliasing

- Nyquist-Theorem besagt, dass ein Audiosignal mindestens mit der doppelten Frequenz gleichmäßig abgetastet werden muss, um das Ursprungssignal hinreichend rekonstruieren zu können
- Beim Alias-Effekt werden Frequenzen, die oberhalb der halben Samplingfrequenz (auch Nyquist-Frequenz genannt) liegen, als niedrigere Frequenzen interpretiert
- Grund: Abtastrate ist nicht doppelt so hoch wie die höchste übertragende Frequenz



zwei unterschiedliche Ansätze Aliasing zu vermeiden:

1. Erhöhung der Abtastrate können Überlappungen von Kopien im Frequenzraum vermieden werden, falls sie bandbegrenzt (höhere Frequenzen als eine vorgegebene Grenzfrequenz kommen nicht vor Tiefpass-/Hochpassfilter) ist → Shannon'sche Abtasttheorie
2. unregelmäßige Abtastabstände verhindern Wahrnehmung von periodischem Signal, zweistufige Abtasthierarchie (unterhalb eines Abstandes, oberhalb dieses Abstandes ist die Summe der Abtastort



gleich) → oberhalb der Grenzfrequenz periodisches Signal, unterhalb als Rauschen wahrgenommen

### Shannon'sches Abtasttheorem

falls aufgenommene Funktion bandlimitiert auf Frequenz  $< W$  ist und gilt:

$$2W \leq 1/\Delta x \Leftrightarrow \Delta x = 1/2W$$

Menschliche Wahrnehmung ist nicht Aliasing empfindlich

### 6. Vorlesung: Bildrestauration

Wiederherstellung eines ungestörten Bildes aus einem gestörten Bild

#### Störoperator

- Störung wird als linearer Operator A beschrieben
  - Modell:  $A^*f = g$
  - falls A bekannt und invertierbar, ungestörte Bild  $f = A^{-1}g$
- Verschiebungsinvarianz
  - leicht regularisierbar
  - verschiebungsinvarianter Operator wirkt im gesamten Bild gleich 
$$A \circ f(x+a, y+b) = [A \circ f](x+a, y+b)$$

#### Bewegungsunschärfe

- Faltungskern ist eine Funktion w mit
 
$$w(t * \cos \alpha, t * \sin \alpha) = \begin{cases} \frac{1}{\Delta s}, & \left| \frac{t}{2} \right| < \Delta s \\ 0, & \text{sonst} \end{cases}$$
- der Winkel  $\alpha$  gibt die Bewegungsrichtung an
- die Strecke  $\Delta s$  gibt die Strecke an, um die sich der Punkt bewegt hat

#### Repräsentation linearer Störungen

- Jede verschiebungsinvariante, lineare Operation wird vollständig durch die Faltungsfunktion beschrieben
- Die Faltungsfunktion beschreibt die Operation für beliebige Bilder
- Die Faltungsfunktion kann als Resultat der Veränderung eines Punkts erzeugt werden
- Punktantwort = Point Spread Function (PSF)

#### Bestimmung einer unbekannten PSF

- Annahme: Störung ist unveränderlich Testaufnahme ist möglich, Aufnahme ist eine Näherung für die PSF
- Wenn Testaufnahme nicht möglich: Näherungsweise Bestimmung der PSF durch Betrachtung von Punkten und Linien im gestörten Bild

## 1D-Kanten

- Stärke einer Kante hängt von der Steigung der Funktion ab:
  - Betrag der ersten Ableitung
- für diskrete Funktionen:
  - Ableitung wird durch Differenz gebildet
  - $f'(x) = f(n+1) - f(n)$

## Kanten im 2D-Raum

- Gradient im kontinuierlichen Raum: Vektor der partiellen Ableitung der Bildfunktion  $\vec{G}(f(x, y)) = \left( \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \right)$
- Richtung: Richtung der größten Steigung:  $\tan^{-1}(G_y/G_x)$
- Länge: Stärke der stärksten Steigung: Betrag  $\sqrt{(G_x^2 + G_y^2)}$
- Approximation des Gradienten: Differential wird durch Differenz approximiert (Rückwärts-, Zentral- und Vorwärtsdifferenz)
- PSF kann auch aus dem Verlauf einer ideal angenommenen Kante gewonnen werden

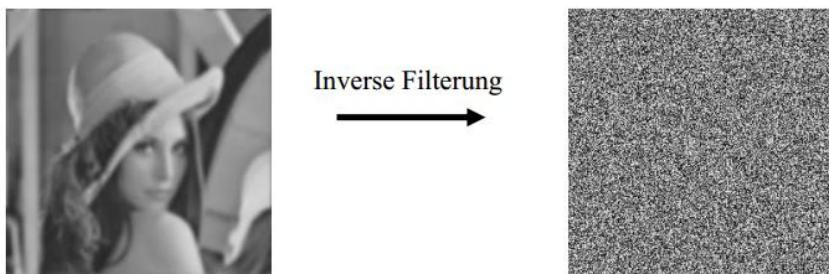
## Invertierung der Störung

- Überführung der Repräsentation in den Frequenzraum  

$$G(u, v) = \mathbf{FT}[g(m, n)] = \mathbf{FT}[[h * f](m, n)] = H(u, v) \cdot F(u, v)$$
- Invertierung:  

$$f(m, n) = \mathbf{FT}^{-1}[G(u, v)/H(u, v)]$$
 (**Inverse Filterung**)

## Rauschen als Störeinfluss



## Bildrauschen

- nicht-deterministisch, nicht wiederholbar
- Beschreibbar als Wahrscheinlichkeit (Impulsrauschen, Quantenrauschen)

## Signal-Rausch-Verhältnis

- Charakterisierung des Grads des Rauschens
- Signal: Unterschied zwischen Objekt und Hintergrund. Falls das Objekt nicht bekannt ist:
  - maximaler Grauwert
  - durchschnittlicher Grauwert
- Rauschen ist über Standardabweichung gegeben

$$SNR_{\max}(f) = \frac{f_{\max}}{\sigma}$$

## Wiener-Filter

- soll Einfluss von Rauschen im Frequenzraum minimieren, wird daher meistens bei schlechtem SNR angewendet

$$H(u,v)^{-1} \frac{|H(u,v)|^2}{|H(u,v)|^2 + \gamma \left( \frac{S_\eta(u,v)}{S_f(u,v)} \right)}$$

- $S_\eta(u, v)$  - Spektrum des Rauschens
- $S_f(u, v)$  - Spektrum des ungestörten und verrauschten Bildes
- $\gamma$  - Steuerungsparameter für die Dämpfung  
( $\gamma = 1$ : Wiener Filter,  $\gamma \neq 1$ : parametrisches Wiener Filter)
- Heuristische Variante (falls Spektrum des Rauschens unbekannt ist)

$$H(u,v)^{-1} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}$$

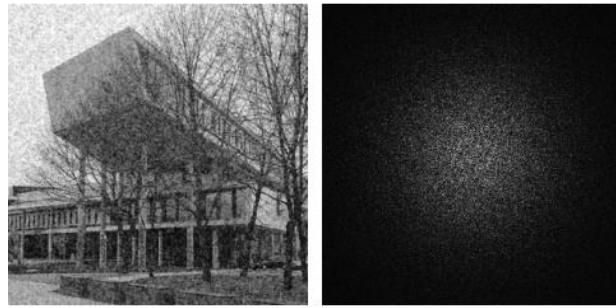
## Impulsrauschen

- nur wenige Pixel betroffen (ca.5%)
- betroffene Pixel nehmen entweder den dunkelsten Grauwert (negativer Impuls) oder den hellsten Grauwert an (positiver Impuls)
- Unterdrückung durch: Medianfilter (Rangordnungsfilter)

## Quantenrauschen

- Quanteneffekte sorgen dafür, dass sich Photonen nicht geradlinig ausbreiten, sondern Anteile der Photonen gar nicht oder nicht an der richtigen Stelle auftreten
- Wahrscheinlichkeit poissonverteilt, lässt sich für eine hohe Anzahl an Photonen der Gauß-Verteilung annähern
- Bild = Kombination von Helligkeitsanteilen aus sich regulär verhaltenen Photonen (Signal) und Rauschen:  $g(m,n) = f(m,n) + \eta(m,n)$

### **Farbiges Rauschen**



- im Frequenzraum wenn z.B. hohe Frequenzen verrauscht sind in der Mitte heller Fleck
- enthält unterschiedliche Frequenzen mit verschieden hoher Amplituden

### **Weißes Rauschen = gleichverteilt**

- Variation der Rauschabweichung zwischen Pixeln völlig zufällig
- Erwartungswert für die Amplitude der FT unabhängig der Frequenz
- Frequenz eines Farbspektrums aufgefasst wäre das Resultat weiß (jede Frequenz tritt mit der gleichen Amplitude auf)
- Restauration: Inverse Filterung mit Verhältnis zwischen Amplitude von H und durchschnittlicher Amplitude zu gewichten  $\frac{1}{H(u,v)} * \frac{|H(u,v)|}{A_N}$

## 7. Vorlesung: Bildkompression

- Bilddaten sind ein Mittel, um Information zu vermitteln.
- Information ist in der Regel redundant kodiert.
- Reduktion von redundanter Daten
- Codierungsredundanz (Anzahl der Bits zu Grauwertkodierung).
  - Interpixelredundanz (Anzahl der Datenpunkte pro Pixel).
  - Psychovisuelle Redundanz (zur Erkennung benötigt Bildinformation)
- Kompressionsrate:  $C = n_1 / n_2$ 
  - $n_1$  und  $n_2$  sind die Anzahl der Informationseinheiten, um dieselbe Information zu kodieren.

### **Codierungsredundanz**

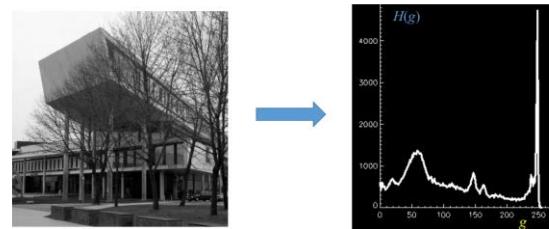
- Anzahl der verfügbaren Codes ist größer als die der benötigten Codes.
- Beispiel (jedes Pixel = ein Byte)
  - nur Grauwerte 1 bis 100 sind im Bild vorhanden.
  - 90% der Pixel haben den Grauwert 100.
- Lösung: 1-Bit-Code für 100, 9-Bit-Code für alle anderen
  - (100): 0, (1): 100000001, (2): 100000010, (3): 100000011, (4): 100000100, ...
  - (100) (100) (100) (17) (15) (100)...: 0001000010011000011110...
  - Kompression:  $8/(0.9*1+0.1*9)=8/1.8\approx4.4$
- Ziele

- Berechnung der Redundanz (Berechnung des Informationsgehalts)
- Reduzierung der Codelänge.

