

# Digitale Bildverarbeitung

Prof. Dr. Sibylle Schwarz  
HTWK Leipzig, Fakultät IM  
Gustav-Freytag-Str. 42a, 04277 Leipzig  
Zimmer Z 411 (Zuse-Bau)

<https://informatik.htwk-leipzig.de/schwarz>  
[sibylle.schwarz@htwk-leipzig.de](mailto:sibylle.schwarz@htwk-leipzig.de)

Sommersemester 2019

# Bildverarbeitung – Motivation

Menschen entnehmen visuellen Eindrücken sehr viel Information (ca. 90%)

oft Bilder dreidimensionaler Szenen

maschinelle Bildverarbeitungsverfahren dienen der

- ▶ Informationsgewinnung aus Bildern  
(Bilderkennung, Bildanalyse)
- ▶ Unterstützung der Informationsgewinnung aus Bildern  
(Bildbearbeitung)
- ▶ Darstellung der Information aus Bildern  
(Szenenrepräsentation)

# Bildwahrnehmung

Szene: Foto vom Laborgebäude

Objekte: Wolken, Gebäude, Kunst

Regionen: Himmel, Fassadenteile

Texturen: Wirbel-, Linienstrukturen

Ecken, Kanten: (Farb-)Kontrast

Pixel: Farbwert

Rohdaten: Fotoaufnahme

# Digitale Bildbearbeitung – Ziele

**Informationsgewinn** Extraktion / Hervorhebung der relevanten  
Bildinhalte zur menschlichen oder maschinellen  
Weiterverarbeitung

**Informationsverlust** beabsichtigt  
Abstraktion von unwichtigen Details  
(entfernen, abschwächen)

# Bildverarbeitung – Einsatzgebiete

**Medizin** , Biologie, Astronomie:

Aufnahme und Auswertung z.B. von Mikroskop-,  
Teleskop-, Röntgen-, Ultraschall-, CT-Bildern

**Zeichenerkennung** (OCR):

Text, Handschrift, Symbole, Noten

**Überwachung** , z.B. industrieller Prozesse, Verkehr,  
Qualitätskontrolle, Materialprüfung

**Luftbilder** und Satellitenaufnahmen

z.B. bei Kartierung, Erkundung, Meteorologie

**Biometrie** (Identifikation zur Authentifizierung, Kriminalistik):  
Finger-, Handabdruck, Iris-Scan zur

**Autonomes Fahren** Bewegungsplanung und -korrektur,  
Orientierung

# Bildverarbeitung – Aufgaben

**Aufnahme** durch technische Systeme

**Erzeugung** künstlicher Bilder (z.B. Texturen)

**Wiedergabe** Bildschirm, Druck

**Analyse** statistischer Merkmale

**Transformation** , z.B. Entrauschen, Kontrastverstärkung

**Restaurierung** gestörter Bilder

**Erkennung** im Bild enthaltener Kanten, Ecken, Objekte

**Interpretation** der dargestellten Szene (Art, Anordnung der Objekte)

z.B. in medizinischer BV, Lokalisierung

**Kompression** zur effizienten Speicherung mit möglichst wenig visuellen Verlusten

digitaler Bilder

# Bildverarbeitung – Verwandte Gebiete

**Signalverarbeitung** , (z.B. Audiosignale) ähnliche Techniken  
(z.B. statistische Merkmale,  
Fourier-Transformation, Filter)

**Bildbearbeitung** häufig mit spezialisierter Software

**Computergraphik** Erzeugung von Bildern aus geometrischen  
Objektmodellen

**Virtual Reality** kombinierte Darstellung realer und künstlicher  
Szenen, z.B. für Spiele, Filme, Anleitungen

**Computer Vision** (Bildverstehen, Bilderkennung)  
Transformation von Bildern in abstrakte  
Beschreibungen der im Bild dargestellten Szene,  
z.B. für Anwendungen in der Robotik

**Algorithmische Geometrie** Datenstrukturen und effiziente  
Algorithmen für typische geometrische Aufgaben  
(z.B. für Gestenerkennung,  
Geoinformationssysteme)

# Einordnung in die Informatik

**Informatik** Lehre von der Darstellung und Verarbeitung von **Information** durch **Algorithmen**

Teilgebiete der Informatik:

**theoretisch**

- ▶ Sprachen zur **Formulierung von Information** und Algorithmen,
- ▶ Möglichkeiten und Grenzen der Berechenbarkeit durch Algorithmen,
- ▶ Grundlagen für technische und praktische (und angewandte) Informatik

**technisch**

- ▶ **maschinelle Darstellung von Information**
- ▶ Mittel zur Ausführung von Algorithmen

(Rechnerarchitektur, Hardware-Entwurf, Netzwerk, ...)

**praktisch** Entwurf und Implementierung von **Datenstrukturen und Algorithmen**

(Betriebssysteme, Compilerbau, SE, ...)

**angewandt** Anwendung von Algorithmen:  
Text- und **Bildverarbeitung**, Datenbanken, **KI**,  
Wirtschafts-, **Medizin-**, **Bio-**, **Medieninformatik**, ...

# Inhalte der Lehrveranstaltung

- ▶ digitale Bilder
- ▶ statistische Merkmale
- ▶ Punktoperationen (Farb-, Helligkeits-, Kontraständerungen)
- ▶ geometrische Operationen (Projektionen, Verschiebungen, Verzerrungen, Koordinatentransformation)
- ▶ Spektrum, Fourier-Transformation
- ▶ Lokale Operationen, Filter (Hochpass-, Tiefpass-)
- ▶ Merkmalerkennung: Ecken, Kanten, Kurven
- ▶ morphologische Operationen, Skelettierung
- ▶ Segmentierung, Objekterkennung, Klassifikation

# Organisation

- Vorlesung:**
- ▶ 3 Wochen je 2 V
  - ▶ 6 Wochen je 1 V + 1 P
  - ▶ 3 Wochen je 2 P Anwendungsprojekte
  - ▶ 2 Wochen je 2 V (incl. Projektpräsentationen und Besprechung der Übungsaufgaben)

**Übungsaufgaben** zur Prüfungsvorbereitung:  
gelegentlich Übungsserien als Hausaufgaben,  
Besprechung in den letzten Vorlesungen

- Praktikum:** ab Ende April
- ▶ zunächst: Praktikum-Aufgaben (ImageJ)
  - ▶ nach Pfingsten: Arbeit an Anwendungsprojekten

**Prüfungsvorleistung:** zwei „Projekte“:

1. Praktikum-Aufgabenserien
2. Anwendungsprojekt mit Selbststudium-Anteil (Recherche, Entwicklung, Dokumentation, Präsentation in letzten LV-Wochen)

**Prüfung:** Klausur (120 min)  
Aufgabentypen wie Übungsaufgaben  
(Notenverbesserung durch Projekt-Bonus möglich)

# Bilder

**Bild** ist Funktion  $B : \text{pos} \rightarrow \text{col}$  mit

**pos** : Menge der **Positionen** im Bild

**col** : Menge der **Werte (Farben)** im Bild mit verschiedenen Bedeutungen, z.B.

- ▶ Intensitäten (Grauwerte),
- ▶ Tupel von Farbtintensitäten,
- ▶ Adressen (Hashwerte) in Lookup-Tabellen,
- ▶ Tiefen-Information (Abstand)

**reale** Bilder:  $\text{pos} \subseteq \mathbb{R}^n$  und  $\text{col} \subseteq \mathbb{R}^{\mathbb{R}}$

**unendliche stetige** Bereiche

**digitale** Bilder:

$\text{pos}$  und  $\text{col}$  **endliche diskrete** Bereiche  $\subseteq \mathbb{N}^n$

$\text{pos}$  meist Rechteckgitter

$\{0, \dots, m_1\} \times \dots \times \{0, \dots, m_n\}$ , z.B.

$\text{pos} = \{0, 1, 2\} \times \{0, \dots, 3\}$

(0, 0)	(0, 1)	(0, 2)	(0, 3)
(1, 0)	(1, 1)	(1, 2)	(1, 3)
(2, 0)	(2, 1)	(2, 2)	(2, 3)

# Digitale 2D-Bilder

digitales Bild (2D) mit rechteckigem Positionen-Raster

$$\text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$$

(Höhe  $m$ , Breite  $n$ ),

kann betrachtet werden als

1. Funktion  $B : \text{pos} \rightarrow \text{col}$   
(ordnet jeder Position einen Farbwert zu)
2. 2D-Matrix  $B \in \text{col}^{m \times n}$   
(mit Farbwerten als Einträgen)

**Pixel** (picture element):

Eintrag in dieser Matrix mit seiner Position,  
Adresse: (Zeile, Spalte)

z.B. Pixel an Adresse (1, 2) hat Wert 1 im Bild  $B$

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \quad B(1,2) = 1$$

# Farbwertbereiche

**Schwarz-Weiß-Bilder** mit Werten aus  $\text{col} = \{0, 1\}$   
meist 0 für schwarz und 1 für weiß

**Grauwertbilder** mit Intensitäten (Helligkeiten) aus  
 $\text{col} = \{0, \dots, k\} \subseteq \mathbb{N}$   
meist 0 für schwarz und  $k$  für maximale Intensität,  
oft  $k = 2^n - 1$  für ein  $n \in \mathbb{N}$ , meist  $n = 8$ , also  
Intensitäten  $\{0, \dots, 255\}$   
(eindimensional)

**Farbbilder** mit  $k$  Farbkomponenten (Spektrum)  
meist  $\text{col} \subseteq \mathbb{N}^k$  oder  $\text{col} \subseteq [0, 1]^k \subseteq \mathbb{R}^k$   
( $k$ -dimensional, Koordinaten bedeuten  
Farbintensitäten)  
Beispiel RGB:  $\text{col} \subseteq \{0, \dots, 255\}^3$

# Auflösungen

Ortsauflösung |pos| (Rasterung):

Anzahl der Positionen (Zeilen und Spalten)

oft **hohe Auflösung notwendig** zum Erkennen der dargestellten Szene

Farb-, Kontrastauflösung |col| (Quantisierung):

Anzahl der Farbwerte,

oft **geringe Auflösung ausreichend** zum Erkennen der dargestellten Szene

# Menschliches Sehen

Auge: Linse bildet Original auf Netzhaut (Retina) ab

Netzhaut enthält Fotorezeptoren:

- ▶  $130 \cdot 10^6$  **Stäbchen** registrieren **Intensität**  
(  $< 100$  Grauwerte,  $2^8 = 256$  Grauwerte genügen für realistischen visuellen Eindruck)
- ▶  $6 \cdot 10^6$  **Zapfen** registrieren **Farben** ( $7 \cdot 10^6$  Farbtöne)  
je ein Zapfentyp für Grün, Rot, Blau (in absteigender Empfindlichkeit)  
Durch Mischung der drei RGB-Farbkanäle  
(Spektralbereiche) lässt sich redundanzarm ein **visueller Eindruck** erzeugen, der dem natürlich wahrgenommenen Bild recht ähnlich ist.

Für Bilder zur automatischen Auswertung sind oft auch andere Spektralbereiche und Auflösungen sinnvoll.

Das im Auge empfangene diskrete Signal wird im Gehirn zu einem stetigen Gesamteindruck interpoliert

# Farben in Kamerasystemen

Jeder Sensor registriert einen festen Spektralbereich (nur Intensität in diesem Bereich).

Multispektralkamera enthält mehrere Sensoren für verschiedene Spektralbereiche

Ausgabe der Intensitäten meist diskret, also  $\subseteq \mathbb{N}$

Farbbereich mit  $k$  Farbkomponenten:  $\text{col} \subseteq \mathbb{N}^k$

Zuordnung von (Sensor-)Werten zu Farben oft in **Lookup-Tabellen** (LUT)

# Farbmodelle

Farbmischung:

**additiv** Kombination von Lichtquellen verschiedener Farben

z.B. RGB (Rot, Grün, Blau)

$$\text{col} = \{0, \dots, 255\}^3 = \{0, \dots, 2^8 - 1\}^3$$

zur Erzeugung realitätsnaher visueller Eindrücke im Display

**subtraktiv** Kombination von Filtern für verschiedene Farben  
z.B. CMYK (Cyan, Magenta, Yellow, Key)

$$\text{col} = \{0, \dots, 255\}^4 = \{0, \dots, 2^8 - 1\}^4$$

zur Erzeugung realitätsnaher visueller Eindrücke beim Druck

für wissenschaftliche Bildverarbeitung auch andere (größere, feinere) Spektralbereiche

- ▶ Falschfarben
- ▶ Kontrasterhöhung
- ▶ Überlagerung verschiedener Informationen

# Was bisher geschah

- ▶ Ziele maschineller Bildverarbeitung
  - ▶ automatische Informationsextraktion aus Bildern
  - ▶ Bildaufbereitung zur Unterstützung der visuellen Informationsextraktion
- ▶ Anwendungsbereiche
- ▶ Aufgaben maschineller Bildverarbeitung:  
(Erzeugung, Wiedergabe,)  
Transformation, Analyse, Interpretation digitaler Bilder
- ▶ digitale Bilder:
  - ▶ Funktion  $B : \text{pos} \rightarrow \text{col}$
  - ▶ Matrix  $B \in \text{col}^{\text{pos}}$

mit den Mengen

**pos** von Positionen (Adressen), meist  
 $\{0, \dots, m\} \times \{0, \dots, n\}$

**col** von Farben, Intensitäten, Tupeln von Intensitäten

- ▶ Orts- und Kontrastauflösung
- ▶ Farbmodelle

# Statistische Bildmerkmale (von Grauwertbildern)

Merkmale der **Menge** aller im Bild(-ausschnitt) vorkommenden Intensitäten (Grauwerte)

- ▶ Extremwerte
- ▶ Mittelwerte
- ▶ Kontrast
- ▶ Helligkeit
- ▶ Histogramme
- ▶ kumulative Histogramme
- ▶ Entropie

Merkmale der **Anordnung** aller im Bild(-ausschnitt) vorkommenden Intensitäten

- ▶ Linien- und Bereichsprofile
- ▶ Co-occurrence-Matrix (Grauwerteübergangs-Matrix)

Was lässt sich daraus über ein Bild schließen?

# Extrem- und Mittelwerte

gegeben:  $B \in \text{col}^{\text{pos}}$  (und evtl. Auswahl  $\text{pos}' \subseteq \text{pos}$ )

geringste / größte Intensität  $\min, \max : \text{col}^{\text{pos}} \rightarrow \text{col}$

$$\min(B) = \min\{B(p) \mid p \in \text{pos}\} \quad \max(B) = \max\{B(p) \mid p \in \text{pos}\}$$

größter Intensitätsunterschied  $\text{maxdiff} : \text{col}^{\text{pos}} \rightarrow \text{col}'$

$$\text{maxdiff}(B) = \max(B) - \min(B)$$

(Kontrastumfang des Bildes)

# Mittelwerte

gegeben:  $B \in \text{col}^{\text{pos}}$  (und evtl.  $\text{pos}' \subseteq \text{pos}$ )

durchschnittliche Intensität  $\text{avg} : \text{col}^{\text{pos}} \rightarrow \text{col}'$

$$\text{avg}(B) = \frac{\sum_{p \in \text{pos}} B(p)}{|\text{pos}|}$$

Gesamtheitigkeit des Bildes

**Median** (Zentralwert)  $\text{med} : \text{col}^{\text{pos}} \rightarrow \text{col}$ :

Wert am mittleren Index der sortierten Folge aller Bildwerte  $B(p)$

# Abweichungen

gegeben:  $B \in \text{col}^{\text{pos}}$  (und evtl. Auswahl  $\text{pos}' \subseteq \text{pos}$ )

mittlere quadratische Abweichung  $\sigma^2 : \text{col}^{\text{pos}} \rightarrow \mathbb{R}$

$$\sigma^2(B) = \frac{1}{|\text{pos}|} \sum_{p \in \text{pos}} (B(p) - \text{avg}(B))^2$$

und Standardabweichung  $\sigma$

Aussagen über den Kontrastumfang des Bildes

# Histogramme

gegeben:  $B \in \text{col}^{\text{pos}}$  (und evtl. Auswahl  $\text{pos}' \subseteq \text{pos}$ )

absolute Häufigkeit  $H : (\text{col}^{\text{pos}} \times \text{col}) \rightarrow \mathbb{N}$  der Farbe  $c \in \text{col}$  im Bild  $B$ :

$$H(B, c) = |\{p \in \text{pos} \mid B(p) = c\}|$$

**Histogramm  $H$**  :  $\text{col}^{\text{pos}} \times \text{col} \rightarrow \mathbb{N}$  mit

$$\forall B \in \text{col}^{\text{pos}} \forall c \in \text{col} : H(B, c) = |\{p \in \text{pos} \mid B(p) = c\}|$$

(Häufigkeiten der Intensitäten im Bild)

Histogramm ist unabhängig vom Ort der Intensitäten

**kumulatives Histogramm  $\bar{H}$**  :  $\text{col}^{\text{pos}} \times \text{col} \rightarrow \mathbb{N}$  mit

$$\forall B \in \text{col}^{\text{pos}} \forall c \in \text{col} : \bar{H}(B, c) = |\{p \in \text{pos} \mid B(p) \leq c\}|$$

# Informationen aus dem Histogramm

- ▶ alle im Bild vorkommenden Intensitäten
- ▶ Intensitäts-Extrema
- ▶ häufig, selten vorkommende Intensitäten
- ▶ mittlere Intensität (Durchschnitt, Median)
- ▶ Kontrastumfang
- ▶ Standard-Abweichung

Aussagen über Belichtung, Ausleuchtung, Dynamik

# Histogramm-Formen

häufige Formen der Funktion  $H$ :

**ausgeglichen** etwa Gleichverteilung aller Intensitäten  
wirkt oft recht hell

**bimodal** zwei deutlich voneinander getrennte lokale  
Maxima  
günstig für Segmentierung  
( Objekt-Hintergrund-Trennung)

# Intensitätsprofile

Profil:

Darstellung der Intensität als „Höhe“ über der Profillinie bzw. Profilfläche (Teilmenge  $\text{pos}' \subseteq \text{pos}$ )

**Linienprofil:** Intensitätsverlauf entlang einer Linie  
(Zeile, Spalte, Gerade, Kurve)  
scharfe Intensitätsunterschiede im Bild  
entsprechen steilen Kanten im Linienprofil

**Integriertes Linienprofil:** Summe der Intensitätsverläufe entlang  
mehrerer benachbarter Linien (Zeilen, Spalten)  
zum schnellen Vergleich von Bildern  
Lokalisierung von Abweichungen durch  
Kombination von Zeilen- und Spaltenprofilen

**Profil über Bildbereich:** 3D-Darstellung der Intensitäten als  
über einem zweidimensionalen Bildbereich

# Co-occurrence-Matrix

gegeben:  $B \in \text{col}^{\text{pos}}$  (und evtl. Auswahl  $\text{pos}' \subseteq \text{pos}$ )

Co-occurrence-Matrix  $C \in \mathbb{N}^{\text{col} \times \text{col}}$  (horizontale Nachbarschaft) mit

$$\forall c, d \in \text{col} : C(c, d) = |\{p = (x, y) \in \text{pos} : B(p) = c \wedge B(x, y + 1) = d\}|$$

Anzahl der Vorkommen von Intensitäts-Paaren zwischen  
Nachbar-Pixeln  
(verschiedene Nachbarschafts-Beziehungen)

Beispiel:

$$B = \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 & 3 & 3 \\ 0 & 1 & 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 3 & 3 & 3 \\ 2 & 2 & 3 & 3 & 3 & 3 \end{pmatrix} \quad C(B) = \begin{pmatrix} 4 & 4 & 0 & 0 \\ 0 & 2 & 5 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

## Co-occurrence-Matrix – Informationen

- ▶ große gleichfarbige Flächen in  $B$   
hohe Werte auf der Hauptdiagonale in  $C(B)$
- ▶ scharfe Kontraste in  $B$   
hohe Werte links unten und rechts oben  $C(B)$
- ▶ Länge von Grenzen zwischen Bereichen gleicher Intensitäten
- ▶ Texturmerkmal (zur texturbasierten Segmentierung)