## Histogramm

Häufigkeit der Grauwerte in einem Bild

## Normiertes Histogramm



- Normierung nach Anzahl der Pixel eines Bildes (Größe  $M \times N$ ):  
 $H_{\text{norm}}(g) = H(g) / (M \cdot N)$
- ein normiertes Histogramm gibt für jeden Grauwert  $g$  die W-keit an, ob ein beliebiges Pixel den Grauwert hat

## Informationsgehalt

messbare Einheit von Informationen mit intuitiver Bedeutung

Informationsgehalt  $IG(E)$  eines Grauwerts  $E$  ist umso höher, je größer die Gesamtanzahl  $M$  der verwendeten Grauwerte ist:

- $IG_M(E) = M$ .
- Informationsgehalt ist unabhängig davon, welcher Grauwert aus der Liste  $E = \{E_0, E_1, \dots, E_{M-1}\}$  übermittelt wurde.
- Informationseinheiten  $I_M(E) = \text{Anzahl der benötigten } n\text{-wertigen Symbole, für die Speicherung des Informationsgehalts, also } I_M(E) = \log_n IG_M(E)$
- Beispiel:
  - Anzahl der Grauwerte: 256
  - Informationsgehalt jedes Grauwerts: 256
  - Symbol: Bit (2-wertig)
  - Benötigte Informationseinheiten:  $\log_2 256 = 8$
- Nachteil: Informationsgehalt eines häufig vorkommenden Grauwerts ist genauso groß wie die eines seltenen vorkommenden Werts
- Informationsgehalt  $IG(E)$  eines Pixelwerts  $E$  unter Berücksichtigung der Häufigkeit von  $E$ :
- Umgekehrt proportional zur Wahrscheinlichkeit  $P(E)$  des Eintreffens  $IG(E) = 1/P(E)$ .
- Anzahl der benötigten Informationseinheiten ist dann  $I(E) = \log_n 1/P(E) = -\log_n P(E)$
- Zur Repräsentation der Information  $IG(E)$  werden dann  $I(E)$  Informationseinheiten benötigt.
- Beispiel für Schwarzweißbilder mit gleicher Anzahl schwarzer und weißer Pixel:

- Wahrscheinlichkeit für Eintreffen von E ist 0.5
- Informationsgehalt ist dann  $I(E) = -\log_2 0.5 = \log_2 2 = 1$

**Entropie**       $E(H) = -\sum_{i=0}^{K-1} H_p(i) \log_2 H_p(i)$

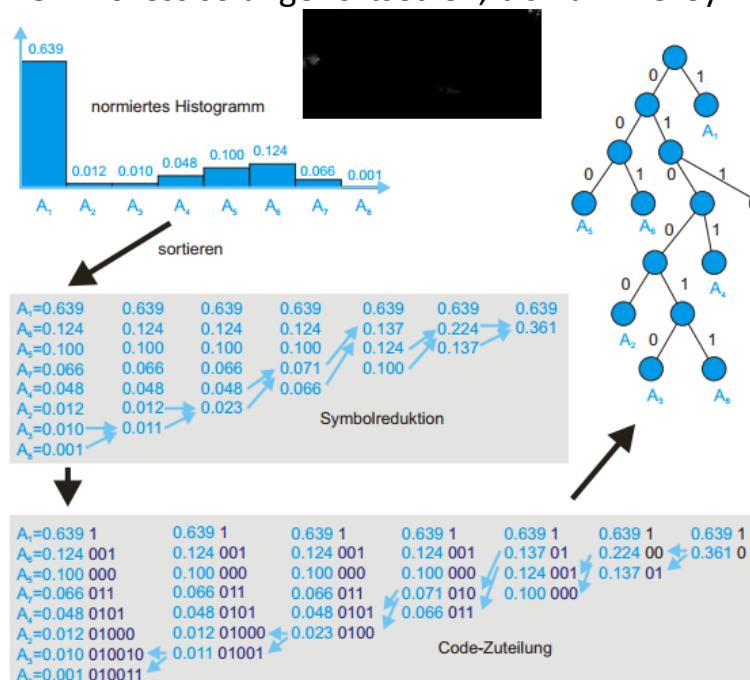
Entropie wird aus dem normierten Histogramm berechnet.

Je gleichmäßiger die Intensitätswerte verteilt sind, desto höher ist die Entropie

Mit der Entropie als durchschnittlicher Anzahl der benötigten Bits für eine Folge von Pixeln haben wir ein Maß für den maximal erzielbaren Kompressionsfaktor durch Reduktion von Signalredundanz.

### Huffman-Codierung

- Für ein Bild f mit der Entropie E, in dem alle Grauwerte gleich häufig vorkommen, werden E Bits pro Pixel benötigt, um f redundanzfrei zu repräsentieren
- ungleichmäßig Verteilung, dann für optimale Kompression jeder Pixelgrauwert g mit  $-\log_2 H_p(g)$  Einheiten repräsentiert werden
- Logarithmus der relativen Häufigkeit selten eine ganze Zahl, so dass nächsthöhere Zahl gewählt werden muss.
- 8 Bit pro Grauwert im Bild  $\rightarrow 1/8$  zu hohen Wert wählen erreichbare Kompressionsrate kann weit verfehlt werden
- Berechne Häufigkeitsverteilung (normiertes Histogramm)
- Eingangssymbole nach Häufigkeit ordnen
- Die zwei seltensten Symbole zu einem kombinierten Symbol zusammensetzen.
- Den Prozess solange fortsetzen, bis nur zwei Symbole existieren.



## Interpixel-Redundanz

- Minderung der Kodierungsredundanz berücksichtigt die Homogenität innerhalb des Bilds nicht
- Nachbarpixel sind nicht unabhängig voneinander
- Interpixel-Redundanz Repräsentation durch Blockspezifikation  
→ Run-Length-Encoding (RLE)



## Run-Length-Encoding

- Bild wird zeilenweise kodiert  $(10, \square), (4, \blacksquare)$
- Eine Zeile mit Grauwertfolge  $g(0), g(2), \dots, g(n)$  wird zerlegt in Zeilenstücke mit gleichem Grauwert (die runs):
- Die Zeilenstücke werden durch Angabe von Länge und Grauwert codiert:  
 $(r_1, g(r_1)), (r_2, g(r_2)), \dots$
- Kompressionsrate für jede Zeile ist  $n/(2 \cdot rd)$ , wobei  $rd$  die durchschnittliche Länge des runs ist.

## RLE auf Bit-Ebene

- RLE ist ineffizient bei verrauschten Bildern (sehr viele kleine Grauwertänderungen = kurze Runs).
- Lösung:
  - Zerlegung in Bit-Ebene
  - Runs Separat auf jede Bit-Ebene
- Problem:
  - Je niedriger die Position des Bits ist, desto kürzer wird allerdings die durchschnittliche Lauflänge auf dieser BitEbene sein
  - den untersten Bit-Ebenen ist die Verteilung fast zufällig, so dass Lauflängen von einem Pixel die Regel sein werden.  
→ hybride Repräsentation bezahlt, nur die höheren Bit-Ebenen nach Lauflängen codiert

## Gray-Code

- Code bei dem sich aufeinanderfolgende Zahlen höchstens um 1Bit unterscheiden
- Generierung für Bitfolge  $b_1, b_2, \dots, b_n$  (höchstwertiges Bit hat niedrigsten Index):
- Beispiel:  
 $b: 000\ 001\ 010\ 011\ 100\ 101\ 110\ 111$   
 $g: 000\ 001\ 011\ 010\ 110\ 111\ 101\ 100$

## Reduktion der Kontrastauflösung

- menschlich unterscheidbare Helligkeitsstufen (30-100) geringer als die repräsentierbare Anzahl
- Bildinformation ist auch noch bei kleinerer Anzahl wahrnehmbar
- Machbandeffekt (Signal wird an Kanten künstlich überhöht wahrgenommen) kann durch Addition von Rauschen vermieden werden

## Reduktion hochfrequenter Anteile

- sehr viel Bildinformation ist in den niedrigen Frequenzen codiert
- Kompression:
  - Transformation
  - Sortierung der Koeffizienten nach Größe
  - Quantisierung, d.h. Abbildung auf ganze Zahlen:

$$q(e) = Q \cdot \text{sign}(e) \cdot (|e| - e_{\min}) / (e_{\max} - e_{\min})$$

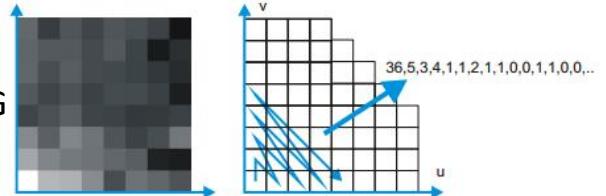
- Kompression der quantisierten Werte durch verlustfreies Verfahren

Dekodierung durch Rücktransformation

## Diskrete Kosinus Transformation (DCT)

- Zerlegung in Wellen unterschiedlicher Frequenzen
- sehr ähnlich zur FFT
- Basisfunktionen sind reell
- Nutzung: Kompressionsverfahren bei JPEG

## DCT vs DFT



- DCT hat nur reelle Komponenten, deren Wert mit Abstand vom Ursprung abnimmt.
- Kompression durch Quantisierung:
  - Werte entlang der Diagonale auslesen,
  - Quantisieren,
  - nur bis zum letzten von Null verschiedenen Wert speichern.

## Zerlegung in Teilbilder

- Anzahl der Informationstragenden Koeffizienten ist geringer, je geringer die Anzahl der Kanten ist
- Zerlegung in Teilbilder
  - Anzahl der Funktionswerte im Frequenzraum entspricht der Anzahl der Pixel des Teilbilds.
  - Wahrscheinlichkeit von Kanten in Teilbildern sinkt mit deren Größe.

- Teilbildgröße: Kompromiss zwischen durchschnittlicher Anzahl der Information tragenden Frequenzen und Gesamtanzahl der Teilbilder.

## **JPEG Kompression von Teilbildern**

- Farbraumtransformation von RGB nach YCrCb
- Tiefpassfilterung und Unter-Abtastung der Farbabweichungssignale Cr und Cb
- Einteilung von 8x8 Blöcke und DCT auf dem Y-Kanal
- Quantisierung der DCT-Koeffizienten
- Umsortierung und Entropiekodierung (Huffman)
- Vorteil: Kompression nach Blockzerlegung ist effizient und ermöglicht hohe Kompressionsraten
- Nachteil: Blockgrenzen sind willkürlich gezogen, Frequenzen die größer als die Blockgröße sind, werden nicht erkannt → Blockartefakte

## **JPEG2000: Wavelet-Koeffizienten**

- Wavelet-Transformation zerlegt das Bild in lokal wirkende Wellen unterschiedlicher Frequenz
- Keine explizite Blockung notwendig
- Kompression wird umso besser sein, umso kompakter die Wavelets in Orts- und Frequenzraum sind

## **8.Vorlesung: Kontrastverbesserung**

### **Pixelbasierte Bildverbesserung**

- Abbildungen von Grau-/Farbwerten unabhängig von ihrem Ort oder ihrer Zuordnung (innerhalb der Grauwerte oder von Grauwert zu Farbwert(Falschfarbdarstellung))
- Figure of Merit (Qualitätsmerkmal): globaler/lokaler Kontrast
- Methoden:
  - Monotone Abbildung der Grauwerte
  - nicht monotone Abbildung
  - Falschfarbdarstellung

### **Kontrast**

- Globaler Kontrast

$$c_{global}(f) = \frac{[max_{m,n}(f(m,n)) - min_{m,n}(f(m,n))]}{g_{range}},$$

mit  $g_{range}$  – Grauwertbereich

- Lokaler Kontrast: z.B. durchschnittlicher Grauwertunterschied zwischen zwei benachbarten Pixeln

$$c_{local}(f) = 1/MN \sum_m \sum_n |f(m, n) - f_{nb}(m, n)|,$$

mit  $f_{nb}(m, n)$  – durchschnittlicher Grauwert in der Umgebung von  $(m, n)$ .

- Grauwertspreizung: Jedem Pixel mit einem Grauwert  $g$  wird über eine Transferfunktion  $g'$  ein von  $[g_{min}, g_{max}]$  abhängiger neuer Grauwert zugewiesen:

$$g'(g) = (g - g_{min}) \cdot \frac{w_{max} - w_{min}}{g_{max} - g_{min}} + w_{min}.$$

$w_{min}$ ,  $w_{max}$  ist der kleinste/größte repräsentierbare Grauwert

### Verbesserung des lokalen Kontrasts

- Bild ist zu helle/dunkel, aber Grauwertbereich nahezu ausgenutzt
- nichtlineare, monotone Transferfunktion.: z.B. Gammakorrektur

$$g'(g) = w_{max} \cdot \left( \frac{g}{w_{max}} \right)^\gamma$$

- $\gamma < 1$  helleren Grauwerte gespreizt, dunkleren gestaucht
- $\gamma > 1$  helleren Grauwerte gestaucht, dunklere gespreizt

### Maximaler Informationsgehalt

- Entropie ist maximal, wenn  $P(g_i) = \text{const}$  für  $i=0, \dots, N-1$
- gesucht Histogrammtransformation zur Maximierung der Entropie
- Annahme: Grauwerte sind kontinuierlich auf einem Bereich von 0 bis 1 definiert → für normiertes & kontinuierliches Histogramm  $g'(g) = \int_0^g H(w) dw$
- Problem: wenn  $g'(g)^*(N-1)$  keine ganze Zahl

### Histogrammlinearisierung

- Transferfunktion für ein diskretes Histogramm

$$E(H(g)) = \lceil N_g \sum_{w=0, g-1} H(w) \rceil - 1$$

- $N_g$  = Anzahl der Grauwerte

Grauwert	0	1	2	3	4	5	6	7
Häufigkeit	50	150	350	250	100	60	30	10
$H(g)$	0.05	0.15	0.35	0.25	0.10	0.06	0.03	0.01
kumulativ	0.05	0.20	0.55	0.80	0.90	0.96	0.99	1.00
Grauwert	0.4	1.6	4.4	6.4	7.2	7.68	7.92	8.00
aufgerundet	1	2	5	7	8	8	8	8
$\lceil \cdot \rceil$	0	1	4	6	7	7	7	7

### Adaptive Histogrammlinearisierung (AHE)

- Histogramm wird für jeden Punkt für eine vorgegebene Umgebung erstellt
- Linearisierung nach diesem Histogramm
- nur Grauwert des betreffenden Punktes wird modifiziert
- Nachteil: extreme Anhebung des Kontrast von homogenen Bildbereichen

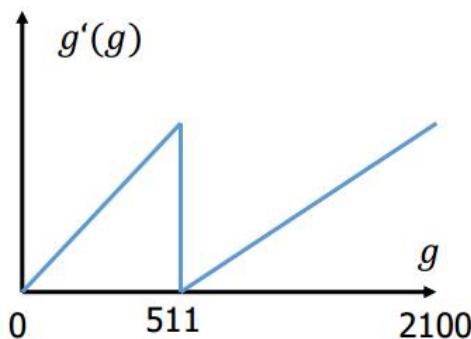
### Kontrast limitierte adaptive Histogrammlinearisierung (CLAHE)

- wie AHE, aber Kontrastverstärkung nur bis zu einem gewissen Maximum
- verhindert die bei der AHE vorkommende Kontrastverstärkung im Bildhintergrund



### Nichtmonotone Grauwertabbildung

- erzeugt künstliche Kanten
- Zuordnung von Helligkeit und Grauwert nicht mehr erkennbar  
Zwei Grauwertfenster in einem Bild.



### Farbe zur Kontrastverstärkung

- es können mehr Farbwerte unterschieden werden als Grauwerte
- Kontrastverstärkung durch drei nicht-lineare, nicht-monotone Abbildungsfunktionen der Grauwerte:  $f_{\text{rot}}(g)$ ,  $f_{\text{grün}}(g)$ ,  $f_{\text{blau}}(g)$
- Zusammenhang zwischen Farbe und Grauwert nicht mehr erkennbar

## 9.Vorlesung: Tiefpassfilter & Kantenhervorhebung

## Rauschen als stochastischer Prozess: Zeitliche Folge

- Annahme
  - Aufnahme mehrerer Bilder über einen gegebenen Zeitraum
  - keine Bewegung, keine Beleuchtungsänderung → Bild ändert sich nicht
  - Erwartungswert E des Rauschen n ist 0

$$\begin{aligned} E\{g(m, n)\} &= E\{f(m, n)\} + E\{n(m, n)\} \\ &= E\{f(m, n)\} + 0 = f(m, n) \end{aligned}$$

## Integration über die Fläche

- Falls für Punkte ( $p_0, \dots, p_n$ )  $f(p_i) = \text{const}$ , dann kann Rauschen mit  $E\{n\} = 0$  durch Mittelung der gemessenen Grauwerte reduziert werden
- Annahme:
  - Bild besteht aus homogenen Bereichen
  - Benachbarte Pixel haben den gleichen Grauwert
- Mittelwertbildung
  - Konvolutionskern: Gleichmäßige Gewichtung der Pixel in einer gegebenen Nachbarschaft

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

3x3 Mittelwertkernel

## Eigenschaften aller RauschunterdrückungsfILTER

- die Summe aller Elemente des Faltungskerns ist 1
  - Gesamthelligkeit bleibt gleich
- Faltungskern ist punktsymmetrisch zum Koordinatenursprung der Faltungsfunktion
  - Bild wird nicht verschoben und keine Richtung wird bevorzugt

## Verhalten an Kanten

- an Kanten ist die Bildinformation nicht lokal konstant
  - Erwartungswert der ungestörten Funktion ergibt sich als Mittel der überdecktem Gebiete
  - Größere Faltungskerne vergrößern den Bereich der Unschärfe

## Ortsraum-Alternative: Binomialfilter

$$B^0 = 1^{-1} \cdot [1]$$

$$B^1 = 2^{-1} \cdot [1 \ 1] [1 \ 1]$$

- Eindimensionaler Binomialfilter  $B^p = [1 \ 1] * [1 \ 1] * \dots * [1 \ 1]$  (p-mal):

$$B^2 = 4^{-1} \cdot [1 \ 2 \ 1]$$

$$B^3 = 8^{-1} \cdot [1 \ 3 \ 3 \ 1]$$

$$B^4 = 16^{-1} \cdot [1 \ 4 \ 6 \ 4 \ 1]$$

- Zweidimensionaler Binomialfilter  $B^p = B^p * (B^p)^T$ :  
 $B^2 = 4^{-1} \cdot [1 \ 2 \ 1]^T \cdot 4^{-1} \cdot [1 \ 2 \ 1] = 16^{-1}$ .

## Butterworth-Filter vs Binomialfilter

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

- Ideales Tiefpassfilter: kompakter Träger im Frequenzraum, aber artefakt-verursachende Ortsraumrepräsentation
- Butterworth-Filter: kontrolliert monoton fallende Funktion im Frequenzraum, deren Ortsraumrepräsentation ebenfalls monoton fällt.
- Mittelwertfilter: kompakter Träger im Ortsraum, aber artefakt-verursachende Frequenzraumrepräsentation
- Binomial-Filter: monoton fallende Funktion mit kompaktem Träger im Ortsraum, deren Frequenzraumrepräsentation monoton fällt

## Binomialfilter und Gaußfunktion

für größere Filterkerne nähert sich der Binomialfilter der Gauß'schen Glockenkurve an

## Gaussfilter

- Funktion wird nie 0
  - Filterkern endlicher Größe „schneidet“ die Funktion ab
  - ausgewählter Filterkern muss normiert werden

## Glätten für Kanten: Unsharp Masking (USM)

- Gaußfilter verwenden um Unschärfe abzuziehen
- Idee:
  - berechne Bild  $f_s$  durch Filterung von  $f$  mit Gaußfunktion mit Standardabweichung  $s$
  - Addiere mit  $p$  gewichtete Differenz  $f-f_s$  auf das Originalbild  $f_{USM}=f + p(f - f_s)$  ( $p$  wird oft in Prozent angegeben)
  - $f_{USM}$  wird für  $f$  eingesetzt, falls der Unterschied zwischen  $f$  und  $f_s$  größer ist, als eine Schwelle  $t$

Parameter:

- $\sigma$ (Radius)
  - je größer der Radius, desto breiter ist der verstärkende Rand an Kanten
  - der Wert ist die Standardabweichung der Gaußfunktion, d.h. die Breite ist 6-mal größer
- $p$  (amount)

- je höher der Wert ist, desto höher ist die Verstärkung an den Rändern
- t (threshold)
  - je höher der Wert ist, desto stärker muss die Kante sein, damit überhaupt ein Sharpening stattfindet
  - jeder Wert  $t > 0$  führt dazu, dass der Filter nicht mehr linear ist (Artefakte sind möglich)

Unsharp Masking ist keine Bildrestauration!!!

- was unscharf war, bleibt unscharf
- nur bei Unschärfe nahe der Auflösungsgrenze führt der künstliche Machbandeffekt zu einem schärfer wahrgenommenen Bildeindruck

## Kanten

- Rauschunterdrückung versagt an Kanten
- Kanten sind wichtige Bestandteile der abgebildeten Information
- Kantenfilter:
  - heben Kanteninformationen gegenüber restlicher Bildinformationen hervor
  - produzieren Kantenpixel, aber noch nicht Kantenzyge
  - verorten die Kantenaspekte im hochfrequenten Bereich des Frequenzspektrums
  - USM erhöht die Bildschärfe durch Addition von  $p(f - f_s)$

## Hochpassfilter = Kantenfilter?

- nicht ganz
  - Hochpassfilter entspricht näherungsweise einem Laplace-Filter
  - Laplace-Filter: Summe der partiellen zweiten Ableitung in x- und y-Richtung
- Laplace-Filter und Kanten:
  - Definieren den Ort einer Kante (Nulldurchgänge)
  - geben nicht Stärke und Richtung einer Kante an (so wie Gradienten)

## Laplace Funktion

- $$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$
- Summe der partiellen zweiten Ableitungen:
  - Nulldurchgänge der Laplace Funktion = zusammenhängende Kurve
  - Approximation durch Kombination einer doppelten Differenzbildung in x- und y-Richtung
    - z.B. Faltung eines Differenzoperators  $[-1 \ 1]$  mit sich selbst:  $[-1 \ 2 \ -1]$

- ergibt für  $[-1 \ 1] * [-1 \ 1] + [-1 \ 1]^T * [-1 \ 1]^T$ :