# Entropie

gegeben:  $B \in \text{col}^{\text{pos}}$  (und evtl. Auswahl  $\text{pos}' \subseteq \text{pos}$ )

Entropie  $S : \text{col}^{\text{pos}} \rightarrow \mathbb{R}$

$$S(B) = - \sum_{c \in \text{col}} (h(B, c) \log_2(h(B, c)))$$

mit relativer Häufigkeit  $h : (\text{col}^{\text{pos}} \times \text{col}) \rightarrow [0, 1]$ :

$$h(B, c) = \frac{|\{p \in \text{pos} \mid B(p) = c\}|}{|\text{pos}|}$$

Aussage über **Informationsgehalt** des Bildes

(oder der Co-occurrence-Matrix oder anderer Bildmerkmale)

Beispiele:

- ▶ einfarbiges  $(256 \times 256)$ -Bild  $B$  mit Grauwert 10 hat Entropie  $S(B) = \dots$
- ▶  $(2^8 \times 2^8)$ -Schwarz-Weiß-Bild mit gleicher Anzahl schwarzer und weißer Pixel hat Entropie  $S(B) = \dots$
- ▶  $(2^8 \times 2^8)$ -Bild mit jeweils gleicher Anzahl an Pixeln in 8 Intensitätswerten hat Entropie  $S(B) = \dots$

untere Schranke für Bit-Anzahl zur Codierung des Bildes

(bietet Information zur verlustfreien Komprimierbarkeit)

# Zentralmomente

Berechnung des Mittelwertes  $\text{avg} : \text{col}^{\text{pos}} \rightarrow \mathbb{R}$  aus relativen Häufigkeiten:

$$\text{avg}(B) = \sum_{c \in \text{col}} (c \cdot h(B, c))$$

Zentralmomente  $m_n : \text{col}^{\text{pos}} \rightarrow \mathbb{R}$

$$m_n(B) = \sum_{c \in \text{col}} ((c - \text{avg}(B))^n h(B, c))$$

Spezialfälle:

$m_1(B)$  = avg( $B$ ) Mittelwert

$m_2(B)$  Varianz ( $m_2(B) = 0$  für einfarbige Bilder  $B$ )

$m_3(B)$  Schiefe, Grad der Asymmetrie

$m_4(B)$  Wölbung (Exzess)

hilfreich zur Texturanalyse

# Was bisher geschah

- ▶ Aufgaben maschineller Bildverarbeitung:  
Transformation, Analyse, Interpretation  
(, Erzeugung, Wiedergabe) digitaler Bilder

- ▶ digitale Bilder:

- ▶ Funktion  $B : \text{pos} \rightarrow \text{col}$
- ▶ Matrix  $B \in \text{col}^{\text{pos}}$

mit den Mengen

- ▶ pos von Positionen (Adressen)
- ▶ col von Farben, Intensitäten

- ▶ Orts- und Kontrastauflösung

- ▶ Farbmodelle

- ▶ statistische Merkmale (Bildanalyse)

- ▶ Bild  $\mapsto$  Wert: Extrema, Mittelwerte, Abweichung, Entropie
- ▶ Bild  $\mapsto$  Funktion  $\text{col} \rightarrow \mathbb{N}$ :  
Histogramme, kumulierte Histogramme,
- ▶ Linien- und Bereichsprofile, Co-occurrence-Matrix

und ihre Aussagen über das Bild

# Bildbearbeitung durch Punktoperationen

**Punktoperation** zwischen verschiedenen Farbbereichen

Funktion  $f : \text{col}_1 \rightarrow \text{col}_2$

(neuer) Wert **eines** Pixels

nur abhängig vom (alten) Wert **dieses** Pixels

Fortsetzung  $f : \text{col}_1^{\text{pos}} \rightarrow \text{col}_2^{\text{pos}}$  der Funktion  $f$  auf Bilder durch **punktweise** Anwendung

$$\forall p \in \text{pos} : f(B)(p) = f(B(p))$$

Beispiel:

$$f : \{3, \dots, 6\} \rightarrow \{0, \dots, 7\} \quad \text{mit} \quad f(x) = \frac{7(x-3)}{3}$$

$$f \begin{pmatrix} 4 & 5 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} f(4) & f(5) \\ f(3) & f(4) \end{pmatrix} = \begin{pmatrix} \frac{7}{3} & \frac{14}{3} \\ 0 & \frac{7}{3} \end{pmatrix} \approx \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}$$

## Punktoperationen – Beispiele

- ▶ Transformation von RGB ( $\text{col}_1 = \{0, \dots, 255\}^3$ ) in Graustufenintensität ( $\text{col}_2 = \{0, \dots, 255\}$ )  
 $f : \{0, \dots, 255\}^3 \rightarrow \{0, \dots, 255\}$  mit  
 $\forall (r, g, b) \in \text{col}_1 : f(r, g, b) = (r + g + b)/3$   
bzw.  $\forall (r, g, b) \in \text{col}_1 : f(r, g, b) = 0.299r + 0.587g + 0.114b$
- ▶ Umrechnung 8-Bit-Intensität ( $\text{col}_1 = \{0, \dots, 255\}$ ) in relative Intensität ( $\text{col}_2 = [0, 1]$ ) in  
 $f : \{0, \dots, 255\} \rightarrow [0, 1]$  mit  $\forall c \in \text{col}_1 : f(c) = c/255$
- ▶ Falschfarbendarstellung von Intensitäten (z.B. für Infrarot-, Tiefeninformation)

$$f : [0, 1] \rightarrow [0, 1]^3 \quad \text{mit} \quad f(x) = \left( \underbrace{\sqrt{x}}_r, \underbrace{x^3}_g, \underbrace{\sin(2\pi x)}_b \right)$$

- ▶ nachträgliche Colorierung von Grauwertbildern  
z.B. <http://richzhang.github.io/colorization>
- ▶ Transformationen für  $\text{col}_1 = \text{col}_2$  (gleicher Farbbereich)  
meist zur Kontrasterhöhung oder Invertierung

# Prominente Punktoperationen

z.B. zur Bildverbesserung durch Kontrasterhöhung

- ▶ lineare Funktionen (Kontrast- und Helligkeitsänderung)

$$f(c) = ac + b \quad f(c) = \max(\text{col}) - ac + b$$

- ▶ Stufenfunktionen (stückweise konstant),  
Abbildung auf weniger Werte  
Spezialfall Binarisierung mit Schwellwert  $\theta \in \text{col}$

$$f(c) = \begin{cases} 1 & \text{falls } c \geq \theta \\ 0 & \text{sonst} \end{cases}$$

- ▶ weitere nichtlineare Funktionen, z.B.

$$f(c) = \frac{c^2}{\max(\text{col})} \quad f(c) = \sqrt{c \max(\text{col})}$$

# Histogrammspreizung (-normalisierung)

gegeben:  $B \in \text{col}^{\text{pos}}$  ( und evtl. Auswahl  $\text{pos}' \subseteq \text{pos}$ )

punktweise Anwendung der (linearen) Punktoperation  
 $f : \text{col} \rightarrow \text{col}$  mit

$$f(c) = (c - \min(B)) \frac{\max(\text{col}) - \min(\text{col})}{\max(B) - \min(B)} + \min(\text{col})$$

sorgt für volles Ausschöpfen der gesamten Intensitätsskala

Clipping:

Ignorieren (häufig informationsarmer) Bereiche  
sehr geringer ( $< \theta_1$ ) und sehr hoher ( $> \theta_2$ ) Intensitäten durch  
Histogrammspreizung des Farbbereiches  $[\theta_1, \theta_2]$

# Histogrammausgleich (-äqualisierung)

gegeben:  $B \in \text{col}^{\text{pos}}$ ,  
Auswahl  $\text{pos}' \subseteq \text{pos}$  von Positionen im Bild  $B$

Punktoperation  $f : \text{col} \rightarrow \text{col}$  mit

$$f(c) = (\max(\text{col}) - \min(\text{col})) \frac{|\{p \in \text{pos} \mid B(p) \leq c\}|}{|\text{pos}|}$$

sorgt für

- ▶ annähernd konstantes Histogramm
- ▶ annähernd lineares kumulatives Histogramm
- ▶ ähnlich häufiges Auftreten aller Intensitäten

# Posterizing

Abbildung auf vorgegebene Anzahl  $k'$  von Intensitätsstufen  
(häufig geringer als Anzahl der Intensitätsstufen des  
Originalbildes)

gegeben:  $B \in \{0, \dots, k\}^{\text{pos}}$

gesucht:  $B' \in \{0, \dots, k'\}^{\text{pos}}$

durch Stufenfunktion mit  $k'$  Schwellwerten  $\theta_i \in \{1, \dots, k'\}$

$$f(x) = \begin{cases} 0 & \text{falls } x < \theta_1 \\ 1 & \text{falls } \theta_1 \leq x < \theta_2 \\ \vdots & \\ k' & \text{falls } \theta_{k'-1} \leq x \leq \theta_{k'} \end{cases}$$

# Binarisierung

Abbildung von Bildern mit Wertebereich  $\text{col} = \{0, \dots, k\}$   
(Intensitäten)

auf Schwarz-weiß-Bilder (mit Wertebereich  $\text{col}' = \{0, 1\}$ )

gegeben:  $B \in \{0, \dots, k\}^{\text{pos}}$

gesucht:  $B' \in \{0, 1\}^{\text{pos}}$

abhängig von einem Schwellwert  $\theta \in \text{col}$

$$f(x) = \begin{cases} 1 & \text{falls } x \geq \theta \\ 0 & \text{sonst} \end{cases}$$

oder abhängig von zwei Schwellwerten  $\theta_1, \theta_2 \in \text{col}$

$$f(x) = \begin{cases} 1 & \text{falls } \theta_1 < x \leq \theta_2 \\ 0 & \text{sonst} \end{cases}$$

# Schwellwertbestimmung

- ▶ mittlerer Grauwert  $\theta = \text{avg}(B)$
- ▶ bei bekanntem prozentuaem Anteil  $p\%$  der (angenommen helleren) Objektpunkte:  
Auswahl des maximalen  $\theta$  mit  $\sum_{c < \theta} h(B, c) \leq 1 - p/100$
- ▶ bei bimodalen Histogrammen:  
 $\theta =$  lokales Minimum zwischen beiden Maxima
- ▶ Schwellwertberechnung nach Otsu

## **adaptive** Schwellwertberechnung

verschiedene Schwellwerte für verschiedene Bildbereiche anhand lokaler Merkmale (z.B. avg, med) der Bildbereiche

# Schwelwertbestimmung nach Otsu

- ▶ Jeder Schwellwert  $\theta \in \text{col}$  zerlegt  $\text{col}$  in zwei Farbbereiche.
  - ▶  $\text{col}_{<\theta} = \{c \in \text{col} \mid c < \theta\}$  (Hintergrund)
  - ▶  $\text{col}_{\geq\theta} = \{c \in \text{col} \mid c \geq \theta\}$  (Vordergrund)

- ▶ statistische Größen dieser beiden Farbbereiche innerhalb  $B$ :

- ▶ Auftrittswahrscheinlichkeit

$$p_{<\theta}(B) = \sum_{c \in \text{col}_{<\theta}} h(B, c) = \frac{|\{p \in \text{pos} \mid B(p) < \theta\}|}{|\text{pos}|}$$

- ▶ Mittelwert  $\text{avg}_{<\theta}(B) = \frac{1}{p_{<\theta}(B)} \sum_{c < \theta} c \cdot h(B, c)$

- ▶ Varianz  $\sigma_{<\theta}^2(B) = \frac{1}{p_{<\theta}(B)} \sum_{c < \theta} (h(B, c)(c - \text{avg}_{<\theta}(B))^2)$

- ▶ gesucht:  $\theta$  mit beiden folgenden Eigenschaften

- ▶ minimale Varianz innerhalb der Klassen

$$\sigma_{\theta, W}^2(B) = p_{<\theta}(B)\sigma_{<\theta}^2(B) + p_{\geq\theta}(B)\sigma_{\geq\theta}^2(B)$$

- ▶ maximale between-class-Varianz

$$\sigma_{\theta, B}^2(B) = p_{<\theta}(B) (\sigma_{<\theta}^2(B) - \text{avg}(B))^2 + p_{\geq\theta}(B) (\sigma_{\geq\theta}^2(B) - \text{avg}(B))^2$$

- ▶ Für jeden Schwellwert  $\theta$  ist  $\sigma_{\theta, W}^2(B) + \sigma_{\theta, B}^2(B) = \sigma^2(B)$  konstant.  
Es existiert ein  $\theta$ , so dass  $\sigma_{\theta, W}^2(B)$  minimal und  $\sigma_{\theta, B}^2(B)$  maximal.

# Kombinationen mehrerer Bilder

Eingabe: Bilder  $B, C \in \text{col}^{\text{pos}}$

Ausgabe: Bild  $D \in \text{col}^{\text{pos}}$

Beispiele für punktweise Operationen  $f : \text{col} \times \text{col} \rightarrow \text{col}$  auf

**Binärbildern** ( $\text{col} = \{0, 1\}$ ): logische Operationen  
(Mengenoperationen)

**Grauwertbilder** ( $\text{col} = \{0, \dots, k\}$ ): arithmetische Operationen  
z.B. min, max, Differenz, Mittelwert

**Farbbildern**

- ▶ Grauwertbild-Operationen auf jeder Farbkomponente einzeln
- ▶ Transformation in andere Farbbereiche (z.B. Grauwerte) durch Operationen zur Kombination mehrerer Farbkomponenten

# Binärbilder als Mengen von Positionen

- ▶ Bild  $B : \text{pos} \rightarrow \{0, 1\}$   
ist charakteristische Funktion der Menge aller weißen Pixel  
(meist als Vordergrund interpretiert)
- ▶ Bild  $B$  repräsentiert die Menge  $B^{-1}(1)$  (Urbild von 1)  
 $B \sim B^{-1}(1)$
- ▶ Transformation von  $B$  durch Anwendung von  
Mengenoperationen auf  $B$

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline \end{array} \sim \left\{ \begin{array}{l} (0, 1), (0, 2), (0, 3), \\ (1, 2), (1, 3), (2, 2), \\ (2, 3), (3, 1), (4, 1) \end{array} \right\}$$

# Mengenoperationen auf Binärbildern

**Komplement** eines Bildes  $B : \text{pos} \rightarrow \{0, 1\}$ :  
 $\bar{B} : \text{pos} \rightarrow \{0, 1\}$  mit  $\bar{B} \sim \overline{B^{-1}(1)} = B^{-1}(0)$   
punktweise:

$$\forall p \in \text{pos} : \bar{B}(p) = \neg B(p) = 1 - B(p)$$

(Farb-Invertierung)

**Vereinigung** zweier Bilder  $B, C : \text{pos} \rightarrow \{0, 1\}$   
 $(B \cup C) : \text{pos} \rightarrow \{0, 1\}$  mit  
 $(B \cup C) \sim B^{-1}(1) \cup C^{-1}(1) = \{p \in \text{pos} \mid B(p) = 1 \vee C(p) = 1\}$   
punktweise:

$$\forall p \in \text{pos} : (B \cup C)(p) = B(p) \vee C(p) = \max(B(p), C(p))$$

**Schnitt** zweier Bilder  $B, C : \text{pos} \rightarrow \{0, 1\}$   
 $(B \cap C) : \text{pos} \rightarrow \{0, 1\}$  mit  
 $(B \cap C) \sim B^{-1}(1) \cap C^{-1}(1) = \{p \in \text{pos} \mid B(p) = 1 \wedge C(p) = 1\}$   
punktweise:

$$\forall p \in \text{pos} : (B \cap C)(p) = B(p) \wedge C(p) = \min(B(p), C(p))$$

# Mengenoperationen auf Binärbildern

**Differenz** zweier Bilder  $B, C : \text{pos} \rightarrow \{0, 1\}$

$(B \setminus C) : \text{pos} \rightarrow \{0, 1\}$  mit

$$\begin{aligned}(B \setminus C) &\sim B^{-1}(1) \setminus C^{-1}(1) \\ &= \{p \in \text{pos} \mid B(p) = 1 \wedge C(p) = 0\}\end{aligned}$$

punktweise:

$$\forall p \in \text{pos} : (B \setminus C)(p) = B(p) \wedge \neg C(p) = \min(B(p), 1 - C(p))$$

**Symmetrische Differenz** zweier Bilder  $B, C : \text{pos} \rightarrow \{0, 1\}$

$(B \Delta C) : \text{pos} \rightarrow \{0, 1\}$  mit

$$\begin{aligned}(B \Delta C) &\sim (B \cup C) \setminus (B \cap C) = (B \setminus C) \cup (C \setminus B) \\ &= \{p \in \text{pos} \mid B(p) = 1 \text{ XOR } C(p) = 0\}\end{aligned}$$

punktweise:

$$\begin{aligned}\forall p \in \text{pos} : (B \Delta C)(p) &= B(p) \text{ XOR } C(p) \\ &= (B(p) + C(p)) \pmod 2\end{aligned}$$

# Arithmetische Operationen auf Grauwertbildern

gegeben:  $B, C \in \text{col}^{\text{pos}}$  mit  $\text{col} = \{0, \dots, k\} \subset \mathbb{N}$

typische Punktoperationen:

$\forall B, C \in \text{col}^{\text{pos}} \forall p \in \text{pos}$

- ▶ Differenz  $(B - C)(p) = B(p) - C(p)$   
zum Vergleich von Bildern, Hintergrund-Entfernung,  
Bewegungsdetektion  
häufig auch  $\max(B(p) - C(p), 0)$  zur Vermeidung negativer  
Farbwerte
- ▶ Mittelwert mehrerer Bilder  $\text{avg}(B_1, \dots, B_n)(p) = \frac{\sum_{i=1}^n B_i(p)}{n}$   
zur Entfernung zufälliger Störungen, Entrauschen
- ▶ Addition  $(B + C)(p) = B(p) + C(p)$   
häufig auch  $\min(B(p) + C(p), \max(\text{col}))$  zur Vermeidung  
von Werten außerhalb des Farbbereiches
- ▶ Multiplikation  $(BC)(p) = B(p)C(p)$   
Verhältnis  $(B/C)(p) = B(p)/C(p)$

# Was bisher geschah

- ▶ digitale Bilder:
  - ▶ Funktion  $B : \text{pos} \rightarrow \text{col}$
  - ▶ Matrix  $B \in \text{col}^{\text{pos}}$
- ▶ statistische Merkmale (Bildanalyse)
  - ▶ Bild  $\mapsto$  Wert: Extrema, Mittelwerte, Abweichung, Entropie
  - ▶ Bild  $\mapsto$  Funktion  $\text{col} \rightarrow \mathbb{N}$ : (kumulierte) Histogramme
  - ▶ Bild  $\mapsto$  Funktion von Teilmenge von  $\text{pos} \rightarrow \mathbb{N}$ :  
Linienprofile, integriertes Linienprofil
  - ▶ Bild  $\mapsto$  Funktion  $\text{col}^2 \rightarrow \mathbb{N}$ : Co-occurrence-Matrix

und ihre Aussagen über das Bild

- ▶ Punktoperationen  $f : \text{col}_1 \rightarrow \text{col}_2$  (Farbtransformation)  
und ihre Fortsetzung  $f : \text{col}_1^{\text{pos}} \rightarrow \text{col}_2^{\text{pos}}$

$$\forall p \in \text{pos} : f(B)(p) = f(B(p))$$

- ▶ Intensitäts-, Kontrasterhöhung
- ▶ Farbtransformationen
- ▶ Binarisierung
- ▶ Kombination mehrerer Bilder

# Nachbarschaftsbeziehungen zwischen Positionen

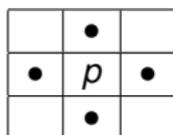
Zur Identifikation zusammenhängender Bildteile (Linien, Flächen) verschiedene Nachbarschaftsbeziehungen üblich, z.B.