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

## 10. Vorlesung: Kanten und Rauschunterdrückung

### Kanten und Rauschen

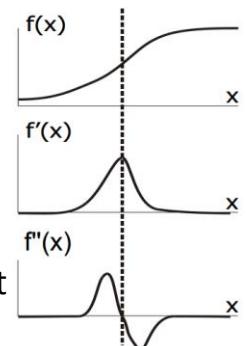
- Bestimmung von Kantenstärke und Kantenrichtung
- Gradienten und Gradientenoperatoren
- Kantenerhaltende Rauschunterdrückung
- Trennung von Rauschen und Kanten
- Medianfilter
- Varianten des Medianfilters

### Kanten und Nulldurchgänge

- Vorzeichenwechsel ist leichter zu erkennen als Minimum oder Maximum
  - Gradient (Länge) als Kennzeichen für wichtige Kanten
  - zweite Ableitung für den Ort der Kante (Nulldurchgang)

### Gradientenschätzung über Konvolution

- Gradient in N dimensionalen Raum ist ein N-dimensionaler Vektor aus N partiellen Ableitungen
- Jede partielle Ableitung kann durch eine Differenz abgeschätzt werden, die durch Konvolution berechnet werden kann



### Sobel-Operator

- Glättung orthogonal zur Differenzierungsrichtung
- Glättung durch gewichtete Mittelung (ähnlich wie Binomialfilter)

$$\begin{array}{ccc} -1 & 0 & 1 \\ -2 & 0 & 2 \end{array} \quad \text{und} \quad \begin{array}{ccc} -1 & -2 & -1 \\ 0 & 0 & 0 \end{array}$$

- Faltungskerne:

$$\begin{array}{ccc} -1 & 0 & 1 \end{array} \quad \begin{array}{ccc} 1 & 2 & 1 \end{array}$$

### Kantendetektion

- Annahme:
  - Nachbarschaft um Kante hat lokale Intensität
  - Rauschen hat Erwartungswert 0
- dann:
  - Glättung so, dass Kantensignal das Rauschen überwiegt
  - Kantenselektion durch Laplace-Operator

- Unschärfe stört bei Kantenbestimmung nicht

## Mehr als eine Kante

- Ungenaue Kantenlokalisierung bei Kantenabstand, der geringer als die Größe des Faltungskerns ist
- Umfang der Glättung ist ein Kompromiss zwischen Rauschunterdrückung und Detailierungsgrad

## Kanten und Rauschunterdrückung

- Grenzen einer Kantenbestimmung durch Glättung + Differenzoperator
  - zu verrausches/detailreiches Bild: kein Kompromiss
  - Kanten soll im Bild verschärft werden → Unschärfe wirkt auf den Betrachter unangenehmer als Bildrauschen
- Alternative
  - Kantenerhaltende Rauschunterdrückung
  - Setzt voraus, dass Unterschiede zwischen Merkmalen von Kanten und von Rauschen herausgearbeitet werden

## Strategie 1: Kombination von linearer Filterung

- Rauschunterdrückung gefolgt von Kantenverstärkung
  - Rauschunterdrückung reduziert Amplitude in hohem Frequenzraum
  - Kantenverstärkung, verstärkt mittleren Frequenzbereich
  - niedriger Frequenzbereich bleibt unverändert
- Beispiel: Unsharp Masking
- Resultat ist heuristisch und produziert „künstlerische“ Effekte vom Originalbilds

## Strategie 2: Charakteristische Grauwertverteilung

- Rauschen und Kanten haben im Frequenzbereich ähnliche Attribute
- Ist ein nichtlineares Filter denkbar, dass für Rauschen und Kanten-eigenschaften unterschiedlich sensitiv ist?
  - Rauschen sind räumlich gleichverteilte Grauwertvariationen
  - Grauwertvariationen an Kanten sind nicht räumlich gleichverteilte
  - Filter muss diesen Unterschied berücksichtigen

## Rangordnungsfilter

- Vorgehen:
  - Sortierung der Elemente in einer Filtermaske
  - Auswahl des an einer bestimmten Stelle eingesortierten Wertes
  - Eintragung des ausgewählten Wertes in die zentrale Position
- Eigenschaften
  - Es entstehen keine neuen Werte,
  - Filter ist nicht kommutativ, nicht assoziativ, nicht linear

## Medianfilter

- Annahmen:
  - Medianfilter ist quadratisch mit ungerader Seitenlänge.
  - Kantensignal ist größer als das Rauschsignal
  - Kante im Filterbereich verläuft in diesem Bereich (nahezu) gerade.
  - Grauwert ist (nahezu) konstant in einer Umgebung von Größe des Filters.
- Kante verläuft durch Filterbereich:
  - Der Mittelpunkt des Filters liegt auf der gleichen Seite der Kante wie die Mehrzahl der Pixel (z.B. rechts von der Kante).
  - Pixel von dieser Seite (z.B. rechts) wird selektiert (kantenerhaltend).
- Keine Kante im Filterbereich:
  - Median nähert sich dem Erwartungswert mit Anzahl der Stichproben (rauschunterdrückend)

## Medianfilter und Impulsrauschen

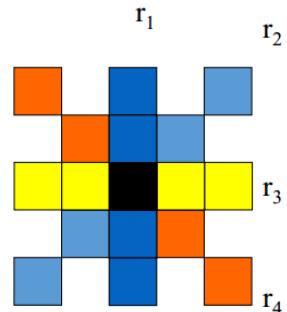
- Impulsrauschen:
  - wenige Pixel sind gestört
  - gestörte Pixel sind entweder maximal positiv oder maximal negativ
- Glättung
  - zu große Nachbarschaft notwendig, um Erwartungswert zu schätzen

## Gewichteter Medianfilter

- Isotrop, mit Abstand zum Zentrum abnehmende Gewichtung
- (Indirekte) Berücksichtigung des Umstandes, dass Korrelation zwischen Pixeln mit der Entfernung abnimmt

- Beispiel:

			ohne Gewichtung
			mit Gewichtung
		1   3   1	1,1,1,1, <b>3</b> ,6,6,6,7
		1   6   6	
		1   6   7	
Gewichtung		Grauwerte	



### Anisotroper Medianfilter

- Zerlegung in verschiedene Richtungen
- Medianfilterung auf allen Richtungen
- Berechnung des Medians aller Mediane
- Besserer Behalt von Bilddetails nach Filterung

## 11. Vorlesung: Datengetriebene Segmentierung 1

### Referenzsegmentierung

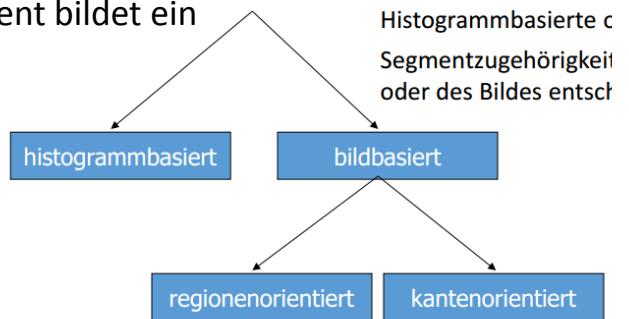
- es existieren meistens Segmente, die jeder als solche bezeichnet, aber
  - was Hintergrund ist, wird oft unterschiedlich gesehen
  - Segmente selbst werden unterschiedlich detailliert gesehen
- Segmentierung ist nicht a-priori lösbar

### Segmente

- Segmente und Semantik
  - Unterschiedliche Segmente haben unterschiedliche Bedeutung
  - Pixel innerhalb eines Segments haben die gleiche Bedeutung
- Segmentierung ist ein subjektiver Prozess
- Segmente können aus dem Bild generiert werden, falls
  - Pixel innerhalb eines Segments andere Eigenschaften haben, als die außerhalb eines Segments
  - jedes Pixel nur einem Segment angehört
- es gibt viele Fälle bei denen diese Annahmen (näherungsweise) zutreffen
- Ziel: Pixelaggregate schaffen
- Pixel gehören zusammen, wenn sie zusammenhängendem Gebiet angehören
- Homogenitätsbedingung ist vom Bildinhalt unabhängig
- modellbasierte Segmentierung: Zwischen Segmentierung und Analyse

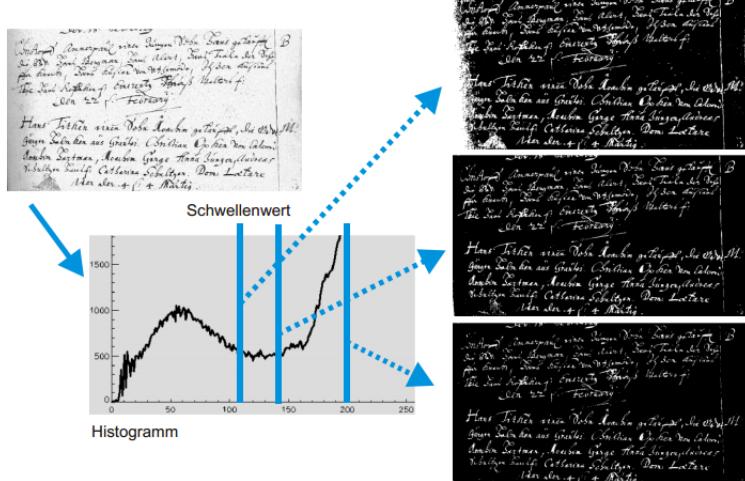
- Eigenschaften:
  - vollständig: jedes Pixel ist einem Segment zugeordnet
  - überdeckungsfrei: ein Pixel ist genau einem Segment zugeordnet
  - zusammenhängend: jedes Segment bildet ein zusammenhängendes Gebiet

## Datenbasierte Segmentierung



## Histogramm basierte Segmentierung

- Annahme: Bild besteht aus zwei Anteilen, die sich durch ihren Grauwert unterscheiden
- Aufgabe: Schwellenwert zwischen den beiden Grauwerten finden



## Shading

- Helligkeitsvariationen zerstören die bimodale Verteilung der Häufigkeiten  
→ Schwellenwert ist nicht mehr für das ganze Bild definiert
- Berücksichtigung:
  - homogenes Bild unter gleichen Bed. aufnehmen
  - Shading-Bild aus dem Bild selbst bestimmen
  - variierende Schwellenwerte
- Bestimmung des Shadingbildes:
  - falls Mehrzahl des Bildes aus Vor- und Hintergrund → Rangordnungsfilter (muss mind. so groß sein, dass ein Vor- und Hintergrundpixel enthalten ist)
- Shading-Korrektur

- berechne aus dem Hintergrundbild Shadingfunktion  $s(i,j)$
- Korrektur:  $f'(i,j) = f(i,j) * s(i,j)$
- Segmentierung aus dem korrigierten Bild

## Region Labelling

- Schwellenwert zerlegt das Bild in Vor- und Hintergrundsegmente
- Region Labelling bestimmt Ort und Anzahl aller zusammenhängenden Gebiete im Binärbild  $b$ 
  1. Nach der Schwellenwertsegmentierung haben Regionen in Abhängigkeit von ihrem Grauwert ein Label erhalten
  2. zusammenhängende Gebiete mit gleichem Label gesucht
  3. jedes Gebiet erhält ein neues, eigenes Label
- Trennung nach Grauwerten wird nicht perfekt sein
- Schwellenwertbildung enthält falsche Regionen (kleine Gebiete, Störungen am Rand)
- Nachverarbeitung:
  - Medianfilter auf den Labeln
  - Entfernung von zu kleinen Regionen
  - Relaxations Labelling

## Regionbasierte Segmentierung

- Homogenität im Inneren des Segments
- Homogenitätskriterium wird ausgewertet
- Homogenität ist relativ zu den Attributen des Segments definiert
- Globale Zusammenhänge über Multiskalenstrategie

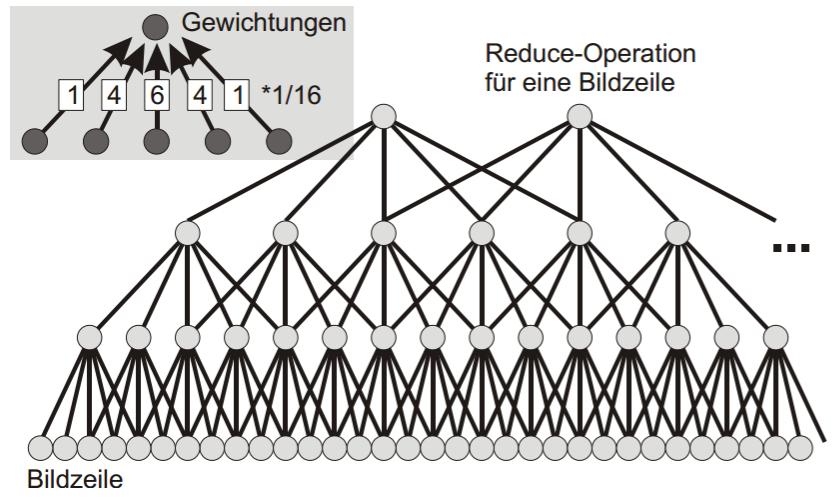
## Multiskalenstrategie

- relative Kriterien für Homogenität können über unterschiedliche Entfernungswerte wirken
- Segmentierung nach Multiskalenstrategie wertet Kriterien auf unterschiedlichen Skalierungen aus
- Multiskalenstrategie
  - explizit auf einer Multiskalenrepräsentation
  - implizit in den Segmentierungsalgorithmus integriert

## Gaußpyramide

- Originalbild wird durch eine „reduce“-Operation skaliert
- jedes Pixel der nächst höheren Stufe repräsentiert 4 Pixel der aktuellen Stufe
- vor der Reduktion wird der Frequenzumfang durch Filterung vermindert

- Filterung mit Gaußfilter oder Binomialfilter



### Expand-Operation

- Um die vorherige Skalierungsstufe zu erzeugen, wird eine „expand“-Operation definiert.
- Pixel der neuen Skalierungsstufe werden durch Interpolation erzeugt:
- Pixelorte, die auf beiden Skalierungsstufen existieren:
- Pixelorte, die nur auf der vorherigen Skalierungsstufe existieren:
- nicht verlustfrei

### Laplace-Pyramide

- Jede Skalierungsstufe  $s$  enthält nur den Unterschied  $f_s - \text{expand}(\text{reduce}(f_s))$
- Redundanzfreie Repräsentation

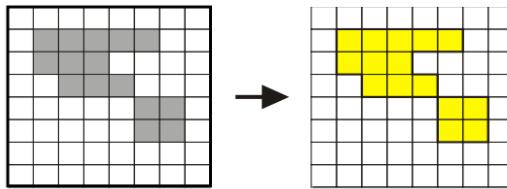
### Region Merging

- Initial wird jedes Pixel zu einem Segment erklärt.
- Zwei benachbarte Regionen werden zusammengefasst, wenn sie auch gemeinsam das Homogenitätskriterium erfüllen.
- Die Segmentierung ist beendet, wenn keine zwei Regionen mehr existieren, die zusammengefasst werden können.
- Zwischenergebnisse werden in einem Region Adjacency Graph (RAG) gespeichert
- Region Labelling kann integriert werden
- Homogenitätskriterium:
  - Grauwertdifferenz innerhalb der Region
  - Wahrscheinlichkeit, dass beide Regionen die gleiche Grauwertverteilung haben

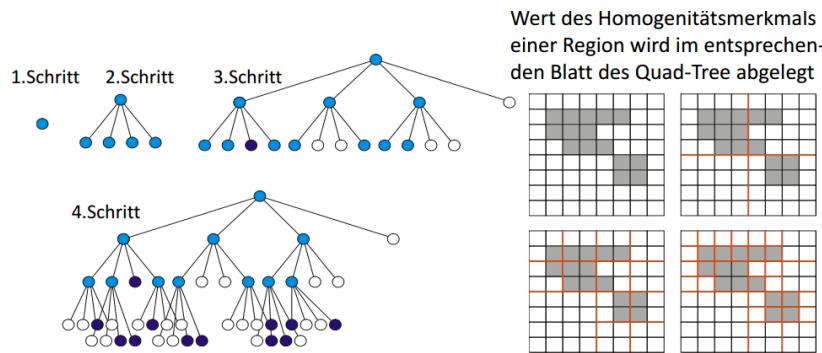
### Split & Merge-Algorithmus

- Startbedingung: das gesamte Bild ist ein Segment

- ein Segment wird solange in 4 Untersegmente zerlegt, wie es ein gegebenes Homogenitätskriterium verletzt (terminiert spätestens auf Pixelebene)
- benachbarte Segmente werden zusammengefasst, wenn sie auch nach der Zusammenfassung das Homogenitätskriterium erfüllen
- Resultat: vollständige, überdeckungsfreie Zerlegung des Bildes (Segmentierung gemäß Definition)



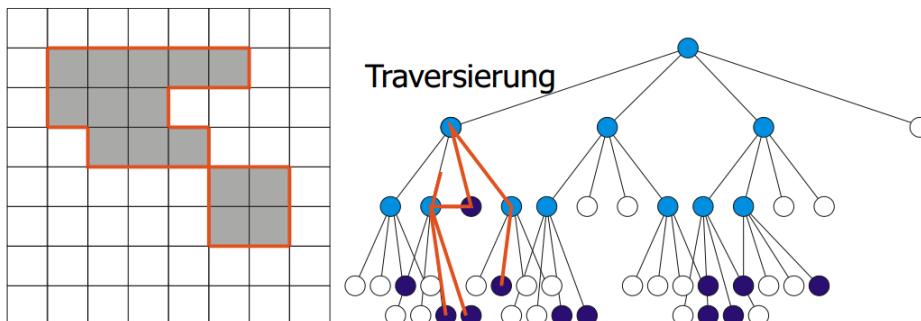
### Zerlegungsschritt (Quad-Tree Präsentation)



### Merging

Quadtree wird traversiert und in einen RAG überführt

auf dem RAG wird ein Region Merging durchgeführt



### Probleme:

- Merge-Schritt ist bei relativen Homogenitätsmerkmal nicht immer eindeutig.
- Keine absoluten Merkmale möglich.

## Kantenbasierte Segmentierung

- Edge Linking und Canny Operator
- Nulldurchgänge
- Wasserscheidetransformation
- Relaxations Labeling zur Nachverarbeitung

## Segmentierung durch Kantenerkennung

- Vorteil: Kantenmerkmale sind robuster gegen Shading
- einfache Methoden:
  - Gradientenberechnung, Kantenpunktdetektion
  - Region Labeling basierend auf Kantenpunkten

## Edge Linking

Anfangs sind alle Kantenpixel frei und nicht untersucht. Edge Linking sucht sich das nächste, nicht untersuchte und freie Kantenpixel und versucht es mit anderen Pixeln zu einem Kantenzug zu verknüpfen.

1. Suche das nächste Kantenpixel, welches noch nicht bereits als „untersucht“ markiert wurde und erkläre es zum Startpixel eines Kantenzugs
2. Falls sich in der Umgebung des Kantenpixels in einer der beiden Richtungen orthogonal zur Kantenrichtung unmarkierte Kantenpixel befinden, die eine ähnliche Gradientenrichtung und -stärke aufweisen:
  - a. Markiere die Pixel als zum Kantenzug zugehörig.
  - b. Erkläre diese Pixel zu neuen Startpixeln.
  - c. Gehe zu 2.
3. Falls sich in der Umgebung markierte Pixel befinden, die den obigen Bedingungen genügen, dann wurde eine Verzweigung von Kanten gefunden
4. Falls kein Kantenpixel gefunden wurde, gehe zurück zu Schritt 1

## Canny Edge Operator

Ziel:

- möglichst viele Kanten vom Hintergrund unterscheiden zu können
- (unverzweigte) Kanten genau lokalisieren
- für jede Kante genau eine Detektorantwort zu liefern

## Canny Operator: Kantenhervorhebung und Erzeugung von Kantenzygen

Optimale Kantenhervorhebung Filterung 1D abgeleitete Gaußfunktion orthogonal zur Kante

Geringfügig schlechtere Ergebnisse 2D Gradientenoperator auf Basis der abgeleiteten Gaußfunktion

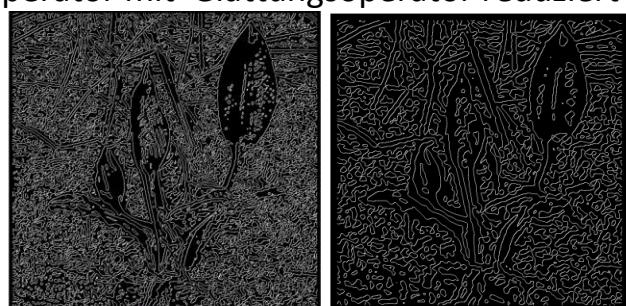
### Canny Edge Detektion

1. Anwendung eines Gradientenoperators (z.B. Sobel-Operator)
2. Non-Maximum Supression (z.B. Nulldurchgänge)
3. Kantenverfolgung- Startpixelsuche
  - es wird immer das Pixel mit der größten Gradientenlänge selektiert
  - Startpixel können nur Pixel sein, deren Gradientenlänge oberhalb einer Signifikanzschwelle T1 liegt
4. Kantenverfolgung – Tracking
  - Neue Kantenpixel haben eine Gradientenlänge  $> T_2$  ( $T_2 < T_1$ ) und sind zu einem bereits gefundenen Kantenpixel benachbart.

Verfahren endet, wenn keine neuen Startpixel gefunden werden.