Zwei **Positionen**  $p, q \in \text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$  heißen genau dann **benachbart**, wenn sie

1. ungleich sind und
2. wenigstens eine gemeinsame

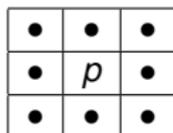
**Kante** besitzen, 4-Nachbarschaft von  $p = (z, s)$ :

$$N_4(p) = \{(z-1, s), (z+1, s), (z, s-1), (z, s+1)\} \cap \text{pos}$$



**Ecke** besitzen, 8-Nachbarschaft von  $p = (z, s)$ :

$$N_8(p) = \left\{ \begin{array}{l} (z-1, s-1), (z, s-1), (z+1, s-1), \\ (z-1, s), (z+1, s), (z-1, s+1), \\ (z, s+1), (z+1, s+1) \end{array} \right\} \cap \text{pos}$$



# Geometrische Operationen in der DBV

Anwendungen z.B. in

- ▶ Geoinformatik:
  - ▶ Satellitenbilder → geographische Daten
  - ▶ Bereinigung um Verzerrungen durch Höhenunterschiede
- ▶ Objekterkennung
- ▶ autonomes Fahren:
  - ▶ Kamerabild → Umgebungskarte
- ▶ Schrifterkennung:
  - ▶ Zeilenausrichtung
  - ▶ Normierung der Zeichengröße
  - ▶ Schrägschrift-Korrektur
- ▶ QR-Codes: Entzerrung
- ▶ Bildvergrößerung und -verkleinerung

# Geometrische Operationen

Operationen auf der Menge  $\text{pos}$ , wobei

- ▶ Intensitäten
- ▶ Nachbarschaftsbeziehungen zwischen Positionen

weitgehend unverändert bleiben

gegeben:  $B \in \text{col}^{\text{pos}_1}$

gesucht:  $C \in \text{col}^{\text{pos}_2}$

Transformationen:

**linear** Kombinationen von

- ▶ Spiegelung
- ▶ Verschiebung
- ▶ Skalierung (Zoom)
- ▶ Drehung

**nichtlinear** z.B. perspektivische Abbildungen, Warping, Swirl

# Transformationen im $\mathbb{R}^2$

Transformation  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

für digitale Bildverarbeitung

(Beschränkung auf Stützstellen im Raster):

$f : \text{pos} \rightarrow \text{pos}'$

hier mit rechteckigen Rastern (Adressen in Matrizen)

$\text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$  und

$\text{pos}' = \{0, \dots, m'-1\} \times \{0, \dots, n'-1\}$

meist Koordinatenverschiebung mit  $(0, 0)$  im Bildmittelpunkt

Probleme bei Transformationen  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ :

- ▶ Zielpositionen evtl. nicht in  $\text{pos}'$
- ▶ nicht jede Position in  $\text{pos}'$  ist Bild einer Position im Original
- ▶ berechnete Zielposition liegt i.A. nicht auf Rasterpunkt
- ▶ Festlegung des Farbwertes im transformierten Bild nötig (Interpolation)

## Translation (Verschiebung) im $\mathbb{R}^2$

Verschiebung (2D) um einen Vektor  $(v_1, v_2) \in \mathbb{R}^2$ :

$V_{(v_1, v_2)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$V_{(v_1, v_2)} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} x_1 + v_1 \\ x_2 + v_2 \end{pmatrix}$$

Verschiebung (2D) als Multiplikation mit Matrix  $V_{(v_1, v_2)} \in \mathbb{R}^{3 \times 3}$   
in **homogenen Koordinaten**:

$$V_{(v_1, v_2)} = \begin{pmatrix} 1 & 0 & v_1 \\ 0 & 1 & v_2 \\ 0 & 0 & 1 \end{pmatrix} \quad V_{(v_1, v_2)} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 + v_1 \\ x_2 + v_2 \\ 1 \end{pmatrix}$$

zu  $V_{(v_1, v_2)}$  inverse Transformation:  $V_{(v_1, v_2)}^{-1} = V_{(-v_1, -v_2)}$

# Spiegelung (Reflexion) im $\mathbb{R}^2$

Spiegelung (2D) an

$x_2$ -Achse  $S_h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$S_h \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 \\ x_2 \end{pmatrix}$$

Spiegelung an  $x_2$ -Achse als Multiplikation mit Matrix  $S_h \in \mathbb{R}^{3 \times 3}$  in homogenen Koordinaten:

$$S_h = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S_h \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} -x_1 \\ x_2 \\ 1 \end{pmatrix}$$

$x_1$ -Achse  $S_v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  analog:

$$S_v = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S_v \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \\ 1 \end{pmatrix}$$

Spiegelungen sind involutiv (selbstinvers):  $S_h^{-1} = S_h$ ,  $S_v^{-1} = S_v$

## Skalierung im $\mathbb{R}^2$

Skalierung (Stauchung, Streckung)

aus dem Ursprung  $(0, 0) \in \mathbb{R}^2$  mit

Skalierungsfaktoren  $s_1 \in \mathbb{R}$  und  $s_2 \in \mathbb{R}$

$S_{(s_1, s_2)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$\forall x \in \mathbb{R}^2 : S_{(s_1, s_2)}(x) = (s_1 x_1, s_2 x_2)$$

Skalierung (2D) als Matrixmultiplikation in homogenen Koordinaten:

$$S_{(s_1, s_2)} = \begin{pmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S_{(s_1, s_2)} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} s_1 x_1 \\ s_2 x_2 \\ 1 \end{pmatrix}$$

zu  $S_{(s_1, s_2)}$  inverse Transformation:  $S_{(s_1, s_2)}^{-1} = S_{(1/s_1, 1/s_2)}$

Spiegelungen sind Skalierungen:  $S_h = S'_{(-1, 1)}$  und  $S_v = S'_{(1, -1)}$

## Rotation (Drehung) im $\mathbb{R}^2$

Drehung um einen Winkel  $\varphi \in [0, 2\pi]$  im mathematischen Drehsinn (entgegen Uhrzeigersinn) um Ursprung  $(0, 0) \in \mathbb{R}^2$

$D_\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$\forall x \in \mathbb{R}^2 : D_\varphi(x) = (x_1 \cos \varphi - x_2 \sin \varphi, x_1 \sin \varphi + x_2 \cos \varphi)$$

Drehung (2D) als Matrix in homogenen Koordinaten:

$$D_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad D_\varphi \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \cos \varphi - x_2 \sin \varphi \\ x_1 \sin \varphi + x_2 \cos \varphi \\ 1 \end{pmatrix}$$

zu  $D_\varphi$  inverse Transformation:  $D_\varphi^{-1} = D_{-\varphi}$

# Affine Transformationen

(z.B. für schräg aufgenommenen Bilder Schriftstücke, QR-Codes)

Matrixdarstellung der linearen Transformationen ermöglicht die Zusammenfassung der Nacheinanderausführung mehrerer Transformationen (Matrixmultiplikation)

allgemeine Form affiner Transformationen im  $\mathbb{R}^2$   
(in homogenen Koordinaten):

$$M = \begin{pmatrix} a_1 & a_2 & v_1 \\ b_1 & b_2 & v_2 \\ 0 & 0 & 1 \end{pmatrix}$$

Jede Transformation (Matrix) dieser Form lässt sich durch eine endliche Verkettung der elementaren Transformationen (Matrix-Multiplikationen)

- ▶ Verschiebung
- ▶ Skalierung
- ▶ Drehung um Ursprung  $(0, 0) \in \mathbb{R}^2$

erzeugen.

für  $M = T_1 \circ \dots \circ T_n$  gilt  $M^{-1} = T_n^{-1} \circ \dots \circ T_1^{-1}$

# Beispiel

	$p$	
	$q$	

Koordinaten  $p = (0, 1)$  und  $q = (1, 1)$

Verschiebung um  $(-1, -1)$ , danach Drehung um  $\pi/4$ ,  
danach Skalierung mit  $(\sqrt{2}, \sqrt{2})$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Schrittweise Anwendung auf  $p = (0, 1)$  und  $q = (1, 1)$

Gesamttransformation:

Matrixmultiplikation in „umgekehrter“ Reihenfolge

$$\begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

# Was bisher geschah

- ▶ digitale Bilder:
  - ▶ Funktion  $B : \text{pos} \rightarrow \text{col}$
  - ▶ Matrix  $B \in \text{col}^{\text{pos}}$
- ▶ statistische Merkmale (Bildanalyse)  
und ihre Aussagen über das Bild
- ▶ Punktoperationen  $f : \text{col}_1 \rightarrow \text{col}_2$  (Farbtransformation)  
und ihre Fortsetzung  $f : \text{col}_1^{\text{pos}} \rightarrow \text{col}_2^{\text{pos}}$   
$$\forall p \in \text{pos} : f(B)(p) = f(B(p))$$
  - ▶ Intensitäts-, Kontrasterhöhung
  - ▶ Histogrammstreckung
  - ▶ Binarisierung
- ▶ Geometrische Operationen  $f : \text{pos}_1 \rightarrow \text{pos}_2$  (Koordinatentransformation)
  - ▶ Verschiebung, Skalierung, Drehung
  - ▶ affine Abbildungen

# Bestimmung affiner Transformationen

**gegeben:** Zuordnung  $\{(p_1, q_1), \dots, (p_n, q_n)\}$  zwischen Positionen  $\{p_1, \dots, p_n\} \in \text{pos}(B)$  im Bild  $B$  und  $\{q_1, \dots, q_n\} \in \text{pos}(A)$  im Original  $A$

**gesucht:** Beschreibung der Transformation  $A \mapsto B$  durch Matrix  $M$  mit  $\forall i \in \{1, \dots, n\} : M(q_i) = p_i$

$$M = \begin{pmatrix} a_1 & a_2 & v_1 \\ b_1 & b_2 & v_2 \\ 0 & 0 & 1 \end{pmatrix}$$

d.h. gesucht sind die sechs Koeffizienten  $a_1, a_2, b_1, b_2, v_1, v_2$  mit

$$\forall i \in \{1, \dots, n\} : M \begin{pmatrix} q_{i1} \\ q_{i2} \\ 1 \end{pmatrix} = \begin{pmatrix} p_{i1} \\ p_{i2} \\ 1 \end{pmatrix}$$

als lineares Gleichungssystem mit Unbekannten  $a_1, a_2, b_1, b_2, v_1, v_2$ :

$$\begin{aligned} q_{i1} a_1 + q_{i2} a_2 + v_1 &= p_{i1} \\ q_{i1} b_1 + q_{i2} b_2 + v_2 &= p_{i2} \end{aligned}$$

eindeutig lösbar für sechs Gleichungen  
= drei (nicht kollineare) Referenzpunkte

## Beispiel

Gesucht ist die affine Transformation  $M = \begin{pmatrix} a_1 & a_2 & v_1 \\ b_1 & b_2 & v_2 \\ 0 & 0 & 1 \end{pmatrix}$ , welche die folgenden Punkte aufeinander abbildet:

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Gleichungssystem

$$a_2 + v_1 = -1 \quad (1)$$

$$b_2 + v_2 = -1 \quad (2)$$

$$a_1 + a_2 + v_1 = 0 \quad (3)$$

$$b_1 + b_2 + v_2 = 0 \quad (4)$$

$$a_1 + v_1 = 2 \quad (5)$$

$$b_1 + v_2 = -1 \quad (6)$$

(3)-(1):  $a_1 = 1$ , (4)-(2):  $b_1 = 1, \dots$

$$M = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

# Projektive Abbildungen

(z.B. zur Lokalisierung autonomer Fahrzeuge aus perspektivische Abbildungen, Verkehrsszenen)

Projektive Abbildung mit

- ▶ affiner Transformation  $M = \begin{pmatrix} a_1 & a_2 & v_1 \\ b_1 & b_2 & v_2 \\ 0 & 0 & 1 \end{pmatrix}$  und
- ▶ Projektionskoeffizienten  $p_1, p_2$

Projektive Abbildung (2D) als Matrix in homogenen Koordinaten:

$$P_{M,p_1,p_2} = \begin{pmatrix} a_1 & a_2 & v_1 \\ b_1 & b_2 & v_2 \\ p_1 & p_2 & 1 \end{pmatrix}$$

$$P_{M,p_1,p_2} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 x_1 + a_2 x_2 + v_1 \\ b_1 x_1 + b_2 x_2 + v_2 \\ p_1 x_1 + p_2 x_2 + 1 \end{pmatrix} = \begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \end{pmatrix}$$

Umrechnung nach  $\mathbb{R}^2$ :  $y_1 = y'_1/y'_3$ ,  $y_2 = y'_2/y'_3$  (nichtlinear)

# Bestimmung projektiver Transformationen

typischer Ausgangspunkt:

- ▶ (projektiv) verzerrtes Bild
- ▶ Menge von Referenzpunkten, für die Position im Ursprungs- und im Zielbild bekannt sind

gesucht:

Transformationsmatrix zur Entzerrung des Bildes  
(8 Koeffizienten  $a_1, a_2, b_1, b_2, v_1, v_2, p_1, p_2$ )

Gleichungssystem mit 8 Unbekannten  
also wenigstens 4 Referenzpunkte nötig

Auflösen der Gleichungen

$$y_1 = \frac{a_1 x_1 + a_2 x_2 + v_1}{p_1 x_1 + p_2 x_2 + 1} \quad y_2 = \frac{b_1 x_1 + b_2 x_2 + v_2}{p_1 x_1 + p_2 x_2 + 1}$$

für jeden Referenzpunkt  $y = (y_1, y_2)$  nach den Koeffizienten

# Rasterung

Abbildung eines Bildes  $B \in \text{col}^{\mathbb{R}^2}$  mit stetiger Positionsmenge (Ebene  $\mathbb{R}^2$ ) auf ein digitales Bild  $C \in \text{col}^{\text{pos}}$  mit diskreter Menge von Positionen  $\text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$  (Rasterpunkte) besteht aus zwei Teilaufgaben:

1. Abbildung  $\mathbb{R}^2 \rightarrow \text{pos}$
2. Bestimmung der Farbwerte  $C(p)$  für Rasterpunkte  $p \in \text{pos}$ , verschiedene Möglichkeiten
  - ▶  $C(p) := B(p)$  (Farbwert an Stützstelle  $p$ )  
nur möglich, wenn Farbwert  $B(p)$  für  $p \in \text{pos}$  gegeben ist
  - ▶  $C(p) := f(U(p))$  als Funktion aller Farben in einer geeignet definierten Umgebung  $U(p) \subset \mathbb{R}^2$  von  $p$   
(z.B. Mittelwert  $C(p) := \frac{\sum_{q \in U(p)} B(q)}{|U(p)|}$ )

# Interpolation der Intensitäten

Das geometrische Transformationen nicht immer Pixel auf Pixel abbilden, wird der Farbwert eines Pixels  $q \in \text{pos}'$  im transformierten Bild häufig aus den Farbwerten einer Umgebung (Menge von Pixeln) von  $T^{-1}(q) \in \text{pos}$  im Original berechnet (interpoliert).

gegeben: digitales (Original-)Bild  $B \in \text{col}^{\text{pos}}$ ,  
geometrische Transformation  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

gesucht: transformiertes digitales Bild  $C \in \text{col}^{\text{pos}'}$

Annahmen:

- ▶ Ursprungs- und Zielbild sind Approximationen desselben stetigen (beliebig fein aufgelösten) Bildes (dessen Werte aber i.A. nur an den Stützstellen  $\text{pos}$  bekannt ist)
- ▶ Positionen werden als Rasterpunkte im  $\mathbb{R}^2$  betrachtet
- ▶ Farbwert eines Rasterpunktes  $q \in \text{pos}'$  im Zielbild lässt sich aus den Farbwerten der Rasterpunkte in einer Umgebung  $U(T^{-1}(q)) \in 2^{\text{pos}}$  des rücktransformierten Rasterpunktes  $T^{-1}(q) \in \mathbb{R}^2$  berechnen durch Interpolationsfunktion  $g : \text{col}^{\text{pos}} \times \mathbb{R}^2 \times 2^{\text{pos}} \rightarrow \text{col}$

Transformation digitaler Bilder:  $T' : \text{col}^{\text{pos}} \times (\mathbb{R}^2 \rightarrow \mathbb{R}^2) \rightarrow \text{col}^{\text{pos}'}$  mit

$$\forall q \in \text{pos}' : T'(B, T, q) = g(B, T^{-1}(q), U(T^{-1}(q)))$$

# Interpolationsfunktionen

prominente Interpolationsfunktionen  $g : \text{col}^{\text{pos}} \times \mathbb{R}^2 \times 2^{\text{pos}} \rightarrow \text{col}$ :

- ▶ Mittelwerte der Umgebungswerte

$$g(B, p, U) = \text{avg}\{B(r) \mid r \in U\} \quad (\text{med}\{B(r) \mid r \in U\})$$

- ▶ Farbe des (eines) nächsten Nachbarn (Rasterpunkt in der Umgebung  $U \subseteq \text{pos}$  mit geringstem Abstand zu  $p$ )

$$\forall p \in \mathbb{R}^2 : g(B, p, U) = B(r) \text{ wobei } d(r, p) = \min\{d(s, p) \mid s \in U\}$$

- ▶ lineare Interpolation (der 1 bis 4 nächsten Punkte im Original)

für  $p = (z', s') \in \mathbb{R}^2$  mit  $z \leq z' \leq z + 1$ ,  $s \leq s' \leq s + 1$

Umgebung  $U(z', s') = \{(z, s), (z, s + 1), (z + 1, s), (z + 1, s + 1)\}$

$$\begin{aligned} g(B, (z', s'), \{(z, s), (z, s + 1), (z + 1, s), (z + 1, s + 1)\}) \\ = a_1 B(z, s) + a_2 B(z, s + 1) + a_3 B(z + 1, s) + a_4 B(z + 1, s + 1) \end{aligned}$$

(gewichtete Summe der Farbwerte) mit

$$\begin{aligned} a_1 &= (1 + z - z')(1 + s - s') \quad , \quad a_2 = (1 + z - z')(s' - s), \\ a_3 &= (z' - z)(1 + s - s') \quad , \quad a_4 = (z' - z)(s' - s) \end{aligned}$$

- ▶ kubische Interpolation

# Eindimensionale Signale

**analog** : Funktion  $\mathbb{R} \rightarrow \mathbb{R}$

**digital** : Funktion  $\text{pos} \rightarrow \text{col}$  mit  $\text{pos}, \text{col} \subseteq \mathbb{N}$ , oft  
 $\text{pos} = \{0, \dots, m-1\} \subseteq \mathbb{N}$ ,  
 $\text{col} = \{0, \dots, n-1\} \subseteq \mathbb{N}$   
(Array / Liste / 1D-Matrix natürlicher Zahlen)

Beispiele:

- ▶ Temperaturverläufe
- ▶ Börsenkurse
- ▶ Töne (Schallwellen)

## Abtastung eindimensionaler Signale

(Transformation eines Signals mit stetigem Definitionsbereich in ein Signal mit diskretem Definitionsbereich)

gegeben: Funktion  $B : \mathbb{R} \rightarrow \mathbb{R}$  (eindimensionales Signal / Bild),  
Menge  $\text{pos} = \{0, \dots, m-1\}$

gesucht: Funktion  $B' : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\forall p \in \mathbb{R} : B'(p) = \begin{cases} B(p) & \text{falls } p \in \text{pos} \\ 0 & \text{sonst} \end{cases}$$

Ziel: Transformation  $f : (\mathbb{R} \rightarrow \mathbb{R}) \rightarrow (\text{pos} \rightarrow \mathbb{R})$  mit

$$\forall B \in (\mathbb{R}^{\mathbb{R}}) \forall p \in \text{pos} : f(B)(p) = B(p)$$

Realisierung von  $f$  durch Multiplikation von  $B$  mit

**Dirac-Impulsen**  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \text{und} \quad \forall x \in \mathbb{R} : x \neq 0 \rightarrow \delta(x) = 0$$

## Dirac-Impuls

Dirac-Impuls  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \text{und} \quad \forall x \in \mathbb{R} : x \neq 0 \rightarrow \delta(x) = 0$$

Approximation von  $\delta$  durch Rechteck-Impulse:  $\forall x \in \mathbb{R} :$

$$f_1(x) = \begin{cases} 1 & \text{falls } -1/2 \leq x \leq 1/2 \\ 0 & \text{sonst} \end{cases}$$

$$f_2(x) = \begin{cases} 2 & \text{falls } -1/4 \leq x \leq 1/4 \\ 0 & \text{sonst} \end{cases}$$