### Nulldurchgänge zur Segmentierung

- Orte der Nulldurchgänge der zweiten Ableitung sind Ränder zusammenhängender Gebiete
- Methode:
  - Laplace Operator
$$\nabla^2 f(i, j) \cdot \nabla^2 \text{shift}[f(i, j)] \leq 0$$
  - Nulldurchgänge bestimmen  
(shift: Verschiebung des Bildes um einen Pixel in jede Richtung)
- Menge der Nulldurchgänge bildet immer geschlossene Kurve = Segment
- Kombination des Laplace Operator mit Glättungsoperator reduziert die



Anzahl der Nulldurchgänge

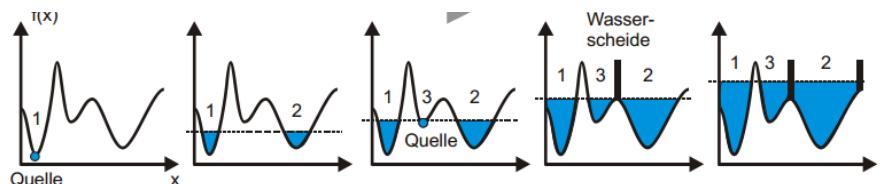
### Wasserscheidentransformation

- Wasserscheide: Menge aller Orte, die die Grenze der Entwässerung in unterschiedlichen Senken sind

- Wasserscheide in der Segmentierung: Generiere ein Höhenprofil so, dass die Wasserscheiden gerade die gesuchten Segmentgrenzen sind
- Wasserscheide soll an Kanten verlaufen
- Wasserscheiden sind „Gebirgskämme“
- Berechnung:
  - Es fällt „Regen“ auf jedes Pixel. Anhand des Gradienten wird entschieden, wohin der Regen entwässert wird
- Flutung:
  - Die „Welt“ wird von den Senken her geflutet. Immer wenn Wasser aus zwei Senken zusammen fließt, entsteht eine Wasserscheide

### Flutungsalgorithmus

Jedes neue Pixel ( $m_f, n_f$ ) ist



- in Isolation:
  - es nicht zu anderen überfluteten Pixeln der Höhe  $h < h_{\text{aktuell}}$  benachbart ist
  - Isolierte Pixel sind Kerne neuer Segmente
- Erweiterung
  - es ist zu andern Pixeln der Höhe  $h < h_{\text{aktuell}}$  mit gleichem Label benachbart
  - Pixel wird dem Segment mit diesem Label zugeordnet
- Wasserscheide
  - zu überflutende Pixel ist zu mind. zwei Regionen benachbart
  - Label „Wasserscheide“

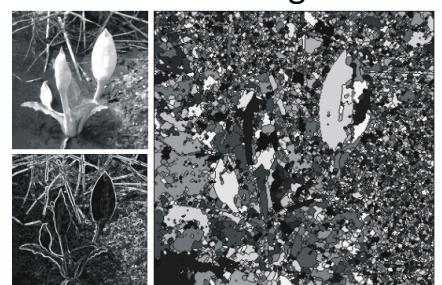
### Resultat der WST

- WST auf Gradienten
- Segmentgrenzen sind lokale Maxima in Gradientenrichtung
- Maxima der ersten Ableitung = Nulldurchgänge der zweiten Ableitung

→ WST sucht nach Nulldurchgängen

### Problem: Übersegmentierung

- für WST sind alle lokalen Minima Senken
- meisten Senken werden durch Rauschen verursacht



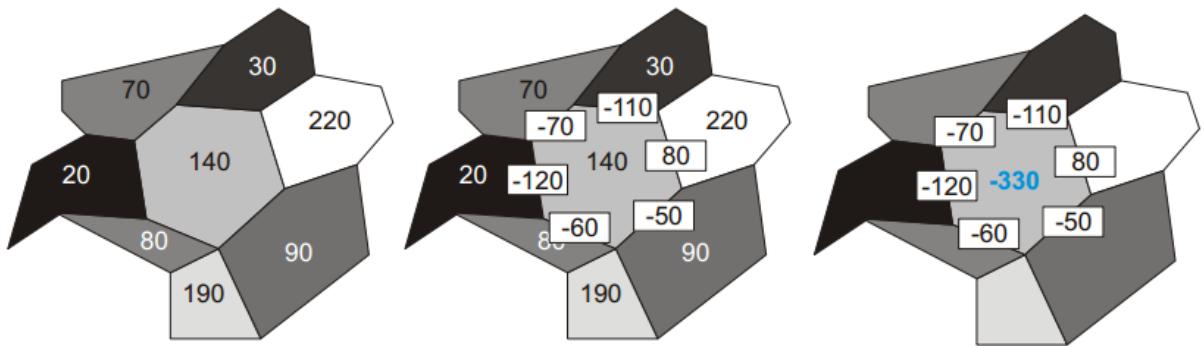
- Senken von Rauschen sind weniger tief als die von Kanten

### **hierarchische WST (hWST)**

- Multiskalenstrategie
- jede Region erhält ihren durchschnittlichen Grauwert als Funktionswert
- erste WST wird hauptsächlich durch Rauschen verursachte Senken finden
- „wahre“ Stufen sollten über mehrere Stufen der Hierarchie erhalten bleiben

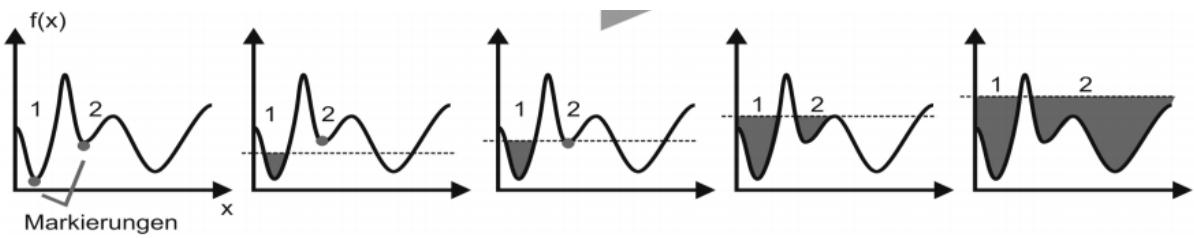
### **Gradienten hWST**

- zu jeder benachbarten Region wird die Differenz gebildet
- Die Länge des Gradienten ist die durchschnittliche Differenz zu allen Regionen.
- Die Richtung ergibt sich aus der (mit der Regionengröße gewichteten) Vektoren zwischen dem Schwerpunkt der Region zu den Schwerpunkten aller benachbarten Regionen



### **Markerbasierte WST**

- Flutung erfolgt nur von vorher definierte Markierungen aus
- Abwandlung des Flutungsalgorithmus:
  - Flutung erfolgt wie vorher von den Senken aus
  - Regionen erhalten das Label „undefined“, wenn sie nicht von einer Markierung aus geflutet werden
  - Wenn eine Region „undefined“ mit einer mit Label versehenen Region zusammenfließt, dann erhält sie dieses Label



## Markierung suchen

- mWST setzt Objektwissen voraus
- Markierungen können z.B. erzeugt werden, falls die gesuchten Objekte mindestens in einem Punkt anhand der Helligkeit identifizierbar sind
- Beispiel: Segmentierung in der Elektrophorese oder Zellsegmentierung

## Nachverarbeitung von Segmenten

- Übersegmentierung: Homogenität ist durch Rauschen gestört
- Charakterisierung der Störung: Rauschen verursacht viele kleine Segmente
- Nachverarbeitung: kleine Segmente werden benachbarten Segmenten zugeordnet
- Problem: Was ist ein „Mastersegment“
- Lösung: relaxierendes Verfahren

## Relaxations Labelling

- jedes Pixel erhält für jedes Label eine vorgegebene W-keit
- (Bsp.: Schwellenwertresultat, Wahrscheinlichkeiten „70% weiß“ und „30% schwarz“ für weiße Pixel; umgekehrt für schwarze Pixel)
- benachbarte kompatible Pixel unterstützen sich
- Relaxationsprozess: Zuordnungsw-keit ändert sich mit dem Maße der Unterstützung
- zu definieren:
  - Kompatibilität
  - Einfluss der Kompatibilität auf die Labelw-keit

## Labelwahrscheinlichkeit

- Umsortierung aller Pixel in Liste  $p_0, \dots, p_n$
- Initiale Labelwahrscheinlichkeit  $P^0$  für jedes Pixel  $p_i$  und jedes Label  $l_k$  vergeben, z.B.

$$P^0(p_i, l_k) = \begin{cases} 0.8 & , \text{ falls } l_k = l(p_i) \\ 0.2 & , \text{ sonst} \end{cases}$$

- Labelwahrscheinlichkeit
  - gibt an, wie sicher man sich nach der Segmentierung über das zugeordnete Label ist
  - darf nicht 0 oder 1 sein (Gewissheiten werden nicht verändert)

## Kompatibilität

- Ein Pixel  $p_i$  mit Label  $l_k$  hat eine Kompatibilität  $r$  mit einem Pixel  $p_j$ , dessen Label  $l_l$  sei:  $r((p_i, l_k), (p_j, l_l))$
- Kompatibilität für Binärbilder (2 Label) z.B.
 
$$r((p_i, l_k), (p_j, l_l)) = r(l_k, l_l) = \begin{cases} 1 & , \text{falls } l_k = l_l \\ 0 & , \text{sonst.} \end{cases}$$

(d.h. nur gleiche Label unterstützen sich)
- Kompatibilität bei mehr als zwei Labeln kann auch bedeuten, dass sich bestimmte Labelpaare unterstützen (auch wenn es unterschiedliche Label sind)

## Confidence Map

- um das Konvergenzverhalte zu beobachten, kann eine Confidence Map erzeugt werden
- Confidence Map: Gibt für jedes Pixel die Zuverlässigkeit der derzeitigen Entscheidungs an
- Für 2-Label-Segmentierung: Differenz zwischen gewählten Label und nicht gewähltem Label (evtl. gewichtet mit der Anzahl der Pixel mit diesem Label).

## Zwischenfazit Segmentierung

- Segmentierungskriterien
  - möglichst homogene Gebiete
  - möglichst kurze Segmente
- Probleme
  - Homogenitätskriterium kann schwer zu definieren sein
  - Homogenität kann in unterschiedlichen Bereichen etwas unterschiedliches bedeuten
  - Segmentgrenzen sind nicht immer Grauwertkanten

## 13.Vorlesung: Modellbasierte Segmentierung

- Interaktive Suche

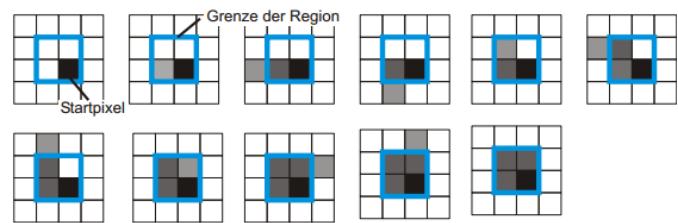
- (markerbasierte WST)
- Region Growing
- Kantenverfolgung
- vollständige Suche
  - Template Matching
  - Hough Transformation

## Modellbasierte Segmentierung

- Segmentierung: Generierung von Symbolen (Bedeutungsträgern) aus Pixeln
- Modell: erwartete Bedeutung  
→ Widerspruch!!
- Mixtur zwischen Segmentierung und Analyse. Mit Vorwissen über ein bestimmtes Objekt wird nach Instanzen gesucht
  - Interaktive Suche: Benutzer gibt Vorwissen interaktiv ein
  - Vollständige Suche: Instanz eines Modells mit wenigen Parametern werden gesucht
  - Iterative Suche: Instanzen eines Modells mit vielen Parametern werden gesucht

## Region Growing

- Flood Fill auf Grauwertbild für einen einzigen (vorgegebenen) Saatpunkt
- Homogenitätsbedingung:
  - Grauwertbereich
  - Grauwertschwankung
- Annahme und Ziel
  - Gebiet ist intern von größerer Homogenität als an der Grenze
  - Homogenitätsverhältnisse an anderen Orten interessieren nicht
- Probleme:
  - „Auslaufen“ der Region
  - zu kleine Region
  - Rauschanfälligkeit
  - Shading im Bild
- interaktives, recht intuitives Werkzeug
  - Startpunktselektion sollte für „Datenexperten“ trivial sein
  - Homogenitätskriterium schwierig zu bestimmen
- geeignet falls das gesuchte Objekt homogen ist und sich überall vom Hintergrund abhebt
- „Auslaufen“ kann durch untersch. Maßnahmen begrenzt werden
- auch in 3d anwendbar, aber wenig intuitiver

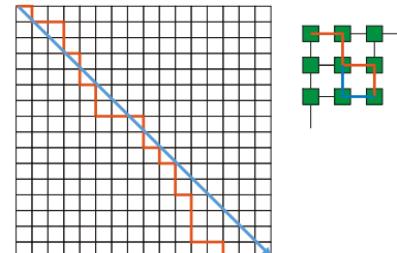


## Gezielte Kantensuche

- Resultat von Kantenfilter
- Liste von Kantenpunkt-kandidaten
- nicht alle Kandidaten gehören zur gesuchten Regionengrenze
- Regionengrenze kann Lücken aufweisen
- Strategie mit Nutzerinteraktion
  - optimale Kantenzüge
  - Hough Transformation

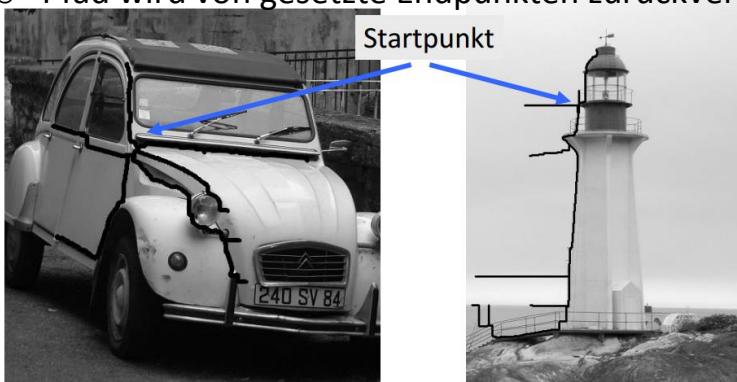
## Optimale Kantenzüge (Pfadsuche)

- Pixel sind Knoten Graphen
- benachbarte Pixel sind durch Kanten verbunden
- Vorteil:
  - Globale Aspekte des Kantenzugs können eingebracht werden (z.B. Gesamtlänge)
- Naheliegende Optimalitätskriterien:
  - Maximierung der Gradientenlänge
  - Minimierung der Abweichungen von einer erwarteten Gradientenlänge
  - Minimierung der Pfadlänge
  - Minimierung der Richtungsänderungen
  - Minimierung der Grauwertänderungen



## Live Wire

- Startpunkt liefert „Modellgradienten“
- Kriterien
  - wenige Abweichungen von der Länge des Modellgradienten
  - kurze Pfadlänge
- Methode
  - Dijkstra-Algorithmus zu allen möglichen Endpunkten
  - Pfad wird von gesetzte Endpunkten zurückverfolgt

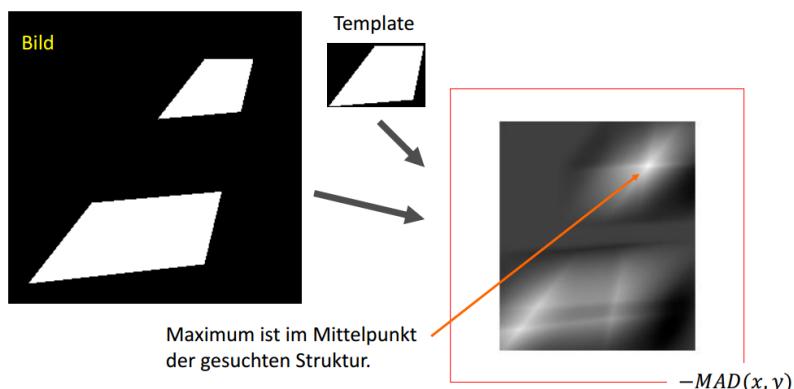


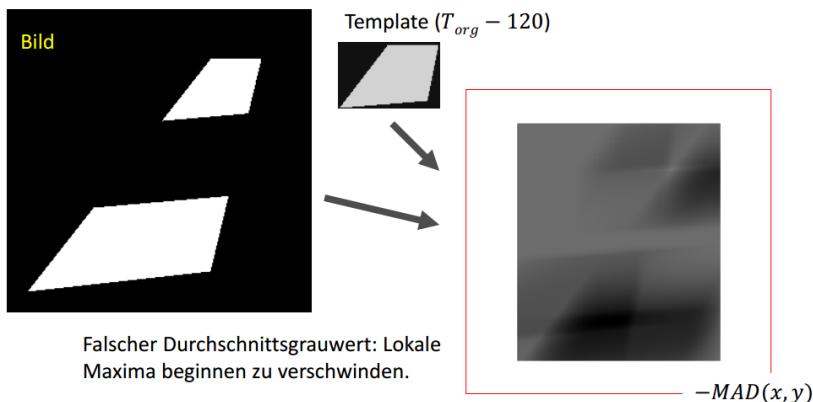
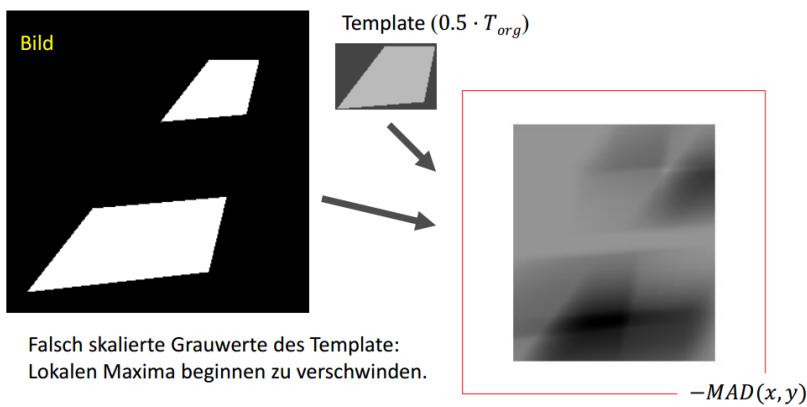
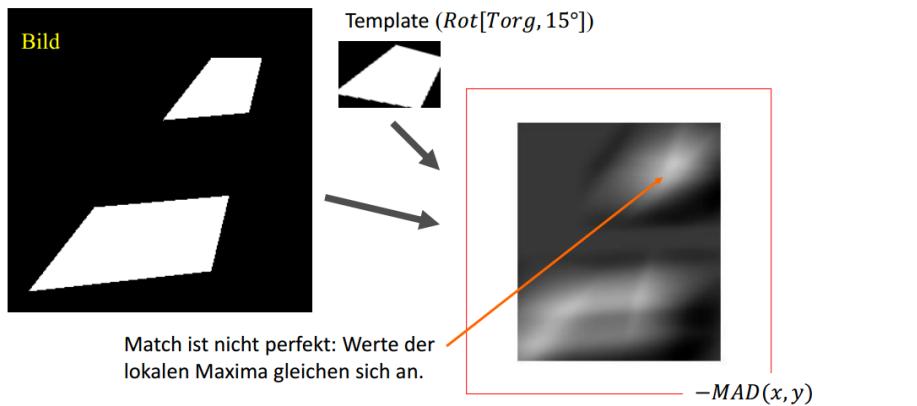
- schnelles, intuitives Werkzeug zur interaktiven Segmentierung
- eignet sich gut, wenn Kanten klar abgrenzbar sind, auch wenn die Gradienten stärke lokal variiert
- kann sogar Lücken in Kettenzügen überbrücken (aber ist rauschanfällig)

## Template Matching

Segmentierung durch vollständige Suche

- gezielte Suche nach einer vorgegebenen Form
- jeder Ort im Bild wird ein Abstandsmaß zwischen Form und Bild berechnet
- Ort an dem das Maß am höchsten ist, ist die gesuchte Position des Objekts
- Abstandsmaße zwischen Bild  $f(x,y)$  und Template  $t(x,y)$ : z.B.
  - MAD: mean absolute difference:
$$MAD(x,y) = \frac{1}{MN} \cdot \sum_{xy} |f(x+m, y+n) - t(x, y)|$$
  - Korrelationskoeffizient  $cc(x,y)$
- unempfindlich gegenüber Rauschen
- Probleme:
  - Werte der lokalen Maxima gleichen sich an
  - falsch skalierte Grauwerte Template: lokale Maxima verschwinden
  - falscher Durchschnittsgrauwert: Lokale Maxima beginnen zu verschwinden





## Template Matching

- schnelles, automatisches Segmentierungsverfahren
- geeignet, wenn die Form des gesuchten Objekts wenig variable ist (Verkehrzeichen, Warenzeichen,...)
- Skalierungs- und Rotationsabhängigkeit kann zusätzliche Strategie für schnelle Berechnung erfordern
- kann auch zur Suche nach Regions of Interest benutzt werden

## Hough Transformation

Modellbasierte Suche:

- Modell einer Regionengrenze ist vorgegeben

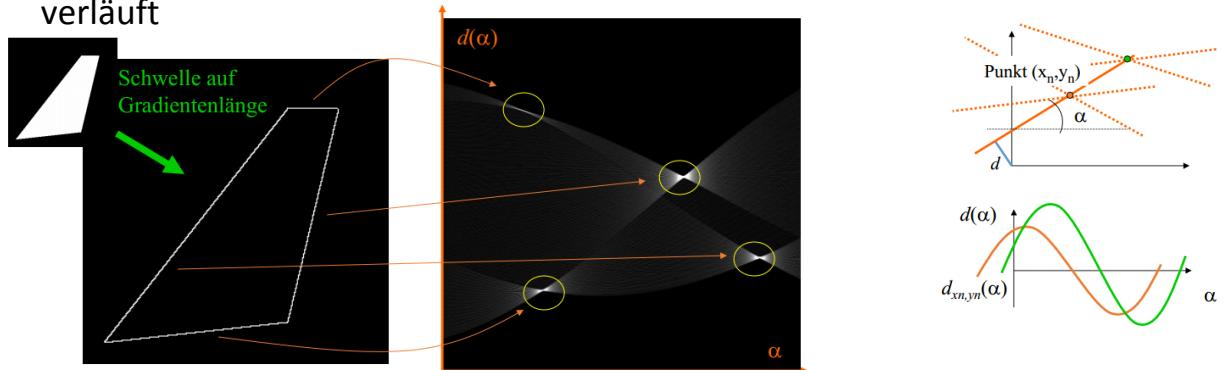
- Information über mögliche Kantenorte ist gegeben
- Zielsetzung: Orte finden, an denen Modell und Bild optimal übereinstimmt