$\vdots$

$$f_i(x) = \begin{cases} 2^{i-1} & \text{falls } -2^{-i} \leq x \leq 2^{-i} \\ 0 & \text{sonst} \end{cases}$$

$\vdots$

$$\forall i > 0 : \int_{-\infty}^{\infty} f_i(x) dx = \int_{-2^{-i}}^{2^{-i}} 2^{i-1} dx = 2^{i-1} \cdot 2 \cdot 2^{-i} = 1$$

# Multiplikation mit Dirac-Impuls

$$\forall i \in \mathbb{N} : f_i(x) = \begin{cases} 2^{i-1} & \text{falls } -2^{-i} \leq x \leq 2^{-i} \\ 0 & \text{sonst} \end{cases}$$

Wirkung der Multiplikation eines Bildes  $B : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f_i$ :

$$\begin{aligned} \int_{-\infty}^{\infty} B(x) f_i(x) dx &= \int_{-2^{-i}}^{2^{-i}} 2^{i-1} B(x) dx \\ &= \text{avg} \left\{ B(x) \mid x \in \left[ -2^{-i}, 2^{-i} \right] \right\} \end{aligned}$$

$$\text{für } i \rightarrow \infty : \delta : \int_{-\infty}^{\infty} B(x) \delta(x) dx = B(0)$$

# Dirac-Kamm

Dirac-Impuls  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \text{und} \quad \forall x \in \mathbb{R} : x \neq 0 \rightarrow \delta(x) = 0$$

Verschiebung von  $B$  um  $p$  entlang der  $t$ -Achse:

$$\int_{-\infty}^{\infty} B(x + p) \delta(x) dx = \int_{-\infty}^{\infty} B(x) \delta(p - x) dx = B(p)$$

(wegen  $\delta(p - x) \neq 0$  gdw.  $x = p$ )

Dirac-Kamm: Summe der Dirac-Impulse an jedem  $p \in \text{pos}$

$$\forall x \in \mathbb{R} : \delta_{\text{pos}}(x) = \sum_{p \in \text{pos}} \delta(p - x)$$

Transformation  $f : (\mathbb{R} \rightarrow \mathbb{R}) \rightarrow (\mathbb{R} \rightarrow \mathbb{R})$  mit

$$\begin{aligned} \forall B \in (\mathbb{R}^{\mathbb{R}}) \quad \forall p \in \text{pos} : f(B)(p) &= \int_{-\infty}^{\infty} B(x) \delta_{\text{pos}}(x) dx \\ &= \begin{cases} B(p) & \text{falls } p \in \text{pos} \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Ziel erreicht: Abtastung der Funktion  $B : \mathbb{R} \rightarrow \mathbb{R}$  an den Stellen  $\text{pos}$

# Beispiele

- ▶  $B : \mathbb{R} \rightarrow \{0, 1\}$  mit  $\forall x \in \mathbb{R} : B(x) = \lfloor x \rfloor \bmod 2$

Abtastung an allen Positionen

- ▶  $\text{pos} = \{n/2 \mid n \in \{0, \dots, 10\}\}$

$$B' = \{(0, 0), (1/2, 0), (1, 1), (3/2, 1), (2, 0), (5/2, 0), \dots\}$$

- ▶  $\text{pos} = \{3n/2 \mid n \in \{0, \dots, 10\}\}$

$$B' = \{(0, 0), (3/2, 1), (3, 1), (9/2, 0), (6, 0), (15/2, 1), \dots\}$$

Unterabtastung

- ▶  $\text{pos} = \{4n \mid n \in \{0, \dots, 10\}\}$

$$B' = \{(0, 0), (4, 0), (8, 0), \dots\}$$

Unterabtastung

- ▶  $B : \mathbb{R} \rightarrow [-1, 1]$  mit  $\forall t \in \mathbb{R} : B(x) = c \sin(x)$

Aliasing: Rekonstruktion nicht im Original vorhandener Frequenzen  
(Alias-Frequenzen, Schwebungen)

# Probleme bei der Abtastung

Durch Digitalisierung (Transformation stetiger Signale in diskrete) entstehen möglicherweise

Informationsverlust bei Unterabtastung  
verfälschte Strukturen : Alias-Effekte

## Nyquist-Shannon-Abtasttheorem:

Zur verlustfreien Rekonstruktion darf das Abtastraster (Abstand  $\Delta$  zwischen den Abtast-Positionen) nicht größer sein als die halbe Größe des kleinsten abzubildenden Details.

alternative Formulierung:

Abtastrate ( $1/\Delta$ ) muss wenigstens doppelt so groß wie die höchste im Signal vorkommende Frequenz sein

# Zusammenfassung Abtastung 1D

Transformation einer stetigen Funktion  $B : \mathbb{R} \rightarrow \mathbb{R}$  (analoges Signal) in eine Funktion  $B' : \text{pos} \rightarrow \mathbb{R}$  (diskretes Signal)  
mathematische Beschreibung dieser Transformation durch Funktion(al)  $f_{\text{pos}} : (\mathbb{R} \rightarrow \mathbb{R}) \rightarrow (\text{pos} \rightarrow \mathbb{R})$

Dirac-Impuls  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \text{und} \quad \forall x \in \mathbb{R} : x \neq 0 \rightarrow \delta(x) = 0$$

Dirac-Kamm  $\delta_{\text{pos}}$  an Abtaststellen  $\text{pos} \subset \mathbb{R}$ :

$$\forall x \in \mathbb{R} : \delta_{\text{pos}}(x) = \sum_{p \in \text{pos}} \delta(p - x)$$

abgetastetes Signal  $f_{\text{pos}}(B)$  an Abtaststellen  $\text{pos} \subset \mathbb{R}$ :  $\forall x \in \mathbb{R}$ :

$$(f_{\text{pos}}(B))(x) = \int_{-\infty}^{\infty} B(x) \delta_{\text{pos}}(x) dx = \sum_{p \in \text{pos}} \int_{-\infty}^{\infty} B(x) \delta(p - x) dx$$

**Abtasttheorem** (Nyquist-Shannon): Zur verlustfreien Rekonstruktion darf das Abtastraster (Abstand zwischen benachbarten Abtast-Positionen) nicht größer sein als die halbe Größe des kleinsten abzubildenden Details.

# Faltung

Verschiebung eines Dirac-Impulses um  $p$  entlang der  $x$ -Achse:

$$\int_{-\infty}^{\infty} B(x+p)\delta(x)dx = \int_{-\infty}^{\infty} B(x)\delta(p-x)dx = (B * \delta)(p)$$

ist Spezialfall einer Operation mit zwei Funktionen:

Für  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  ist die **Faltung** (Convolution) von  $f$  und  $g$  definiert durch

$$\forall p \in \mathbb{R} : (f * g)(p) = \int_{-\infty}^{\infty} f(p-x)g(x)dx$$

## Faltung – Beispiel

$$f(x) = \chi_{[-1,1]} = \begin{cases} 1 & \text{falls } -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

$$g(x) = 2\chi_{[0,2]} = \begin{cases} 2 & \text{falls } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

an Stellen  $p \in \{-2, -1, 0, 1, 2, 3, 4\}$  ausrechnen:

$$\forall p \in \mathbb{R} : (f * g)(p) = \int_{-\infty}^{\infty} f(x)g(p-x)dx = \max(0, 4 - 2|p-1|)$$

anschaulich:

1. Spiegelung von  $g$  an  $y$ -Achse
2. gespiegeltes  $g$  wird entlang der  $x$ -Achse geschoben
3.  $(f * g)(p)$  ist der Flächeninhalt der vom Produkt beider Funktionen überdeckten Fläche

# Was bisher geschah

- ▶ digitales Bild = Funktion  $B : \text{pos} \rightarrow \text{col}$
- ▶ statistische Merkmale (Bildanalyse)  
und ihre Aussagen über das Bild
- ▶ Punktoperationen  $f : \text{col}_1 \rightarrow \text{col}_2$  (Farbtransformation)  
und ihre Fortsetzung  $f : \text{col}_1^{\text{pos}} \rightarrow \text{col}_2^{\text{pos}}$
- ▶ Geometrische Operationen  $f : \text{pos}_1 \rightarrow \text{pos}_2$   
(Koordinatentransformation)

eindimensionale Signale  $f : \mathbb{R} \rightarrow \mathbb{R}$

- ▶ Dirac-Impuls, Dirac-Kamm, Abtast-Funktion
- ▶ Digitalisierung  $f' : \text{pos} \rightarrow \mathbb{R}$
- ▶ Alias-Effekte
- ▶ Nyquist-Shannon-Abtasttheorem

# Zusammenfassung Abtastung 1D

Transformation einer stetigen Funktion  $B : \mathbb{R} \rightarrow \mathbb{R}$  (analoges Signal) in eine Funktion  $B' : \text{pos} \rightarrow \mathbb{R}$  (diskretes Signal)  
mathematische Beschreibung dieser Transformation durch Funktion(al)  $f_{\text{pos}} : (\mathbb{R} \rightarrow \mathbb{R}) \rightarrow (\text{pos} \rightarrow \mathbb{R})$

Dirac-Impuls  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \text{und} \quad \forall x \in \mathbb{R} : x \neq 0 \rightarrow \delta(x) = 0$$

Dirac-Kamm  $\delta_{\text{pos}}$  an Abtaststellen  $\text{pos} \subset \mathbb{R}$ :

$$\forall x \in \mathbb{R} : \delta_{\text{pos}}(x) = \sum_{p \in \text{pos}} \delta(p - x)$$

abgetastetes Signal  $f_{\text{pos}}(B)$  an Abtaststellen  $\text{pos} \subset \mathbb{R}$ :  $\forall x \in \mathbb{R}$ :

$$(f_{\text{pos}}(B))(x) = \int_{-\infty}^{\infty} B(x) \delta_{\text{pos}}(x) dx = \sum_{p \in \text{pos}} \int_{-\infty}^{\infty} B(x) \delta(p - x) dx$$

**Abtasttheorem** (Nyquist-Shannon): Zur verlustfreien Rekonstruktion darf das Abtastraster (Abstand zwischen benachbarten Abtast-Positionen) nicht größer sein als die halbe Größe des kleinsten abzubildenden Details.

## Wiederholung: Faltung (1D)

Verschiebung eines Dirac-Impulses um  $p$  entlang der  $x$ -Achse:

$$\int_{-\infty}^{\infty} B(x+p)\delta(x)dx = \int_{-\infty}^{\infty} B(x)\delta(p-x)dx = (B * \delta)(p)$$

ist Spezialfall einer Operation mit zwei Funktionen:

Für  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  ist die **Faltung** (Convolution) von  $f$  und  $g$  definiert durch

$$\forall p \in \mathbb{R} : (f * g)(p) = \int_{-\infty}^{\infty} f(p-x)g(x)dx$$

## Faltung – Beispiel

$$f(x) = \chi_{[-1,1]} = \begin{cases} 1 & \text{falls } -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

$$g(x) = 2\chi_{[0,2]} = \begin{cases} 2 & \text{falls } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

$$(f * g)(p) = \int_{-\infty}^{\infty} f(x)g(p-x)dx = \int_{-1}^1 g(p-x)dx = \max(0, 4 - 2|p-1|)$$

an Stellen  $p \in \{-2, -1, 0, 1, 2, 3, 4\}$  berechnen (Tafel):

$$(f * g)(-2) = (f * g)(-1) = 0, (f * g)(0) = 2, (f * g)(1) = 4, (f * g)(2) = 2, \dots$$

anschaulich:

1. Spiegelung von  $g$  an  $y$ -Achse
2. gespiegeltes  $g$  wird entlang der  $x$ -Achse um  $p$  verschoben
3.  $(f * g)(p)$  ist der Flächeninhalt der vom Produkt beider Funktionen überdeckten Fläche

## Faltung – Eigenschaften

- ▶ kommutativ, d.h.  $\forall f, g : f * g = g * f$
- ▶ assoziativ, d.h.  $\forall f, g, h : (f * g) * h = f * (g * h)$
- ▶ Dirac-Impuls  $\delta$  ist neutrales Element bzgl.  $*$ :  
Für jede Funktion  $B : \mathbb{R} \rightarrow \mathbb{R}$  gilt

$$\forall x \in \mathbb{R} : (B * \delta)(p) = \int_{-\infty}^{\infty} B(x)\delta(p-x)dx = B(p)$$

also  $B * \delta = B$

- ▶ für verschobenen Dirac-Impuls  $\delta_p$  mit  
 $\forall x \in \mathbb{R} : \delta_p(x) = \delta(p-x)$  und  
jede Funktion  $B : \mathbb{R} \rightarrow \mathbb{R}$  gilt

$$\forall x \in \mathbb{R} : (B * \delta_p)(x) = B(p-x)$$

- ▶ bilinear, d.h. linear in jeder Komponente  
 $(af + bf') * g = a(f * g) + b(f' * g)$  (analog für  $ag + bg'$ )

# Diskrete Faltung

stetig:

Für  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  ist die Faltung von  $f$  und  $g$  definiert durch

$$\forall p \in \mathbb{R} : (f * g)(p) = \int_{-\infty}^{\infty} f(x)g(p-x)dx$$

diskret:

Für  $f, g : \mathbb{Z} \rightarrow \mathbb{Z}$  ist die Faltung (Convolution) von  $f$  und  $g$  definiert durch

$$\forall p \in \mathbb{Z} : (f * g)(p) = \sum_{x=-\infty}^{\infty} f(x)g(p-x)$$

diskret mit endlichem Definitionsbereich pos:

Für  $f, g : \text{pos} \rightarrow \mathbb{Z}$  ist die **Faltung** (Convolution) von  $f$  und  $g$  definiert durch

$$\forall p \in \text{pos} : (f * g)(p) = \sum_{x \in \text{pos}} f(x)g(p-x)$$

## Diskrete Faltung – Beispiel

$f, g : \mathbb{Z} \rightarrow \mathbb{N}$  mit

$$f(x) = \begin{cases} 3 & \text{falls } x = 1 \\ 1 & \text{falls } x = 2 \\ 0 & \text{sonst} \end{cases} \quad g(x) = \begin{cases} 2 & \text{falls } x = 1 \\ 1 & \text{falls } x = 2 \\ 0 & \text{sonst} \end{cases}$$

als Vektoren mit endlichem Definitionsbereich  $\text{pos} = \{0, \dots, 3\}$ :

$$f = (0, 3, 1, 0), g = (0, 2, 1, 0)$$

$$(f * g)(p) = \begin{cases} 6 & \text{falls } p = 2 \\ 5 & \text{falls } p = 3 \\ 1 & \text{falls } p = 4 \\ 0 & \text{sonst} \end{cases}$$

als Vektor:  $f * g = (0, 6, 5, 1)$  (relevanter Teil)

# Zweidimensionale Signale

**analog** : Funktion  $\mathbb{R}^2 \rightarrow \mathbb{R}$

**digital** : Funktion  $\text{pos} \rightarrow \text{col}$  mit  $\text{col} \subseteq \mathbb{N}$  und  $\text{pos} \subseteq \mathbb{N}^2$ ,  
meist

$\text{pos} = \{0, \dots, m - 1\} \times \text{pos} = \{0, \dots, n - 1\} \subseteq \mathbb{N}^2$

und  $\text{col} = \{0, \dots, k - 1\} \subseteq \mathbb{N}$

(Matrix natürlicher Zahlen, Bilder)

## Abtastung (2D)

Transformation eines Signals mit stetigem Definitionsbereich in ein Signal mit diskretem Definitionsbereich (Raster)

gegeben: Funktion  $B : \mathbb{R}^2 \rightarrow \mathbb{R}$

gesucht: Funktion  $B' : \text{pos} \rightarrow \mathbb{R}$  mit

$$\text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$$

Transformation  $f : (\mathbb{R}^2 \rightarrow \mathbb{R}) \rightarrow (\text{pos} \rightarrow \mathbb{R})$  mit

$\forall B : \mathbb{R}^2 \rightarrow \mathbb{R} \forall p = (p_x, p_y) \in \text{pos} :$

$$f(B)(p_x, p_y) = \begin{cases} B(p_x, p_y) & \text{falls } (p_x, p_y) \in \text{pos} \\ 0 & \text{sonst} \end{cases}$$

Realisierung durch Multiplikation von  $B$  mit

**Dirac-Impulsen**  $\delta : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) dx dy = 1 \quad \text{und}$$

$$\forall (x, y) \in \mathbb{R}^2 : (x, y) \neq (0, 0) \rightarrow \delta(x, y) = 0$$

## Dirac-Kamm (2D)

Dirac-Impuls  $\delta : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) dx dy = 1 \text{ und } \forall (x, y) \in \mathbb{R}^2 : (x, y) \neq (0, 0) \rightarrow \delta(x, y) = 0$$

Verschiebung in der Ebene um  $(p_x, p_y) \in \mathbb{R}^2$ :

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(x + p_x, y + p_y) \delta(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(x, y) \delta(p_x - x, p_y - y) dx dy \\ &= B(p_x, p_y) \end{aligned}$$

2D-Dirac-Kamm: Summe mehrerer verschobener Dirac-Impulse

$$\forall (x, y) \in \mathbb{R}^2 : \delta_{\text{pos}}(x, y) = \sum_{p=(p_x, p_y) \in \text{pos}} \delta(p_x - x, p_y - y)$$

Abtastung der Funktion  $B : \mathbb{R}^2 \rightarrow \mathbb{R}$  an den Stellen  $\text{pos} \subseteq \mathbb{R}^2$ :

$$\forall (x, y) \in \text{pos} : B'(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(x, y) \delta_{\text{pos}}(x, y) dx dy = B(x, y)$$

## Faltung (2D)

Für  $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$  ist die **Faltung** (Convolution) von  $f$  und  $g$  definiert durch

$$\forall (p_x, p_y) \in \mathbb{R}^2 : (f * g)(p_x, p_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) g(p_x - x, p_y - y) dx dy$$

diskreter Fall (Bilder):

Für  $f, g : \{0, \dots, m-1\} \times \{0, \dots, m-1\} \rightarrow \text{col}$  ist die **Faltung** (Convolution) von  $f$  und  $g$  definiert durch

$\forall (p_x, p_y) \in \{0, \dots, m-1\} \times \{0, \dots, m-1\} :$

$$(f * g)(p_x, p_y) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) g(p_x - x, p_y - y)$$

# Transformation stetiger Signale

bisher:

- ▶ Diskretisierte Funktion  $f : \text{pos} \rightarrow \text{col}$  als **Linearkombinationen** (gewichtete Summe) von verschobenen **Dirac-Impulsen**  
 $\delta : \mathbb{R} \rightarrow \mathbb{R}$  (bilden Basis)
- ▶ Transformationen durch Operationen auf
  - col** : (punktweise) Farbtransformationen
  - pos** : geometrisch (Verschiebung, Skalierung, ...)

alternativ:

- ▶ Beschreibung von Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  als **Linearkombinationen** (gewichtete Summe) festgelegter (stetiger) **Basisfunktionen**  $g_i : \mathbb{R} \rightarrow \mathbb{R}$  (geeignet skaliert und verschoben)

$$\forall x \in \mathbb{R} : f(x) = \sum_{i \in I} a_i g_i(x)$$

- ▶ Darstellung von Bildern als Linearkombinationen (stetiger) „Basisbilder“ (evtl. skaliert, verschoben und rotiert)
- ▶ Transformationen durch **Operationen auf den Koeffizienten**  $a_i$  der Standardfunktionen

# Periodische Funktionen (1D, Wiederholung)

Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt **periodisch** gdw.

$$\exists p \in \mathbb{R} \forall x \in \mathbb{R} \forall i \in \mathbb{Z} : f(x) = f(x + ip)$$

$p \in \mathbb{R}$  heißt Periode von  $f$

Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt

**gerade** gdw.  $\forall x \in \mathbb{R} : f(x) = f(-x)$   
(Symmetrie an  $y$ -Achse)

z.B.  $x \mapsto x^2, x \mapsto \cos(x)$

**ungerade** gdw.  $\forall x \in \mathbb{R} : f(x) = -f(-x)$

z.B.  $x \mapsto x^3, x \mapsto \sin(x)$

## Fourierreihe (1D) – Wiederholung

Darstellung von (periodischen) Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$

als Linearkombination (gewichtete Summe)

periodischer zueinander orthogonaler „Standard“-Funktionen

mit Perioden  $p_k = \frac{2\pi k}{p}$  für Periode  $p$  von  $f$  und alle  $k \in \mathbb{Z}$ .