## Hough Transformation (HT)

- Voting-Mechanismus, bei dem jeder Ort in Abhängigkeit der lokalen Information für das Modell stimmt
- entwickelt für Geraden, erweiterbar für beliebige Formen
- Suche von Geraden in einem Binärbild
- Geradenrepräsentation:  $x \cos(\alpha) + y \sin(\alpha) - d = 0$
- Suche alle Parameter  $(\alpha, d)$  für Geraden, die durch einen Punkt  $(x_n, y_n)$  gehen
- $d(\alpha) = x_n \cos(\alpha) + y_n \sin(\alpha)$
- Der Raum, der durch  $(\alpha, d)$  aufgespannt wird, heißt Hough-Raum

## Berechnung der HT

- Erzeugung eines Kantenbildes durch Schwellenwertsetzung auf Gradientenlänge
- Diskretisierung des  $(\alpha, d)$ -Raums (Zerlegung in Akkumulatoren)
- für jede Punkt  $x_n, y_n$  wird eine Kurve im  $(\alpha, d)$ -Raum diskretisiert
- jeder Akkumulator wird inkrementiert, sobald eine Kurve durch ihn verläuft



## Hough Transformation

- Automatisches Verfahren, das Segmentränder und Teile von Segmenträndern findet
  - flexibler als Template Matching
  - kann aber Nachverarbeitung erfordern (Zusammensetzung von Randteilen)
- Voting-System ist interessant für Echtzeitanwendung
  - bereits Teilmenge der Stimmen erlaubt Vorhersage auf Endergebnis

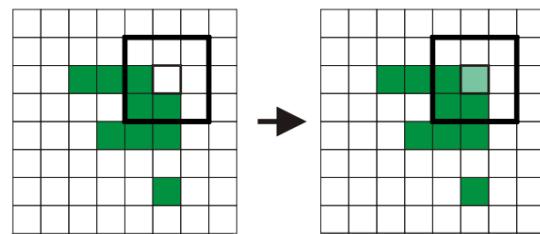
- Stimmabgabe kann gesteuert werden, sodass „wichtige“ Stimmen zuerst abgegeben werden

## 14. Vorlesung: Morphologische Operationen

- Erosion und Dilatation
- Opening und Closing
- Ränder und Distanzen, Morphing
- Hit- or –Miss – Operator
- Skelettierung
- die äußere Gestalt betreffend
- morphologische Operationen
  - Operation auf der Gestalt von Objekten  
→ setzt die Extraktion einer Gestalt voraus
  - in erster Linie Operation auf Segmente
- Wozu ist es gut?
  - Veränderung der Gestalt, um Störung nach einer Segmentierung zu beseitigen
  - Berechnung von Formmerkmalen
  - Suche nach bestimmten Formen

### Dilatation (Ausdehnung)

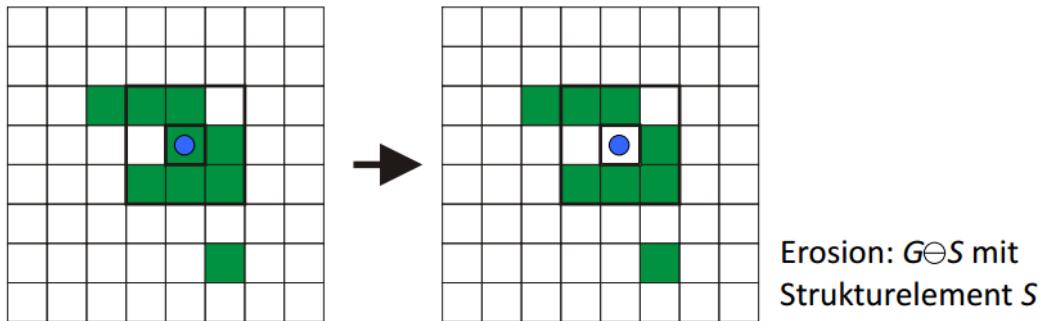
- $G \oplus S$  mit Strukturelement  $S$
  - Dilatation wird für einen Ankerpunkt ausgeführt
    - verbindet Strukturen
    - füllt Löcher
    - vergrößert
- $$g(m,n) = \bigvee_{(m_k, n_k) \in S} b(m + m_k, n + n_k)$$



### Erosion

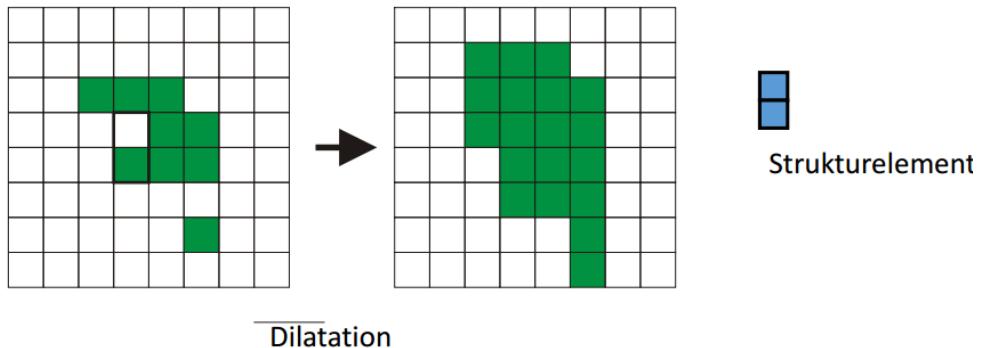
- löst Strukturen auf
- entfernt Details
- verkleinert

$$g(m, n) = \bigwedge_{(m_k, n_k) \in s} b(m + m_k, n + n_k).$$



## Strukturelemente

- entspricht dem Faltungskern bei einer Konvolution
- mit gezielt geformten Strukturelementen können genau definierte Formveränderungen definiert werden



## Eigenschaften von Morphologischen Operationen

- Verschiebungsinvarianz: Wegen der Beschreibung von Erosion/Dilatation als Faltung sind beide Operationen genau wie eine Faltung verschiebungsinvariant
- Kommutativität und Assoziativität:  $M_1 \oplus M_2 = M_2 \oplus M_1$  aber  $M_1 M_2 \neq M_2 M_1$   
es gilt jedoch  $(G M_1) M_2 = G (M_1 M_2) = (G M_2) M_1$
- Dualität:  $\overline{G} \ominus M = \overline{G \oplus M}$  und  $\overline{G} \oplus M = \overline{G \ominus M}$

## Opening

- Kombination von Erosion gefolgt von einer Dilatation am Ankerpunkt gespiegeltem Strukturelement  $S'$   $\mathbf{G} \circ S = (G \ominus S) \oplus S'$

- Ziel:
  - Erosion: Entfernung aller (Teil-)strukturen, die kleiner als das Strukturelement sind
  - Dilation: Wiederherstellung der ursprünglichen Größe des Objekts mit Ausnahme der vollständigen entfernten Teilstrukturen

## Closing

- Kombination von Dilation gefolgt von einer Erosion am Ankerpunkt  

$$G \bullet S = (G \oplus S) \ominus S'$$
- Ziel:
  - Dilatation: Schließen von kleinen Löchern
  - Erosion: Wiederherstellung der ursprünglichen Größe des Objekts

## Extraktion von Rändern

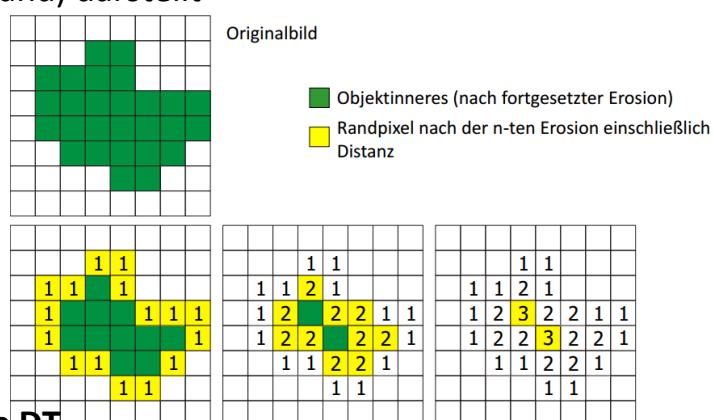
- Erosion von  $S_{b4}$  bzw.  $S_{b8}$  entfernt alle Objektpixel, in deren 4- und 8-Nachbarschaft sich Hintergrundpixel befinden
- Der Rand kann nur durch Differenz zwischen Ursprungsbild und erodiertem Bild erzeugt werden  

$$\partial G_0 = G \setminus (G \ominus S_b)$$

## Distanztransformation

- Fortgesetzte Extraktion von immer weiter vom Rand entfernten Linien und Multiplikation der jeweiligen Resultate mit der aktuellen Entfernung überführt das Binärbild in ein Distanzbild D:  

$$D = \bigcup_{n=1, \infty} [(G \ominus S_b^{n-1}) \setminus (G \ominus S_b^n)] \cdot n$$
 wobei die Operation
  - die punktweise Multiplikation der n-ten Randkurve mit der Zahl n (dem aktuellen Abstand) darstellt



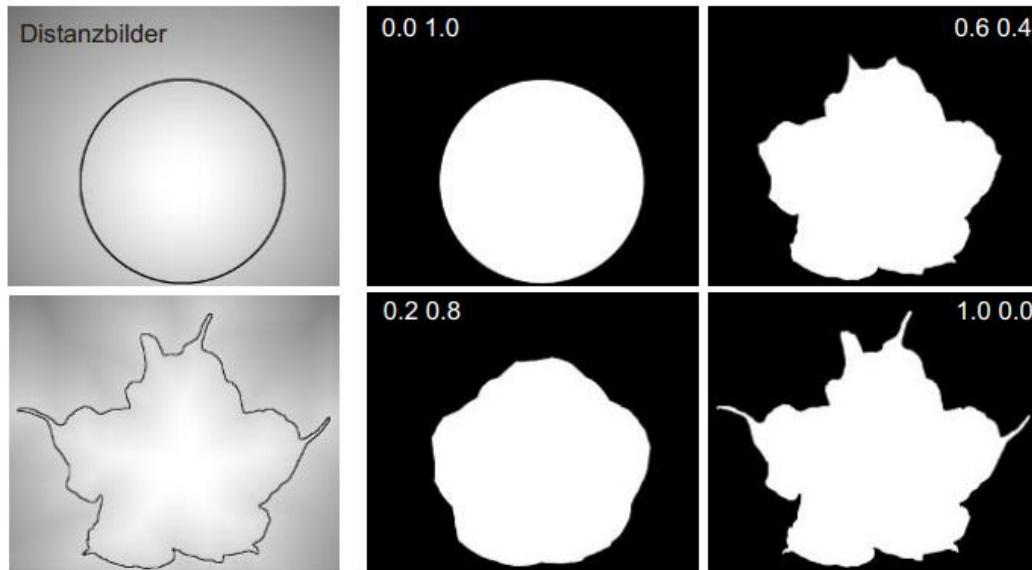
## Vorzeichenbehaftete DT

- Distanztransformation auf dem Vordergrund
- DT auf dem Hintergrund (Operation auf dem negierten Bild)
- Kombination beider DT's

- Übernahme der Ergebnisse auf dem Vordergrund
- Hintergrunddistanzen werden negativ eingetragen

## Morphing

vorzeichenbehaftete DT auf Binärbildern durchführen



## Hit-or-miss-Operator

- Separierung gestaltbasiert definierte Aspekte eines Segments