Für Fourierreihen trigonometrische Funktionen mit Periode  $2\pi$

cos gerader Anteil

sin ungerader Anteil

$$\forall x \in \mathbb{R} : f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

mit Koeffizienten  $a_k, b_k \in \mathbb{R}$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \quad \left( \text{insbesondere } a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx \right)$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx$$

## Beispiel: cos

Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\forall x \in \mathbb{R} : f(x) = 3 \cos(2x)$$

$$\forall x \in \mathbb{R} : 3 \cos(2x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

mit Koeffizienten:

$$a_2 = 3,$$

$$\forall k \in \mathbb{N} \setminus \{2\} : a_k = 0,$$

$$\forall k \in \mathbb{N} : b_k = 0$$

## Beispiel: Stufenfunktion (Rechteck)

Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\forall x \in \mathbb{R} : f(x) = \begin{cases} 1 & \text{falls } \exists k \in \mathbb{Z} : x \in [2k\pi, (2k+1)\pi) \\ -1 & \text{falls } \exists k \in \mathbb{Z} : x \in [(2k-1)\pi, 2k\pi) \end{cases}$$

(stückweise konstante Funktion mit Periode  $2\pi$ )

$$\forall x \in \mathbb{R} : f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

mit Koeffizienten  $\forall k \in \mathbb{N} : a_k = 0,$

$$\forall k \in \mathbb{N} : b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx = \begin{cases} \frac{4}{k\pi} & \text{falls } k \equiv 1 \pmod{2} \\ 0 & \text{sonst} \end{cases}$$

also

$$\forall x \in \mathbb{R} : f(x) = \frac{4}{\pi} \sum_{\substack{k \in \mathbb{N} \\ k \equiv 1 \pmod{2}}} \frac{\sin(kx)}{k}$$

# Was bisher geschah

- ▶ digitale Bilder
  - ▶ Funktion  $B : \text{pos} \rightarrow \text{col}$
  - ▶ Matrix  $B \in \text{col}^{\text{pos}}$
- ▶ statistische Merkmale (Bildanalyse)
- ▶ Punktoperationen  $f : \text{col}_1 \rightarrow \text{col}_2$  (Farbtransformation)
- ▶ Geometrische Operationen  $f : \text{pos}_1 \rightarrow \text{pos}_2$  (Koordinatentransformation)
- ▶ Interpolation der Intensitäten
- ▶ 1D- und 2D-Signale (analog, diskret)
- ▶ Digitalisierung (Abtastung) von 1D- und 2D-Signalen
- ▶ Alias-Effekte
- ▶ Faltung von 1D- und 2D-Funktionen (analog, diskret)
- ▶ Wiederholung Fourierreihen

## Fourierreihe (1D) – Wiederholung

Darstellung von (periodischen) Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$

als Linearkombination (gewichtete Summe)

periodischer zueinander orthogonaler „Standard“-Funktionen

mit Perioden  $p_k = \frac{2\pi k}{p}$  für Periode  $p$  von  $f$  und alle  $k \in \mathbb{Z}$ .

Für Fourierreihen trigonometrische Funktionen mit Periode  $2\pi$

cos gerader Anteil

sin ungerader Anteil

$$\forall x \in \mathbb{R} : f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

mit Koeffizienten  $a_k, b_k \in \mathbb{R}$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \quad \left( \text{insbesondere } a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx \right)$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx$$

# Komplexe Darstellung

Transformation von Linearkombinationen reellwertiger trigonometrischer Funktionen in  $\mathbb{C}$ -wertige Funktionen mit Hilfe der Eulerschen Formel:

$$\begin{aligned}\forall x \in \mathbb{R} : e^{ix} &= \cos(x) + i \sin(x) \\ e^{ikx} &= \cos(kx) + i \sin(kx) \\ e^{-ikx} &= \cos(kx) - i \sin(kx)\end{aligned}$$

Darstellung von  $f : \mathbb{R} \rightarrow \mathbb{R}$  als **Fourier-Reihe**:

$$\forall x \in \mathbb{R} : f(x) = \sum_{-\infty}^{\infty} c_k e^{ikx}$$

mit den **Fourier-Koeffizienten**

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx = \begin{cases} a_0/2 & \text{falls } k = 0 \\ (a_k - ib_k)/2 & \text{falls } k > 0 \\ (a_{-k} - ib_{-k})/2 & \text{falls } k < 0 \end{cases}$$

# Fourier-Reihen periodischer Funktionen (1D)

$f : \mathbb{R} \rightarrow \mathbb{R}$  stückweise stetig und periodisch, d.h.

$$\exists p \in \mathbb{R} \forall x \in \mathbb{R} \forall n \in \mathbb{Z} : f(x) = f(x + np)$$

Darstellung von  $f : \mathbb{R} \rightarrow \mathbb{R}$  als Fourier-Reihe:  $\forall x \in \mathbb{R} :$

$$f(x) = \sum_{k \in \mathbb{Z}} \left( a_k \cos \left( \frac{2\pi kx}{p} \right) + b_k i \sin \left( \frac{2\pi kx}{p} \right) \right) = \sum_{k \in \mathbb{Z}} \left( c_k e^{i \frac{2\pi kx}{p}} \right)$$

mit den **Fourier-Koeffizienten**  $\forall k \in \mathbb{Z} : c_k \in \mathbb{C}$  mit

$$c_k = \frac{1}{p} \int_0^p f(x) e^{-i \frac{2\pi kx}{p}} dx = \frac{1}{p} \int_{-p/2}^{p/2} f(x) e^{-i \frac{2\pi kx}{p}} dx$$

Grundfrequenz von  $f$ :  $f_p = 1/p$

$c_k \neq 0$  gdw. Frequenz  $k/p$  kommt in  $f$  mit Amplitude  $|c_k|$  vor  
(diskretes) **Fourier-Spektrum** von  $f$ :

$c : \mathbb{Z} \rightarrow \mathbb{C}$  mit  $\forall k \in \mathbb{Z} : c(k) = c_k$  (Folge  $(c_k)_{k \in \mathbb{Z}}$ )

anschaulich:

- ▶  $f$  hat einen Frequenzanteil  $k/p$  gdw.  $c_k \neq 0$
- ▶  $|c_k|$  ist die Amplitude des Frequenzanteils  $k/p$

# Fourier-Transformation nicht-periodischer Funktionen

Idee: wachsende Periodenlänge  $\rightarrow$  geringere Frequenz

Grenzübergang von diskreten Frequenzen  $2\pi k/p \in \mathbb{R}$

zu kontinuierlichen Frequenzen  $u \in \mathbb{R}$

Falls zu einer (nicht notwendig periodischen) Funktion

$f : \mathbb{R} \rightarrow \mathbb{R}$  für jedes  $u \in \mathbb{R}$  das Integral

$$\int_{\mathbb{R}} f(x) e^{-iux} dx$$

existiert,

**Fourier-Transformation**  $F : \mathbb{R}^{\mathbb{R}} \rightarrow \mathbb{C}^{\mathbb{R}}$  mit

$$\forall u \in \mathbb{R} : (F(f))(u) = \int_{\mathbb{R}} f(x) e^{-iux} dx$$

$F(f)$  heißt die Fourier-Transformierte von  $f$

(kontinuierliches Fourier-Spektrum)

**inverse** Fourier-Transformation  $F^{-1} : \mathbb{C}^{\mathbb{R}} \rightarrow \mathbb{R}^{\mathbb{R}}$

$$\forall x \in \mathbb{R} : (F^{-1}(d))(x) = \int_{\mathbb{R}} d(u) e^{iux} du$$

# Diskrete Fourier-Transformation (1D)

gegeben:  $f : \text{pos} \rightarrow \mathbb{R}$  mit  $\text{pos} = \{0, \dots, m-1\}$

Idee:

periodische Fortsetzung von  $f$  auf  $\mathbb{Z}$

(mit Periode  $m$ , Frequenz  $1/m$ )

diskrete Fourier-Transformation von  $f : \text{pos} \rightarrow \mathbb{R}$  (1-DFT):

$F : \mathbb{R}^{\text{pos}} \rightarrow \mathbb{C}^{\text{pos}}$  mit

$$\forall u \in \{0, \dots, m-1\} : F(f)(u) = \sum_{k=0}^{m-1} f(k) e^{-2\pi i k u / m}$$

inverse Fourier-Transformation von  $g : \text{pos} \rightarrow \mathbb{C}$ :

$F^{-1} : \mathbb{C}^{\text{pos}} \rightarrow \mathbb{R}^{\text{pos}}$  mit

$$\forall x \in \{0, \dots, m-1\} : F^{-1}(g)(x) = \sum_{k=0}^{m-1} g(k) e^{2\pi i k x / m}$$

# Eigenschaften der 1D-DFT

1D-DFT  $F : \mathbb{R}^{\text{pos}} \rightarrow \mathbb{C}^{\text{pos}}$

$$F(f)(x) = \sum_{k=0}^{m-1} f(k) e^{-2\pi i k x / m}$$

für alle  $f : \text{pos} \rightarrow \mathbb{R}$  gilt:

- ▶  $F(f)(x) = F(f)(-x)$  (symmetrisch)
- ▶  $|F(f)(x)| = |F(f)(-x)|$
- ▶  $F(f)(x + m) = F(f)(x)$  (unendlich periodisch)

## Diskrete Fourier-Transformation (2D)

gegeben:  $f : \text{pos} \rightarrow \mathbb{R}$  mit  $\text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$

periodische Fortsetzung von  $f$  auf  $\mathbb{Z}^2$

diskrete Fourier-Transformation von  $f : \text{pos} \rightarrow \mathbb{R}$  (2-DFT):

$F : \mathbb{R}^{\text{pos}} \rightarrow \mathbb{C}^{\text{pos}}$  mit

$$\forall (u_x, u_y) \in \text{pos} : F(f)(u_x, u_y) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) e^{-2\pi i(xu_x/m + yu_y/n)}$$

inverse Fourier-Transformation von  $g : \text{pos} \rightarrow \mathbb{C}$ :

$F^{-1} : \mathbb{C}^{\text{pos}} \rightarrow \mathbb{R}^{\text{pos}}$  mit

$$\forall (u_x, u_y) \in \text{pos} : F^{-1}(g)(u_x, u_y) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} g(x, y) e^{2\pi i(xu_x/m + yu_y/n)}$$

## Fourier-Spektren (2D)

Fourier-Transformation  $F : \mathbb{R}^{\text{pos}} \rightarrow \mathbb{C}^{\text{pos}}$  mit

$$\forall (u_x, u_y) \in \text{pos} : F(f)(u_x, u_y) = \sum_{x=0}^{m-1} \sum_{y=0}^{m-1} f(x, y) e^{-2\pi i(xu_x/m + yu_y/n)}$$

**Fourier-Spektrum:** Matrix  $S(f) \in \mathbb{R}^{\text{pos}}$  mit

$$\forall (x, y) \in \text{pos} : S(f)(x, y) = |F(f)(x, y)|$$

**Leistungs-Spektrum:** Matrix  $L(f) \in \mathbb{R}^{\text{pos}}$  mit

$$\forall (x, y) \in \text{pos} : P(f)(x, y) = |F(f)(x, y)|^2$$

**Phasen-Spektrum:** Matrix  $\phi(f) \in \mathbb{R}^{\text{pos}}$  mit

$$\forall (x, y) \in \text{pos} : \phi(f)(x, y) = \tan^{-1} \left( \frac{\Im(F(f)(x, y))}{\Re(F(f)(x, y))} \right)$$

# Eigenschaften der 2D-DFT

Fourier-Transformation  $F : \mathbb{R}^{\text{pos}} \rightarrow \mathbb{C}^{\text{pos}}$  mit

$$\forall (u_x, u_y) \in \text{pos} : F(f)(u_x, u_y) = \sum_{x=0}^{m-1} \sum_{y=0}^{m-1} f(x, y) e^{-2\pi i(xu_x/m + yu_y/n)}$$

Für alle Funktionen  $f : \text{pos} \rightarrow \mathbb{R}$  gilt:

- ▶  $F(f)(0, 0) = \text{max}(f)$  (mitunter auch  $\text{avg}(f)$ , je nach Faktor)
- ▶  $F(f)(x, y) = F(f)(-x, -y)$  (symmetrisch)
- ▶  $|F(f)(x, y)| = |F(f)(-x, -y)|$
- ▶  $\forall f, g : \text{pos} \rightarrow \mathbb{R} \forall a, b \in \mathbb{R}$ :  
 $F(af + bg)(x, y) = aF(f)(x, y) + bF(g)(x, y)$  (linear)
- ▶  $F(f)(x + m, y + n) = F(f)(x + m, y) = F(f)(x, y + n) = F(f)(x, y)$  (unendlich periodisch in beide Richtungen)

# DFT von Bildern

Bild  $B \in \text{col}^{\text{pos}}$

Fourier-Transformierte  $F(B) \in \mathbb{C}^{\text{pos}}$

$$\forall (x, y) \in \text{pos} : F(B)(x, y) = \sum_{x=0}^{m-1} \sum_{y=0}^{m-1} B(x, y) e^{-2\pi i(xu_x/m + yu_y/n)}$$

mit

Ortsfrequenz  $F(B)(x, y) \in \mathbb{C}$

Spektrum  $|F(B)(x, y)| \in \mathbb{R}$

Phase  $\phi(x, y) = \tan^{-1} \left( \frac{\Im(F(B)(x, y))}{\Re(F(B)(x, y))} \right) \in \mathbb{R}$

# Anwendung in der Bildverarbeitung

**Erkennung** typischer Bildeigenschaften,  
z.B. Größe, dominante Richtungen der Strukturen

**Transformation** eines (Gesamt-)Bildes  $B$  in drei Schritten:

1. Fourier-Transformation  $C = F(B) \in \mathbb{C}^{\text{pos}}$
2. Änderung der Fourier-Transformierten zu  $C'$   
(z.B. Löschen von Teilbildern)
3. inverse Fourier-Transformation  $B' = F^{-1}(C')$

**Kanten-Betonung** durch Entfernen (Wert  $\mapsto 0$ ) niedriger  
Frequenzen (um Ursprung)

**Glättung** durch Entfernen (Wert  $\mapsto 0$ ) hoher Frequenzen  
(entfernt vom Ursprung)

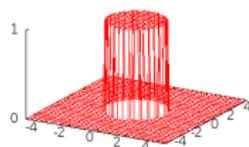
# Transformationen im Frequenzraum

punktweise Transformation  $f : \text{pos} \rightarrow \mathbb{C}$  der Werte der Fourier-Transformierten  $C \in \mathbb{C}^{\text{pos}}$

durch **Multiplikation** mit einer Funktion  $f : \text{pos} \rightarrow \mathbb{R}$ .

Beispiel ( mit Abstand  $d$ , z.B.  $d(u, v) = \sqrt{u^2 + v^2}$  ):

**Tiefpass** mit Schwellwert (Cut-Off-Frequenz)  $\theta \in \mathbb{R}$

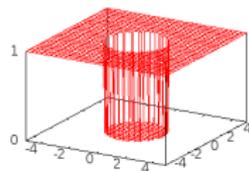


$$f_{T(\theta)}(u, v) = \begin{cases} 1 & , \text{ falls } d(u, v) \leq \theta \\ 0 & \text{sonst} \end{cases}$$

Glättung, aber Ringing-Artefakte möglich

# Mehr Beispiele

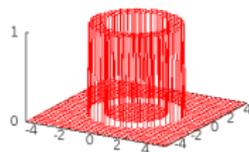
**Hochpass** mit Schwellwert (Cut-Off-Frequenz)  $\theta \in \mathbb{R}$



$$f_{H(\theta)}(u, v) = \begin{cases} 0 & , \text{ falls } d(u, v) \leq \theta \\ 1 & \text{sonst} \end{cases}$$

Hervorhebung von Kanten

**Bandpass** mit Schwellwerten  $\theta_1, \theta_2 \in \mathbb{R}$



$$f_{B(\theta_1, \theta_2)}(u, v) = \begin{cases} 1 & , \text{ falls } \theta_1 \leq d(u, v) \leq \theta_2 \\ 0 & \text{sonst} \end{cases}$$

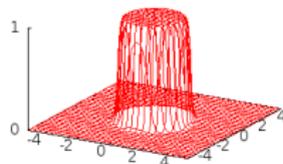
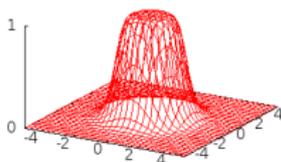
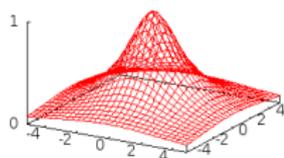
**Bandstop** (analog) mit Schwellwerten  $\theta_1, \theta_2 \in \mathbb{R}$

# Vermeidung von Ringing-Artefakten

Idee:

Wert (Amplitude) bei steigenden / fallenden Frequenzen stetig  
(statt stufenförmig) ändern

Beispiel: Butterworth-Filter



$$f_{T(\theta)}(u, v) = \frac{1}{1 + \left(\frac{d(u,v)}{\theta}\right)^{2n}} \quad (\text{Tiefpass})$$

analog:  $f_{H(\theta)}(u, v) = \frac{1}{1 + \left(\frac{\theta}{d(u,v)}\right)^{2n}} \quad (\text{Hochpass})$

# Ziel: Bildtransformationen im Ortsraum

**Transformation** des Bildes  $B \in \text{col}^{\text{pos}}$  **in drei Schritten**

1. Fourier-Transformation  $F(B) = C \in \mathbb{C}^{\text{pos}}$
2. Änderung der Fourier-Transformierten  $C \in \mathbb{C}^{\text{pos}}$  durch Multiplikation mit einer Funktion  $h'$ :

$$\forall (u, v) \in \text{pos} : C'(u, v) = F(B)(u, v) \cdot h'(u, v)$$

3. inverse Fourier-Transformation (mit  $h' = F(h)$ )

$$B'(x, y) = F^{-1}(F(B) \cdot F(h))(x, y)$$

ersetzen durch

**eine Operation im Ortsraum**  $h : \text{col}^{\text{pos}} \rightarrow \text{col}^{\text{pos}}$  mit

$$\forall B \in \text{col}^{\text{pos}} \forall (x, y) \in \text{pos} : h(B)(x, y) = F^{-1}(F(B) \cdot F(h))(x, y)$$

# Faltungssatz

Wiederholung: Faltung  $f * g$  der Funktionen

$f, g : \{0, \dots, m-1\} \times \{0, \dots, m-1\} \rightarrow \mathbb{R}$  ist definiert durch

$\forall (p_x, p_y) \in \{0, \dots, m-1\} \times \{0, \dots, m-1\} :$

$$(f * g)(p_x, p_y) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f(p_x - x, p_y - y)g(x, y)$$

(assoziativ, kommutativ, linear)

Es gilt der **Faltungssatz**:

$\forall f, g : \text{pos} \rightarrow \text{col} :$

$$\forall (p_x, p_y) \in \text{pos} : F(f * g)(p_x, p_y) = F(f)(p_x, p_y) \cdot F(g)(p_x, p_y)$$

**Faltung im Ortsraum** entspricht **Multiplikation im Frequenzraum**

also  $\forall B : \text{pos} \rightarrow \text{col} \forall h : \text{pos} \rightarrow \mathbb{R} \forall (p_x, p_y) \in \text{pos} :$

$$F(B * h)(p_x, p_y) = F(B)(p_x, p_y) \cdot F(h)(p_x, p_y) = (F(B) \cdot F(h))(p_x, p_y)$$

und damit  $\forall B : \text{pos} \rightarrow \text{col} \forall h : \text{pos} \rightarrow \mathbb{R}$

$$(B * h) = F^{-1}(F(B * h)) = F^{-1}(F(B) \cdot F(h))$$

also Ziel erreicht: Faltung  $B * h$  ersetzt die Operation (drei Schritte)

$$F^{-1}(F(B) \cdot F(h)) : \text{col}^{\text{pos}} \rightarrow \text{col}^{\text{pos}}$$

# Lineare Filter

$$(B * h)(p_x, p_y) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} B(p_x - x, p_y - y) h(x, y)$$

mit Koeffizienten  $h(i, j)$

häufiger Spezialfall:

- ▶ Koeffizienten  $h(x, y)$  verschwinden mit größerem Abstand von  $(0, 0)$ , und außerhalb eines Bereiches  $(x, y) \in \{-k, \dots, k\}^2$  gilt  $h(x, y) = 0$
- ▶ Angabe von  $h$  auf dem (quadratischen) Ausschnitt  $(i, j) \in \{-k, \dots, k\}^2$  als  $(2k + 1) \times (2k + 1)$ -Matrix genügt Filterkern, Filtermaske, Filtermatrix, Faltungskern (zentriert um Ursprung, oft symmetrisch) dargestellt als Funktion  $h : \text{pos}' \rightarrow \mathbb{R}$  mit  $\text{pos}' = \{-k, \dots, k\}^2 \rightarrow \mathbb{R}$

# Lineare Filter

Faltung von Bild  $B$  : pos  $\rightarrow$  col mit Filtermatrix

$$h = \begin{pmatrix} h(-k, -k) & \dots & h(-k, 0) & \dots & h(-k, k) \\ \vdots & & \vdots & & \vdots \\ h(0, -k) & \dots & h(0, 0) & \dots & h(0, k) \\ \vdots & & \vdots & & \vdots \\ h(k, -k) & \dots & h(k, 0) & \dots & h(k, k) \end{pmatrix}$$

transformiert jeden Wert  $B(x, y)$  zu

$$\begin{aligned} (B * h)(x, y) &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} B(x - i, y - j)h(i, j) \\ &= \sum_{i=-k}^k \sum_{j=-k}^k B(x - i, y - j)h(-i, -j) \\ &\stackrel{\substack{(i \rightarrow -i) \\ (j \rightarrow -j)}}{=} \sum_{i=-k}^k \sum_{j=-k}^k B(x + i, y + j)h(-i, -j) \end{aligned}$$

(gespiegelte Filtermatrix)

# Lineare Filter – Beispiele

Glättungsfiler ( $3 \times 3$ ) mit normalisierter Filtermatrix, z.B.

$$h_{\text{Box}} = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad h_{\text{Circ}} = \frac{1}{5} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$h_{\text{Bin}} = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad h_{\text{Gauss}} = \frac{1}{40} \begin{pmatrix} 3 & 5 & 3 \\ 5 & 8 & 5 \\ 3 & 5 & 3 \end{pmatrix}$$

Differenzenfilter ( $5 \times 5$ ) mit normalisierter Filtermatrix, z.B.

$$h_{\text{Laplace}} = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

# Lineare Filter

Faltung von Bild  $B$  : pos  $\rightarrow$  col mit  $(3 \times 3)$ -Filtermatrix

$$h = \begin{pmatrix} h(-1, -1) & h(-1, 0) & h(-1, 1) \\ h(0, -1) & h(0, 0) & h(0, 1) \\ h(1, -1) & h(1, 0) & h(1, 1) \end{pmatrix}$$

transformiert jeden Wert  $B(x, y)$  zu

$$B'(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 B(x+i, y+j)h(-i, -j)$$

gewichtete Summe aller Bildpunkte im Bereich

$(x-1, x, x+1) \times (y-1, y, y+1)$  (lineare Operation)

meist **normalisierte Filtermatrizen**  $h$  mit  $\sum_{i=-k}^k \sum_{j=-k}^k h(i, j) = 1$

Normierung (für Beispiel  $k = 1$ ):

$$\frac{1}{\sum_{i=-k}^k \sum_{j=-k}^k h(i, j)} \begin{pmatrix} h(-1, -1) & h(-1, 0) & h(-1, 1) \\ h(0, -1) & h(0, 0) & h(0, 1) \\ h(1, -1) & h(1, 0) & h(1, 1) \end{pmatrix}$$

# Faltung am Bildrand

$$B'(x, y) = \sum_{i=-k}^k \sum_{j=-k}^k B(x+i, y+j)h(-i, -j)$$

Problem:

- ▶ an Randpunkten  $(x, y)$  hängt  $B'(x, y)$  von Werten  $B(x', y')$  an Positionen  $(x', y') \notin \text{pos}$  ab.
- ▶  $B(x', y')$  ist dort nicht definiert.

mögliche Lösungen:

- ▶ Randbereiche nicht filtern (Werte  $B(x, y)$  bleiben)
- ▶ spezielle Rand-Filtermasken
- ▶ konstante Fortsetzung auf Positionen außerhalb pos
- ▶ schwarzer Rand (Zero-Padding):  $\forall (x, y) \notin \text{pos} : B(x, y) := 0$
- ▶ gespiegelter Rand:  $\forall (x, y) \notin \text{pos} : B(x, y) := B(-x, y)$   
(bzw. einem der Werte  $B(-x, y)$ ,  $B(-x, -y)$ ,  $B(x, -y)$ )
- ▶ periodisch fortgesetzte Werte
- ▶ Interpolation

# Filter zur Kantenverstärkung

Ziel: Unterschiede zwischen benachbarten Pixeln verstärken

(positive und negative Koeffizienten, Summe aller Koeffizienten = 0)

Beispiele:

Roberts-Operator:  $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  und  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  (Summe)

Prewitt-Operator:  $h: \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$   $v: \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$

Sobel-Operator:  $h: \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$   $v: \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$

Differenz-Filter:  $\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$

# Nichtlineare Filter

Funktion  $f : \text{col}^{\text{pos}} \rightarrow \text{col}^{\text{pos}}$  mit

$$\forall (x, y) \in \text{pos} : f(B)(x, y) = f(B(x - k, y - k), \dots, B(x + k, y + k))$$

nicht durch Faltung mit einer Matrix (lineare Funktion) gegeben,  
sondern nichtlinear

Beispiele: Rangordnungsfilter

$$\text{Minimum: } f(B)(x, y) = \min\{B(x - k, y - k), \dots, B(x + k, y + k)\}$$

$$\text{Maximum: } f(B)(x, y) = \max\{B(x - k, y - k), \dots, B(x + k, y + k)\}$$

$$\text{Median: } f(B)(x, y) = \text{med}\{B(x - k, y - k), \dots, B(x + k, y + k)\}$$

gewichteter Medianfilter, abhängig von Matrix  $h : \{-k, \dots, k\}^2 \rightarrow \mathbb{N}$ ,  
Gewicht  $(i, j)$  bestimmt Anzahl der Vorkommen des Wertes  
 $B(x + i, y + j)$  in der Medianliste

# Was bisher geschah

- ▶ digitale Bilder:
  - ▶ Funktion  $B : \text{pos} \rightarrow \text{col}$
  - ▶ Matrix  $B \in \text{col}^{\text{pos}}$
- ▶ statistische Merkmale (Bildanalyse)
- ▶ Punktoperationen  $f : \text{col}_1 \rightarrow \text{col}_2$  (Farbtransformation)
- ▶ Geometrische Operationen  $f : \text{pos}_1 \rightarrow \text{pos}_2$  (Koordinatentransformation)
- ▶ Interpolation der Intensitäten
- ▶ Digitalisierung (Abtastung) von 1D- und 2D-Signalen
- ▶ Aliasing-Effekte
- ▶ Faltung von Signalen (Funktionen)
- ▶ Fourier-Transformierte / inverse Fourier-Transformierte und darin enthaltene Informationen über das Bild
- ▶ Operationen im Frequenzbereich
- ▶ Filter zur Glättung, Kanten-Hervorhebung

# Wiederholung Filter

Glättungsfiler: Tiefpassfilter (mit Normierung)

Mittelwert

$$\frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Gauß

$$\frac{1}{10} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

Median (nichtlinear)

# Gradienten-Filter

(Approximation der 1. oder 2. Ableitung, Hochpass)  
zur Kanten-Hervorhebung, z.B.

horizontale / vertikale Differenzen, z.B.

$$\begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}$$

Roberts: diagonale Differenzen, z.B.

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Prewitt: Differenzen und Mittelung senkrecht dazu

$$\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

Sobel: Differenzen und Gauß senkrecht dazu

$$\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix}$$

oft Linearkombinationen (punktweise Addition der Filterkerne)

# Template-Matching

- ▶ Filterung mit mehreren rotierten Filterkernen (für Kanten verschiedener Richtungen) nacheinander
- ▶ Ergebnisbild: punktweise Maximum der Beträge

geeignete Filterkerne (liefern oft bessere Ergebnisse als Prewitt, Sobel)

## Kompass-Gradient

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{pmatrix} \dots$$

## Kirsch-Operator

$$\begin{pmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{pmatrix} \quad \begin{pmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{pmatrix} \quad \begin{pmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{pmatrix} \dots$$

# Laplace-Operator

Differenzenfilter, Approximation der 2. Richtungsableitungen

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Linearkombinationen, z.B.

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

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

# Transformation von Binärbildern

Bild = Matrix  $B \in \text{col}^{\text{pos}}$

hier Binärbilder mit Farben  $\text{col} = \{0, 1\}$   
(0 für schwarz, 1 für weiß)

Annahmen:

- ▶ Vordergrund und Hintergrund getrennt  
(z.B. durch Binarisierung mit Schwellwert)
- ▶ Vordergrund weiß (1), Hintergrund schwarz (0)

Ziel: Merkmale (z.B. Form) der (weißen) Regionen

- ▶ hervorheben,
- ▶ um Störungen bereinigen,  
(einzelne Punkte, Löcher, zerklüftete Ränder)
- ▶ erkennen
- ▶ Mustern zuordnen (klassifizieren)

# Morphologische Operationen

Idee:

Transformation von Binärbildern  $B \in \{0, 1\}^{\text{pos}}$  durch Verknüpfung von  $B$  mit bestimmten (geeignet transformierten) **Strukturelementen** (Masken) durch Mengenoperationen

Strukturelement  $M : \text{pos}' \rightarrow \{0, 1\}$  definiert

- ▶ Bezugspunkt  $(0, 0) \in \text{pos}'$  (hot spot)
- ▶ relevante Positionen  $p \in \text{pos}$  mit  $M(p) = 1$  (an der Mengenoperation mit  $B$  beteiligt)
- ▶ irrelevante Positionen  $p \in \text{pos}$  mit  $M(p) = 0$

Beispiele:

0	1	0
0	1	0
0	1	0

0	1	0
1	1	1
0	1	0

1	0	0
1	1	0
0	0	1

1	1	0
1	0	0
1	1	1

1	1
---	---

# Strukturelemente

Form und Bezugspunkt abhängig von gewünschter Wirkung

prominente Formen

(häufig symmetrisch mit Bezugspunkt in der Mitte):

- ▶ Kreuz (z.B. 4-Nachbarschaft)
- ▶ Quadrat (z.B. 8-Nachbarschaft)
- ▶ Rechtecke mit verschiedenen Positionen des Bezugspunktes
- ▶ Kreise, Ellipsen (diskrete Approximationen)

Wiederholung (geometrische Transformationen):

Verschiebung  $V_p(B) : \{0, 1\}^{\text{pos}} \rightarrow \{0, 1\}^{\text{pos}}$  um

$p = (p_x, p_y) \in \text{pos}$

$$\forall (x, y) \in \text{pos} : V_p(B)(x, y) = (x + p_x, y + p_y)$$

# Erosion

gegeben:  $B : \text{pos} \rightarrow \{0, 1\}$

Strukturelement  $M : \text{pos}' \rightarrow \{0, 1\}$

Erosion: Bildpunkt  $p$  weiß (Vordergrund) gdw.

**alle** relevanten Positionen des mit dem Bezugspunkt in  $p$  verschobenen Strukturelementes weiß

$$B \ominus M = \{p \in \text{pos} \mid V_p(M) \subseteq B\} = \{p \in \text{pos} \mid V_p(M) \cap \bar{B} = \emptyset\}$$

$$\begin{aligned} \forall p \in \text{pos} : (B \ominus M)(p) &= \min\{B(q) \mid q \in V_p(M)\} \\ &= \min\{B(q) \mid (V_p(M))(q) = 1\} \\ &= \begin{cases} 1 & , \text{ falls } \forall q \in V_p(M) : B(q) = 1 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

# Erosion – Beispiel

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline \end{array}$$

$$M = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array}$$

$$B \ominus M = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & \mathbf{0} & 0 \\ \hline 0 & 0 & 1 & \mathbf{0} & 0 \\ \hline 0 & 0 & 1 & \mathbf{0} & 0 \\ \hline 0 & \mathbf{0} & 0 & 0 & 0 \\ \hline 0 & \mathbf{0} & 0 & 0 & 0 \\ \hline \end{array}$$

# Erosion – Eigenschaften

- ▶ nicht kommutativ:  $\exists B, M : B \ominus M \neq M \ominus B$
- ▶ nicht assoziativ:  $\exists B, M, M' : (B \ominus M) \ominus M' \neq B \ominus (M \ominus M')$

Wirkung: Abtragen der Vordergrundregionen an

- ▶ Randpunkten
- ▶ kleinen isolierten Punktgruppen
- ▶ konvexen Randbereichen
- ▶ Erosion mit quadratischem Strukturelement:  
Minimum-Filter

# Dilatation

gegeben:  $B : \text{pos} \rightarrow \{0, 1\}$

Strukturelement  $M : \text{pos}' \rightarrow \{0, 1\}$

Dilatation: Bildpunkt  $p$  weiß gdw.

**wenigstens eine** relevante Position des mit dem Bezugspunkt in  $p$  verschobenen Strukturelementes weiß

$$B \oplus M = \{p \in \text{pos} \mid V_p(M) \cap B \neq \emptyset\}$$

$$\begin{aligned} \forall p \in \text{pos} : (B \oplus M)(p) &= \max\{B(q) \mid q \in V_p(M)\} \\ &= \max\{B(q) \mid (V_p(M))(q) = 1\} \\ &= \begin{cases} 1 & , \text{ falls } \exists q \in V_p(M) : B(q) = 1 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

# Dilatation – Beispiel

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline \end{array}$$

$$M = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array}$$

$$B \oplus M = \begin{array}{|c|c|c|c|c|} \hline \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \hline 0 & \mathbf{1} & 1 & 1 & 0 \\ \hline 0 & \mathbf{1} & 1 & 1 & 0 \\ \hline \mathbf{1} & 1 & 0 & 0 & 0 \\ \hline \mathbf{1} & 1 & 0 & 0 & 0 \\ \hline \end{array}$$

# Dilatation – Eigenschaften

- ▶ kommutativ:  $\forall B, M : B \oplus M = M \oplus B$
- ▶ assoziativ:  $\forall B, M, M' : (B \oplus M) \oplus M' = B \oplus (M \oplus M')$

Wirkung: Auffüllen von Vordergrundregionen an

- ▶ konkaven Randbereichen
- ▶ Löchern
- ▶ Dilatation mit quadratischem Strukturelement:  
Maximum-Filter

# Beziehungen zwischen Erosion und Dilatation

Für alle Bilder  $B \in \{0, 1\}^{\text{pos}}$  und  
alle Strukturelemente  $M, M' \in \{0, 1\}^{\text{pos}'}$  mit 1 im Referenzpunkt  
gilt:

$$\begin{aligned} B \ominus M &\subseteq B \subseteq B \oplus M \\ (B \ominus M) \ominus M' &= B \ominus (M \oplus M') \\ \overline{B \ominus M} &= \overline{B} \oplus M \\ \overline{B \oplus M} &= \overline{B} \ominus M \end{aligned}$$

Kombinationen von Erosions- und Dilatationschritten:  
Opening, Closing

# Morphologisches Öffnen (Opening)

gegeben:  $B : \text{pos} \rightarrow \{0, 1\}$

Strukturelement (Maske)  $M : \text{pos}' \rightarrow \{0, 1\}$

Opening

$$B \circ M = (B \ominus M) \oplus M'$$

$M'$ : Strukturelement  $M$  am Referenzpunkt gespiegelt  
(um  $180^\circ$  gedreht)

Wirkung:

- ▶ Dilatation macht die Verkleinerung der die Erosion überlebenden (hellen) Regionen rückgängig
- ▶ Beseitigung von kleinen oder schmalen hellen Bereichen (falsch als Vordergrund interpretierte Stellen)
- ▶ Trennung von gerinfügig verbundenen (hellen) Regionen

## Opening – Beispiel

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline \end{array} \quad M = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array} \quad B \ominus M = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$B \circ M = (B \ominus M) \oplus M' \quad \text{mit} \quad M' = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array} \quad B \circ M = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & \mathbf{0} & 0 & 0 & 0 \\ \hline 0 & \mathbf{0} & 0 & 0 & 0 \\ \hline \end{array}$$

# Morphologisches Schließen (Closing)

gegeben:  $B : \text{pos} \rightarrow \{0, 1\}$

Strukturelement (Maske)  $M : \text{pos}' \rightarrow \{0, 1\}$

Closing

$$B \bullet M = (B \oplus M) \ominus M'$$

$M'$ : Strukturelement  $M$  am Referenzpunkt gespiegelt  
(um  $180^\circ$  gedreht)

Wirkung:

- ▶ Erosion macht die Ausdehnung der (hellen) Regionen bei Dilation rückgängig
- ▶ Beseitigung von kleinen oder schmalen dunklen Bereichen (Störungen in Vordergrund-Regionen)
- ▶ Verbindung nahe benachbarter Vordergrundregionen

## Closing – Beispiel

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline \end{array} \quad M = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array} \quad B \oplus M = \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 1 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \\ \hline \end{array}$$

$$B \bullet M = (B \oplus M) \ominus M' \quad \text{mit} \quad M' = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array} \quad B \bullet M = \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ \hline \end{array}$$

# Beziehungen zwischen morphologischem Öffnen und Schließen

Für alle Bilder  $B, C \in \{0, 1\}^{\text{pos}}$  und alle Strukturelemente  $M \in \{0, 1\}^{\text{pos}'}$  gilt

$$\begin{array}{ll} B \circ M \subseteq B & B \subseteq B \bullet M \\ (B \circ M) \circ M = (B \circ M) & (B \bullet M) \bullet M = (B \bullet M) \\ \overline{A \circ M} = \overline{A} \bullet M & \overline{A \bullet M} = \overline{A} \circ M \\ A \circ M = \overline{\overline{A} \bullet M} & A \bullet M = \overline{\overline{A} \circ M} \\ B \subseteq C \text{ gdw.} & (B \circ M \subseteq C \circ M) \wedge (B \bullet M \subseteq C \bullet M) \end{array}$$

Beseitigung größerer Störungen durch geeignete Folgen von Öffnen- und Schließen-Schritten

# Morphologische Operationen auf Grauwertbildern

Bild  $\text{col}^{\text{pos}}$  mit Farbbereich  $\text{col} = \{0, \dots, k\}$  statt  $\{0, 1\}$

Übergang von charakteristischer Funktion  $B : \text{pos} \rightarrow \{0, 1\}$  zu Funktion  $B : \text{pos} \rightarrow \{0, \dots, k\}$

geeignete Erweiterungen der Mengenoperationen  
(punktweise Boolesche Operationen)  
vom Farbbereich  $\{0, 1\}$  auf  $\{0, \dots, k\}$ , z.B.:

$$\wedge \mapsto \min$$

$$\vee \mapsto \max$$

$$\neg \mapsto (x \mapsto k - x)$$

(für Spezialfall  $k = 1$  genau die Booleschen Operationen)

## Beispiel: Erosion von Grauwertbildern

gegeben:  $B : \text{pos} \rightarrow \{0, \dots, k\}$

Strukturelement (Maske)  $M : \text{pos}' \rightarrow \{0, 1\}$

Erosion:

$$\forall p \in \text{pos} : (B \ominus M)(p) = \min\{B(p + v) \mid M(v) = 1\}$$

Wirkung:

Verdunklung der hellen Regionen (Vordergrund)  
(ähnlich Minimum-Filter mit Strukturelement)

Dilatation:

$$\forall p \in \text{pos} : (B \oplus M)(p) = \max\{B(p + v) \mid M(v) = 1\}$$

Opening, Closing damit analog zu Binärbildern

# Was bisher geschah

- ▶ digitale Bilder  $B : \text{pos} \rightarrow \text{col}$
- ▶ statistische Merkmale (Bildanalyse)
- ▶ Punktoperationen  $f : \text{col}_1 \rightarrow \text{col}_2$  (Farbtransformation)
- ▶ Geometrische Transformationen  $f : \text{pos}_1 \rightarrow \text{pos}_2$
- ▶ Interpolation der Intensitäten
- ▶ Digitalisierung (Abtastung) von 1D- und 2D-Signalen
- ▶ Faltung von Signalen (Funktionen)
- ▶ Fourier-Transformierte / inverse Fourier-Transformierte
- ▶ Operationen im Frequenzbereich, Filter
- ▶ Morphologische Operationen auf Binärbildern:  
Erosion, Dilatation, Opening, Closing

# Hit-or-Miss-Transformation auf Binärbildern

Ziel:

Identifizierung von im Bild vorkommenden Pixelmustern mit Toleranz, z.B. Linien der Länge von ... bis ...

Idee: Schnitt von

1. Erosion des Bildes mit Positiv-Maske  $M_1$  zur Beseitigung aller zu kleinen Vordergrund-Regionen (Hit-Transformation)
2. Erosion des invertierten Bildes mit Negativ-Maske (Umriss)  $M_2$  zur Beseitigung aller zu großen Vordergrund-Regionen (Miss-Transformation)

gegeben:  $B : \text{pos} \rightarrow \{0, 1\}$

Strukturelemente  $M_1 : \text{pos}_1 \rightarrow \{0, 1\}$ ,

$M_2 : \text{pos}_2 \rightarrow \{0, 1\}$  mit  $M_1 \cap M_2 = \emptyset$

Hit-or-Miss-Transformation:

$$B \otimes (M_1, M_2) = (B \ominus M_1) \cap (\overline{B} \ominus M_2)$$

# Beispiel

$$B =$$

0	0	0	0	0
0	0	1	1	0
0	0	0	1	0
0	0	0	0	0
0	1	1	1	0
0	0	0	0	0
0	0	1	1	0
0	0	0	0	0

$$M_1 =$$

0	0	0	0
0	1	1	0
0	0	0	0

$$B \ominus M_1 =$$

0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0
0	1	1	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0

$$\bar{B} =$$

1	1	1	1	1
1	1	0	0	1
1	1	1	0	1
1	1	1	1	1
1	0	0	0	1
1	1	1	1	1
1	1	0	0	1
1	1	1	1	1

$$M_2 =$$

0	1	1	0
1	0	0	1
0	1	1	0

$$\bar{B} \ominus M_2 =$$

1	0	0	0	1
0	0	0	0	0
1	0	0	0	0
1	0	0	0	1
0	0	0	0	0
0	0	0	0	1
0	0	1	0	0
1	0	0	0	1

(hier mit Randfortsetzung in  $B$  durch 0 und  $\bar{B}$  durch 1)

$$B \otimes (M_1, M_2) = (B \ominus M_1) \cap (\bar{B} \ominus M_2) = \{(6, 2)\}$$

# Hit-or-Miss-Transformation

Hit-or-Miss-Transformation:

$$\begin{aligned} B \otimes (M_1, M_2) &= (B \ominus M_1) \cap (\overline{B} \ominus M_2) \\ &= (B \ominus M_1) \cap \overline{B \oplus M_2} \end{aligned}$$

Zusammenfassung von

- ▶ Hit-Maske  $M_1 : \text{pos}' \rightarrow \{0, 1\}$  und
- ▶ Miss-Maske  $M_2 : \text{pos}' \rightarrow \{0, 1\}$

mit  $\text{pos}' = \{-m, \dots, m\} \times \{-n, \dots, n\}$

(gleiche Größe, ggf. um 0 erweitern)

zu spezieller Hit-or-Miss-Maske  $M : \text{pos}' \rightarrow \{0, 1, *\}$  möglich:

$$\forall p \in \text{pos}' : M(p) = \begin{cases} M_1(p) & , \text{ falls } M_2(p) = 1 - M_1(p) \\ * & , \text{ sonst} \end{cases}$$

# Hit-or-Miss-Transformation – Beispiel

$$M_1 = \begin{array}{|c|c|c|c|} \hline 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$M_2 = \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 1 & 0 \\ \hline \end{array}$$

$$M = \begin{array}{|c|c|c|c|} \hline 0 & 0 & 0 & * \\ \hline * & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & * \\ \hline \end{array}$$

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$B \otimes M = \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

(hier mit Randfortsetzung in  $B$  durch 0)

# Hit-or-Miss-Transformation

ermöglicht Erkennung spezieller Strukturen im Bild, z.B.

$$M_1 = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad \text{isolierte Punkte}$$

$$M_2 = \begin{array}{|c|c|c|} \hline * & 1 & 0 \\ \hline 1 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad \text{untere rechte Ecken}$$

$$M_3 = \begin{array}{|c|c|c|} \hline 0 & * & 1 \\ \hline 0 & 1 & 1 \\ \hline 0 & * & 1 \\ \hline \end{array} \quad \text{linker Rand, Zusammenhang-erhaltend}$$

$$M_4 = \begin{array}{|c|c|c|} \hline 1 & 1 & * \\ \hline 1 & 0 & * \\ \hline 1 & * & 0 \\ \hline \end{array} \quad \text{Thickening: } B \cup (B \otimes M) \quad (\text{konvexe Hülle})$$

$$M_5 = \begin{array}{|c|c|c|} \hline 0 & * & * \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad \text{untere Endpunkte}$$

# Wirkung morphologischer Operationen

gegeben: Bild  $B : \text{pos} \rightarrow \{0, 1\}$   
Strukturelement  $M : \text{pos}' \rightarrow \{0, 1\}$ ,  
(markiert relevante Pixel)

**Erosion:** Abtragung (Verkleinerung der Vordergrund-Regionen)  
Minimum über alle relevanten Pixel

**Dilatation:** Anlagerung (Vergrößerung der Vordergrund-Regionen)  
Maximum über alle relevanten Pixel

**Opening:** morphologisches Öffnen,  

1. Entfernen kleiner Vordergrundeobjekte
2. Wiederherstellung der größeren Objekte

auch zum Finden von Regionen bestimmter Form

**Closing:** morphologisches Schließen  

1. Entfernen kleiner Vordergrund-Störungen
2. Wiederherstellung des Vordergrund ohne Störung

**Hit-or-Miss:** Finden von Regionen bestimmter Form  
mit Toleranzbereich

## Regionen im Bild

Region  $R \subseteq \text{pos}$  im Bild  $B : \text{pos} \rightarrow \{0, 1\}$

(bzgl. einer Nachbarschaft zusammenhängend)

Annahmen:

- ▶ jede Region hat eine Bedeutung,
- ▶ Positionen in einer Region gehören zur selben Bedeutung

Erkennung von Regionen (semantische Einheiten) durch

**Gemeinsamkeiten** aller Positionen(-gruppen) innerhalb der  
Region, z.B. Intensität, Farbe, Textur

**Unterschieden** zwischen benachbarten Positionen(-gruppen),  
Erkennung der Grenzen der Region

Merkmale von Regionen  $R$ :

**Kontur**  $K(R)$ : Menge der Randpixel  
(verschiedene Codierungen,  
z.B. Freeman-Kettencode, Differential-Code)

**Skelett**  $S(R)$ : topologieerhaltendes Bild aus  
Linien von 1 Pixel Dicke

# Beschreibung von Regionen

Darstellung als

- ▶ Menge aller zur Region gehörenden Positionen
- ▶ Kontur (Positionen am Rand der Region)
- ▶ Skelett (zentrale Positionen der Region)

Verfahren zur Erkennung zusammenhängender Regionen

Menge von Positionen:

1. Schwellwert-Verfahren,
2. Identifikation zusammenhängender Positionsmengen

**Kontur:** Konturbestimmung

**Skelett:** Skelettierung

morphologische Operationen hilfreich

# Konturen von Regionen

Region  $R \subseteq \text{pos}$

Rand (Kontur) von  $R$ :

$K(R)$  Menge aller Positionen  $p \in \text{pos}$  mit

1.  $p \in R$  und
2. in der (geeignet definierten) Nachbarschaft von  $p$  existiert (wenigstens) ein  $q$  mit  $q \notin R$

Rand in

- ▶ 4-Nachbarschaft:  
geschlossener Pfad in 8-Nachbarschaft
- ▶ 8-Nachbarschaft:  
geschlossener Pfad in 4-Nachbarschaft

# Konturbestimmung

$$K = B \setminus (B \ominus M) = B \cap \overline{B \ominus M}$$

Beispiel:

Bestimmung des Randes in 4- bzw. 8-Nachbarschaft durch

1. Erosion mit Strukturelement

$$M_4 = \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 1 & 1 & 1 \\ \hline 0 & 1 & 0 \\ \hline \end{array}$$

$$M_8 = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

(entfernt alle Pixel, deren 4- bzw. 8-Nachbarschaft wenigstens einen Punkt außerhalb der Region enthält)

2. Differenz zwischen Original- und erodiertem Bild enthält genau alle Randpunkte (innerhalb der Region)

analog lässt sich definieren:

Randpunkte außerhalb der Region

$$K' = (B \oplus M) \setminus B = (B \oplus M) \cap \overline{B}$$

# Konturbestimmung – Beispiel

$$B = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	0	0	0	1	1	1	1	1	0
0	1	1	1	1	1	0	0	0	1	1	0	0	0
0	0	1	1	1	0	0	0	1	1	1	0	0	0
0	1	1	1	0	0	1	0	0	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0

$$B \ominus M_4 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	**0**	**0**	**0**	**0**	0	0	0	**0**	**0**	**0**	**0**	**0**	0
0	**0**	1	1	1	**0**	0	0	0	**0**	**0**	0	0	0
0	0	**0**	1	**0**	0	0	0	**0**	1	**0**	0	0	0
0	**0**	**0**	**0**	0	0	**0**	0	0	**0**	**0**	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0

$$B \setminus (B \ominus M_4) = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	0	0	0	1	1	1	1	1	0
0	1	0	0	0	1	0	0	0	1	1	0	0	0
0	0	1	0	1	0	0	0	1	0	1	0	0	0
0	1	1	1	0	0	1	0	0	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0

# Abstände zwischen Positionen

Abstandsfunktion  $d : \text{pos} \rightarrow \mathbb{R}$  mit den Eigenschaften:

- ▶  $\forall p, q \in \text{pos} : d(p, q) = 0$  gdw.  $p = q$
- ▶  $\forall p, q \in \text{pos} : d(p, q) = d(q, p)$  (kommutativ)
- ▶  $\forall p, q, r \in \text{pos} : d(p, q) + d(q, r) \geq d(p, r)$  (Dreiecksungleichung)

Beispiele: für  $\text{pos} = \{0, \dots, m\} \times \{0, \dots, n\}$  und  $k \in \mathbb{N}$ :

$$\forall p = (p_z, p_s), q = (q_z, q_s) \in \text{pos} : d_k(p, q) = \sqrt[k]{(p_z - q_z)^k + (p_s - q_s)^k}$$

- ▶ für  $k = 2$ :  $d_2(p, q) = \sqrt{(p_z - q_z)^2 + (p_s - q_s)^2}$   
Euklidischer Abstand
- ▶ für  $k = 1$ :  $d_1(p, q) = |p_z - q_z| + |p_s - q_s|$   
Manhattan-Metrik (minimale Anzahl der zu einem  $p$  und  $q$  verbindenden Pfad in 4-Nachbarschaft notwendigen Positionen)
- ▶ für  $k \rightarrow \infty$ :  $d_\infty(p, q) = \max(|p_z - q_z|, |p_s - q_s|)$   
Maximum-Metrik (minimale Anzahl der zu einem  $p$  und  $q$  verbindenden Pfad in 8-Nachbarschaft notwendigen Positionen)

## Distanztransformation von Binärbildern

ordnet jeder Position innerhalb einer Region den kürzesten Abstand zum Rand der Region zu

(Abstand bzgl. gegebener Nachbarschaftsrelation)

Idee:

- ▶ wiederholte Erosion des Binärbildes  $B \subseteq \text{pos}$
- ▶ Distanzmaß ordnet jeder Position  $p$  in einer Region  $R \subseteq B^{-1}(1)$  die Anzahl der Erosionen bis zur Entfernung von  $p$  aus  $B$  zu  
(geeignet definierter Mindestabstand zum Rand  $K(R)$ )

$d_B : \text{pos} \rightarrow \mathbb{N}$  mit

$$\forall p \in \text{pos} : d_B(p) = \begin{cases} 0 & , \text{ falls } p \notin B \\ 1 & , \text{ falls } p \in B \setminus B \ominus M \\ \vdots & \\ k + 1 & , \text{ falls } p \in B(\ominus M)^k \setminus B(\ominus M)^{k+1} \end{cases}$$

übliche Visualisierung: Distanzbild

Intensität von  $p \in B$  umgekehrt proportional zu  $d_B(p)$

# Distanztransformation – Beispiel

	×	×	×	×				×	×	×	×	×	
	×	×	×	×	×	×			×	×			
		×	×	×	×	×		×	×	×			
	×	×	×	×	×		×	×	×	×			
	×	×	×	×					×	×			

0	1	0
1	1	1
0	1	0

	1	1	1	1				1	1	1	1	1	
	1	2	2	2	1	1			1	1			
		1	2	3	2	1		1	2	1			
	1	2	2	2	1		1	1	2	1			
	1	1	1	1					1	1			

# Skelette von Regionen

Ziel:

Reduktion der Vordergrundregionen zu Linien unter Beibehaltung wichtiger topologischer Merkmale (z.B. Zusammenhang, Löcher)

Anwendungen zur Beschreibung von Regionen, für die das Vorhandensein von „Linien“ relevant ist, deren Breite jedoch nicht

z.B. Straßen-, Gesten-, Zeichenerkennung (OCR)

# Skelett von Regionen – Anforderungen

- ▶ Skelett ist Teilmenge der Region
- ▶ Erhaltung der Zusammenhangseigenschaften der Region
- ▶ Skelett ist minimal, d.h. keine weiteren Punkte können entfernt werden, ohne Zusammenhang zu zerstören
- ▶ bevorzugt Mittelachsen  $M$  der Vordergrundregionen (gleicher Abstand zu den nächsten Randpunkten)

## Skelettierung – Idee

- ▶ mehrfach wiederholtes Entfernen von (weißen Vordergrund-) Positionen, bis nur noch Skelett weiß
- ▶ nur Entfernung der Punkte, die nicht zum Skelett gehören
- ▶ Identifikation typischer Pixelmuster für Skelettlinien (Enden, Verzweigungen von Linien von 1 Pixel Breite)
- ▶ Identifikation typischer Pixelmuster (Enden, Verzweigungen) für Linien größerer Breite
- ▶ Hit-or-Miss-Transformation mit geeigneten Strukturelementen (Pixelmustern) zur Identifikation „zu breiter“ Bereiche
- ▶ Entfernen der gefundenen Positionen aus dem Bild

# Thinning (Verdünnung)

Strukturelemente zur Identifikation „zu breiter“ Bereiche

$$M_o = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline * & 1 & * \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$$M_r = \begin{array}{|c|c|c|} \hline 1 & * & 0 \\ \hline 1 & 1 & 0 \\ \hline 1 & * & 0 \\ \hline \end{array}$$

$$M_u = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline * & 1 & * \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

$$M_l = \begin{array}{|c|c|c|} \hline 0 & * & 1 \\ \hline 0 & 1 & 1 \\ \hline 0 & * & 1 \\ \hline \end{array}$$

$$M_{or} = \begin{array}{|c|c|c|} \hline * & 0 & 0 \\ \hline 1 & 1 & 0 \\ \hline * & 1 & * \\ \hline \end{array}$$

$$M_{ur} = \begin{array}{|c|c|c|} \hline * & 1 & * \\ \hline 1 & 1 & 0 \\ \hline * & 0 & 0 \\ \hline \end{array}$$

$$M_{ul} = \begin{array}{|c|c|c|} \hline * & 1 & * \\ \hline 0 & 1 & 1 \\ \hline 0 & 0 & * \\ \hline \end{array}$$

$$M_{ol} = \begin{array}{|c|c|c|} \hline 0 & 0 & * \\ \hline 0 & 1 & 1 \\ \hline * & 1 & * \\ \hline \end{array}$$

Thinning:

Wiederholte Folge dieser 8 Hit-or-Miss-Transformationen nacheinander (4 achsenparallel und 4 diagonal) zur Entfernung von Randpositionen ohne Verletzung des Zusammenhanges

## Thinning – Schritte

gegeben: Region  $R \subseteq \text{pos}$

Verfahren:

- ▶ in jedem Schritt:  
für jede der 8 Hit-or-Miss-Masken  $M_i$  (nacheinander):
  1. Hit-or-Miss-Transformation:  $(B \otimes M_i)$
  2. passende Bereiche aus  $B$  entfernen:  $B \setminus (B \otimes M_i)$
- ▶ Verfahren endet, wenn in einem Schritt  
(Anwendung aller 8 Hit-or-Miss-Transformationen)  
kein Pixel geändert wurde

mögliche Vor- und Nachbereitung:

- ▶ vor Beginn: Distanztransformation  
(Speichern der Distanzwerte für jedes  $p \in R$ )
- ▶ vor Ende:  
Distanzwert jedes Skelettpunktes  $p$  ist Radius des größten  
Kreises um  $p$ , der vollständig in  $B$  enthalten ist
- ▶ häufige Nachbearbeitung bei Skelettierung:  
Entfernung von Verzweigungen an den Linienenden

# Thinning – Anwendungen

- ▶ Skelettierung
- ▶ nach Detektion von Randbereichen (z.B. mit Sobel-Filter):  
Verdünnung der Kanten
- ▶ Schrifterkennung

## Probleme:

- ▶ empfindlich gegenüber kleinen Störungen,  
z.B. Rauschen, Binarisierungsungenauigkeiten
- ▶ erzeugtes Skelett enthält kurze Ausläufer an den Enden

# Thickening (Verdickung)

analog zum Thinning

Auffüllen konkaver Randbereiche durch wiederholte Anwendung einer Folge von 8 Hit-or-Miss-Transformationen, z.B.

$$\begin{array}{l} M_1 = \begin{array}{|c|c|c|} \hline * & * & 0 \\ \hline 1 & 0 & * \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad M_2 = \begin{array}{|c|c|c|} \hline 0 & * & * \\ \hline * & 0 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad M_3 = \begin{array}{|c|c|c|} \hline 1 & * & 0 \\ \hline 1 & 0 & * \\ \hline 1 & 1 & * \\ \hline \end{array} \quad M_4 = \begin{array}{|c|c|c|} \hline 0 & * & 1 \\ \hline * & 0 & 1 \\ \hline * & 1 & 1 \\ \hline \end{array} \\ M_5 = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 0 & * \\ \hline * & * & 0 \\ \hline \end{array} \quad M_6 = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline * & 0 & 1 \\ \hline 0 & * & * \\ \hline \end{array} \quad M_7 = \begin{array}{|c|c|c|} \hline * & 1 & 1 \\ \hline * & 0 & 1 \\ \hline 0 & * & 1 \\ \hline \end{array} \quad M_8 = \begin{array}{|c|c|c|} \hline 1 & 1 & * \\ \hline 1 & 0 & * \\ \hline 1 & * & 0 \\ \hline \end{array} \end{array}$$

Ergebnis: Approximation der konvexen Hüllen von Regionen

# Was bisher geschah

- ▶ Definition digitaler Bilder  $B : \text{pos} \rightarrow \text{col}$
- ▶ Bildanalyse, statistische Merkmale
- ▶ Signale im Orts- und Frequenzraum
- ▶ Bildbearbeitung durch
  - ▶ Punktoperationen (Farbtransformation)  $f : \text{col}_1 \rightarrow \text{col}_2$   
(punktweise Fortsetzung auf Gesamtbild)
  - ▶ geometrische Transformationen  
(Koordinatentransformation)  $f : \text{pos}_1 \rightarrow \text{pos}_2$
  - ▶ lokale Operationen (abhängig von Nachbarschaft):  
Filter, morphologische Operationen
- ▶ Merkmale von Regionen im Bild:
  - ▶ Kontur
  - ▶ Distanzbild
  - ▶ Skelett

# Regionen

**Region:** zusammenhängende Gruppe von Positionen

**Segmentierung:** Extraktion und Trennung von Regionen im Bild

Annahmen:

- ▶ jede Region hat eine Bedeutung,
- ▶ Positionen in einer Region haben dieselbe Bedeutung,
- ▶ verschiedene Regionen haben verschiedene Bedeutung  
z.B. enthalten verschiedene Objekte, Teile

Erkennung von Regionen (semantische Einheiten) durch

- ▶ gemeinsame Eigenschaften aller Pixel oder Pixelgruppen der Region, z.B. Intensität, Farbe, Textur
- ▶ signifikanten Unterschieden zwischen benachbarten Positionen  
(Erkennung der Konturen)

# Segmentierung – Anforderungen

gegeben: Bild  $B : \text{pos} \rightarrow \text{col}$

gesucht: endliche Menge  $R = \{r_1, \dots, r_n\}$  von Regionen  
mit  $\forall i \in \{1, \dots, n\} : r_i \subseteq \text{pos}$   
und den folgenden Eigenschaften:

- ▶ Menge  $R = \{r_1, \dots, r_n\}$  aller Regionen bildet eine **Zerlegung** von  $\text{pos}$ , d.h.

1.  $\bigcup_{i \in \{1, \dots, n\}} r_i = \text{pos}$

2.  $\forall i, j \in \{1, \dots, n\} : i \neq j \rightarrow r_i \cap r_j = \emptyset$

- ▶ Jede Region  $r \in R$  ist **zusammenhängend**  
(bildet eine Zusammenhangskomponente im  
Nachbarschaftsgraphen)

Darstellung der Segmentierung (Zerlegung):

- ▶ Region  $r \subseteq \text{pos}$ : endliche Menge von Bild-Positionen,  
Segmentierung  $R = \{r_1, \dots, r_n\}$ : endliche Menge von Regionen  
(endlichen Mengen von Bild-Positionen),
- ▶ Zuordnung  $r : \text{pos} \rightarrow R$  von Bild-Positionen  $p \in \text{pos}$  zu Regionen

# Segmentierung – Verfahren

**Punkt-basiert:** Zuordnung  $pos \rightarrow R$  anhand des Wertes  $B(p) \in col$  und evtl. Merkmale des Bildes / der Region  
(Schwellwertverfahren, evtl. adaptiv, verschiedene Schwellwerte für verschiedene Bildbereiche)

**Zusammenhangs-basiert:** Zuordnung  $pos \rightarrow R$  anhand von Merkmalen in Nachbarschaft von  $p \in pos$

**Regionen-orientiert:** Erkennung zusammenhängender Regionen (Homogenität)  
meist aus dem Inneren heraus

**Form-orientiert:** Finden von Regionen bekannter Form und Ausdehnung  
(z.B. durch morphologische Operationen mit geeigneten Strukturelementen:  
Opening, Hit-or-Miss)

**Kanten-orientiert:** Erkennung der Regionengrenzen (Diskontinuität)  
Erkennung von Randpositionen z.B. durch Fourier-Transformation, Filter, morphologische Operationen

# Homogenitätskriterien

## Ähnlichkeiten

- ▶ einzelner Positionen, z.B.
  - ▶ Intensität
  - ▶ Farbe (Kombination von Intensitäten)  
z.B. Mittelwerte, Intensitäts-Verhältnisse
- ▶ Texturen (charakteristische Intensitätsverteilungen) von Gruppen von Positionen  
häufig verwendete Texturmerkmale:
  - ▶ statistische Merkmale
  - ▶ Co-Occurrence-Matrix
  - ▶ Fourier-Spektrum

## Gebietswachstum (Region Growing, flood fill)

zur Erkennung zusammenhängender Regionen

Idee: Wiederholung der folgenden Schritte, solange noch Positionen ohne Regionenzuordnung existieren:

1. Beginn mit beliebiger (noch nicht einer Region zugeordneter) **Startposition**  $p \in \text{pos}$ ,  
für die auch noch kein Nachbar einer Region zugeordnet ist,
2. neue Region  $r$  initialisieren,  $p$  zu  $r$  hinzufügen
3. rekursiv:  
alle zu  $p$  **ähnlichen Nachbarn** von  $p$  zu  $r$  hinzufügen

abhängig von:

- ▶ Nachbarschaftsrelation  $N \subseteq \text{pos} \times \text{pos}$
- ▶ Ähnlichkeitsrelation  $S \subseteq \text{col} \times \text{col}$  (bzw.  $\text{col}^{m \times n} \times \text{col}^{m \times n}$ )  
z.B. Intensität, Farbe, Textur
- ▶ Startpositionen  
z.B. Position mit minimalem Abstand von  $(0, 0) \in \text{pos}$ ,  
interaktiv, zufällig, Erfahrungswerte, Cluster-Schwerpunkte

je nach Anwendung, Bildqualität, ... verschieden

# Markierung zusammenhängender Regionen

(Region Labeling, Connected component labeling)

- ▶ Bestimmen und gleichzeitiges Zählen der Regionen
- ▶ ordnet jeder Region (Vordergrund) eines Binärbildes eine eindeutige Markierung zu

Idee:

- ▶ einfache Modifikation des Region-Growing-Verfahrens
- ▶ vor Beginn Initialisierung aller Positionen mit Markierung 0 (noch nicht zugeordnet)
- ▶ für jede neu angelegte Region Markierung (Zähler) um 1 erhöhen
- ▶ Markierung an hinzugefügte Nachbarn übergeben

praktisch häufig (zeilenweise) Markierung in zwei Phasen:

1. provisorische Markierung
2. Verschmelzen verbundener Regionen (Ummarkierung)

# Multiskalen-Strategien

Ziel: Bildanalyse in verschiedenen Größen (Auflösungsstufen)

- ▶ wenig aufwendige Untersuchung grober Strukturen
- ▶ feinere Untersuchung feiner Strukturen im Bild

Idee: Multiskalenraum (ähnlich menschlicher Wahrnehmung)

enthält Originalbild  $B_0$  und weniger detaillierte Versionen  $B_k$

Erzeugung der Multiskalen-Bilder (Pyramiden) ( $B_0, B_1, B_2, \dots$ )

durch wiederholte Ausführung der Schrittfolge

1. Glättung (durch Tiefpass-Filter, z.B. Gauß-Filter)
2. **reduce**: Komprimierung durch geringere Abtastrate, z.B. Gauß-Pyramide: Löschen jeder zweiten Zeile und Spalte
3. **expand**: Umkehrung durch Interpolation (nicht verlustfrei) zur Vergleichbarkeit der Bilder verschiedener Auflösungen

Laplace-Pyramide (redundanzarme Darstellung):

Schichten enthalten Differenz zwischen Bild  $B_k$  und

$\text{expand}(\text{reduce } B_k)$

# Region-Merging

zur Erkennung zusammenhängender Regionen

Idee: Regionen-Adjazenz-Graph (RAG)

- ▶ Beginn: jede Bildposition ist eine Region (Knoten) und hat Kanten zu jedem Nachbarn
- ▶ wiederholte Vereinigung benachbarter Regionen mit gleichen / ähnlichen Eigenschaften (Homogenitätsbed.), (Verschmelzen der Knoten benachbarter ähnlicher Regionen)
- ▶ Ende, sobald keine Änderung mehr eintritt (nur noch Kanten zwischen unähnlichen Regionen)

nebenbei möglich:

Zuordnung eindeutiger Markierungen zu den Regionen  
(Region Labeling)

und damit auch das Zählen der Regionen  
(Anzahl der verbliebenen Markierungen)

# Split-and-Merge-Verfahren

zur Erkennung zusammenhängender Regionen

Idee: Region-Merging mit Multiskalen-Strategie

Region-Merging mit Regionen verschiedener Größen

- ▶ Beginn: gesamtes Bild bildet eine Region
- ▶ wiederholte Ausführung folgender Schritte:
  - split**: rekursive Zerlegung inhomogener Regionen in Teilbilder (Quadranten)  
(Quad-Tree mit Regionen als Blättern)
  - merge** : Vereinigung benachbarter Regionen mit gemeinsamer Eigenschaft, solange möglich
- ▶ Ende, sobald keine Änderung mehr eintritt

# Extraktion von Konturen

Darstellung von Konturen:

gegeben: Menge von Positionen  $P \subseteq \text{pos}$   
(mit großen Unterschieden zu Nachbarpositionen)

gesucht: geschlossene Kontur  
zusammenhängende Menge  
(bzgl. einer Nachbarschaft)

Repräsentation von Konturen:

- ▶ Menge aller Kontur-Positionen  $K \subseteq \text{pos}$
- ▶ Freeman-Kettencode
- ▶ Polygonzug (Folge von Streckenabschnitten)
- ▶ Folge von Kurvenabschnitten

# Repräsentation von Konturen durch Kettencodes

Folge von Richtungen beim Umrunden der Kontur  
(entgegen Uhr)

$$N_4(p) = \begin{array}{|c|c|c|} \hline & 1 & \\ \hline 2 & \bullet & 0 \\ \hline & 3 & \\ \hline \end{array} \quad N_8(p) = \begin{array}{|c|c|c|} \hline 3 & 2 & 1 \\ \hline 4 & \bullet & 0 \\ \hline 5 & 6 & 7 \\ \hline \end{array}$$

Beispiel:

		x	x	x	
	x	x	x	x	
	x	x	x		
		x	x	x	
	x	x	x	x	
	x	x	x	x	

		2	2	•	
	2	3	1	0	
	3	0	1		
		3	2	1	
	2	3	x	1	
	3	0	0	0	

		4	4	•	
	5	x	x	1	
	6	x	3		
		7	x	2	
	5	x	x	2	
		7	0	0	

		0	0	•	
	1	x	x	-2	
	1	x	1		
		1	x	0	
	-2	x	x	2	
		2	1	0	

$$C_4(R) = 22323033230001121101, \quad C_8(R) = 44567570022312$$

Differentieller Kettencode (Richtungswechsel):

$$D_8(R) = 00111(-2)210201(-2)1$$

ermöglicht Kompression und einfache Nachbearbeitung, z.B.

Konturglättung, Bestimmung von Randpunkten hoher Krümmung

# Was bisher geschah

- ▶ Definition digitaler Bilder  $B : \text{pos} \rightarrow \text{col}$
- ▶ Bildanalyse, statistische Merkmale
- ▶ Signale im Orts- und Frequenzraum
- ▶ Bildbearbeitung durch
  - ▶ Punktoperationen (Farbtransformation)  $f : \text{col}_1 \rightarrow \text{col}_2$   
(punktweise Fortsetzung auf Gesamtbild)
  - ▶ geometrische Transformationen  
(Koordinatentransformation)  $f : \text{pos}_1 \rightarrow \text{pos}_2$
  - ▶ lokale Operationen (abhängig von Nachbarschaft):  
Filter, morphologische Operationen
- ▶ Segmentierung: Erkennung von Regionen im Bild
  - ▶ punktbasiert, z.B. Schwellwertverfahren
  - ▶ regionenbasiert, z.B. region growing
- ▶ Regionen-Adjazenz-Graph

# Modellbasierte Segmentierung

gezielte Suche nach Regionen mit bekannten Merkmalen im Bild

übliche Verfahren verwenden

- ▶ Merkmale der gesuchten Regionen  
(z.B. durch Hit-and-Miss-Transformation, Ähnlichkeit der topologischen, geometrischen und statistischen Merkmale)
- ▶ Merkmale der Kontur der gesuchten Regionen,  
(z.B. durch Ähnlichkeit der Kettencodes)
- ▶ Merkmale von Standard-Konturteilen,  
z.B. (Ausschnitte von) Geraden, Kreise, Ellipsen  
(Hough-Transformation)

## Finden von Geraden – Idee

Annahme (Modell):

gesuchte Konturteile = Geraden(abschnitte) im Bild

Idee (Voting-Verfahren):

- ▶ Gerade in  $\mathbb{R}^2$ :  $g = \{(x, y) \in \mathbb{R}^2 \mid y = mx + n\}$
- ▶ **Voting**: für jede potentielle Gerade  $g$  im Bild  $B$ :  
Zählen der Kontur-Positionen  $(p_x, p_y) \in \text{pos}$  mit  
 $p_y \approx mp_x + n$
- ▶ Geraden, auf denen sehr viele Konturpunkte liegen,  
enthalten wahrscheinlich Streckenabschnitte im Bild

Zählen in **Akkumulator-Zellen**:

Punkt  $(m, n) \in \mathbb{R}^2$  repräsentiert Gerade

$$g = \{(x, y) \in \mathbb{R}^2 \mid y = mx + n\}$$

Problem:

senkrechte Geraden nicht darstellbar,  
kommen aber häufig in Bildern vor

# Hough-Raum (für Geraden)

Idee (Hough, 1962): alternative Geraden-Darstellung  
(Hesse-Normalform)

$$g = \{ \cos(\alpha)x + \sin(\alpha)y - d = 0 \}$$

mit

$\alpha \in [0, 2\pi]$  – Winkel von  $g$  zur  $x$ -Achse (Normale)

$d \in \mathbb{R}_{\geq 0}$  – Abstand von  $g$  zu  $(0, 0) \in \mathbb{R}^2$

(maximaler Abstand durch Bildgröße bestimmt)

**Hough-Raum:**  $H = [0, 2\pi] \times \mathbb{R}_{\geq 0}$

- ▶ Geraden im  $\mathbb{R}^2$  repräsentiert durch Punkte  $(\alpha, d) \in H$
- ▶ Jeder Punkt  $p = (p_x, p_y) \in \mathbb{R}^2$  liegt auf unendlich vielen Geraden im  $\mathbb{R}^2$ .
- ▶ Jeder Punkt  $p = (p_x, p_y) \in \mathbb{R}^2$  entspricht der Kurve  $d(\alpha) = p_x \cos(\alpha) + p_y \sin(\alpha)$  im Hough-Raum  $H$ .
- ▶ Die Gerade durch  $p$  und  $q$  in  $\mathbb{R}^2$  entspricht dem Schnittpunkt der  $H$ -Kurven der Punkte  $p$  und  $q$ .
- ▶ Schnittpunkte vieler Kurven im Hough-Raum deuten auf Geraden(abschnitte) im Bild hin

# Hough-Transformation (für Geraden)

modifiziertes Voting-Verfahren:

- ▶ Akkumulator-Zellen (Repräsentation der Geraden):  
Hough-Raum  $H = [0, 2\pi] \times \mathbb{R}_{\geq 0}$
- ▶ für jede (potentielle) Kantenposition  $p \in \text{pos}$ :  
Wert jeder Akkumulatorzelle auf  $H$ -Kurve von  $p$  erhöhen
- ▶ hohe akkumulierte Werte in  $(\alpha, d) \in H$  entsprechen Geraden im Bild

Hough-Transformation ist robust gegenüber

- ▶ Unterbrechungen, verdeckte Bereiche,
- ▶ Rauschen

und damit geeignet z.B. zur

- ▶ Erkennung
- ▶ Ergänzung (Fortsetzung)

teilweise sichtbarer gerader Konturbereiche  
(z.B. Fahrbahnmarkierungen)

# Weitere Hough-Transformationen

analoge Verfahren für andere geometrische Formen

- ▶ Kreise

$$c = \{(x, y) \in \mathbb{R}^2 \mid (x - m_x)^2 + (y - m_y)^2 - r = 0\}$$

Parameterraum  $(m_x, m_y, r) \in \text{pos} \times \mathbb{R}_{\geq 0}$  (3dimensional)  
(Dimensionsreduktion bei bekanntem Radius möglich)

- ▶ Ellipsen (5dimensional)

$$c = \{(x, y) \in \mathbb{R}^2 \mid ax^2 + by^2 + cxy + dx + ey - 1 = 0\}$$

(Reduktion auf kleinere Dimensionen möglich)

- ▶ andere parametrisch darstellbare Kurven

# Was bisher geschah

- ▶ Definition digitaler Bilder  $B : \text{pos} \rightarrow \text{col}$
- ▶ Bildanalyse, statistische Merkmale
- ▶ Signale im Orts- und Frequenzraum
- ▶ Bildbearbeitung durch
  - ▶ Punktoperationen (Farbtransformation)  $f : \text{col}_1 \rightarrow \text{col}_2$   
(punktweise Fortsetzung auf Gesamtbild)
  - ▶ geometrische Transformationen  
(Koordinatentransformation)  $f : \text{pos}_1 \rightarrow \text{pos}_2$
  - ▶ lokale Operationen (abhängig von Nachbarschaft):  
Filter, morphologische Operationen
- ▶ Segmentierung: Erkennung von Regionen im Bild
  - ▶ punktbasiert, z.B. Schwellwertverfahren
  - ▶ regionenbasiert, z.B. region growing
  - ▶ modellbasiert, z.B. Hough-Transformation

# Kantenbasierte Segmentierung

Kanten / Konturen:

meist starke Unterschiede in Intensität, Farbe oder Textur innerhalb kleiner Nachbarschaften

Vorsicht:

Unterschiede haben mitunter andere Ursachen, z.B. Schatten

Schritte der kantenbasierten Segmentierung:

1. Kantenextraktion (Finden potentieller Kantenpunkte)  
z.B. durch Filter, morphologische Operationen  
Ergebnis: Menge von Positionen (evtl. gewichtet)
2. Kantenverfolgung:  
Zusammenfügen von Kanten zu geschlossenen Konturen von Regionen (dabei evtl. Lücken schließen)  
z.B. Edge linking, Canny-Edge-Detector
3. Bestimmung / Markierung der Regionen innerhalb der Konturen

# Edge-Linking

- gegeben: Bild  $B : \text{pos} \rightarrow \text{col}$ ,  
Menge  $P \subseteq \text{pos}$  von **potentiellen** Kantenpositionen  
(z.B. durch Filter, morphologische Operationen)
- gesucht: zusammenhängende Menge  $K \subseteq \text{pos}$   
von Positionen (Konturen)

Idee:

- ▶ systematische Untersuchung aller potentiellen Kantenpositionen in  $P$
- ▶ jeweils Suche nach Kantenpositionen in einer Nachbarschaft in bisheriger Kantenrichtung (parametrisiert durch Richtungsabweichung und Abstand)
- ▶ dabei Konstruktion einer zusammenhängenden Menge  $K$  der tatsächlichen Kantenpositionen  
evtl. auch durch Einfügen kurzer Verbindungen zwischen Kantenpositionen

# Edge-Linking-Algorithmus

- ▶ Beginn an beliebiger (potentieller) Kantenposition  $p \in P$
- ▶ Wiederholung, solange  $P \neq \emptyset$ 
  1. Auswahl eines beliebigen  $p \in P$ ,
  2.  $p$  zu (neuem) Konturabschnitt  $K$  hinzufügen,
  3. zu  $p$  (in bisheriger Kantenrichtung) benachbarte Positionen  $q \in \text{pos}$  mit ähnlichen Kanteneigenschaften (Richtung, Intensität) zu  $P$  und  $K$  hinzufügen,
  4.  $p$  aus  $P$  entfernen,
- ▶ Ende, wenn alle Kantenpositionen in  $P$  verarbeitet

Ergebnis abhängig von

- ▶ (richtungsabhängiger) Nachbarschaftrelation
- ▶ Ähnlichkeitbedingung

Nachbearbeitung möglich, z.B.

- ▶ isolierte Kantenpositionen entfernen
- ▶ Kantenverdünnung / Skelettierung

# Canny-Edge-Detector

1. Glättung mit Gauß-Filter (Tiefpass, Entrauschen)

z.B.  $\frac{1}{16}$ 

1	2	1
2	4	2
1	2	1

 oder  $\frac{1}{24}$ 

1	3	1
3	8	3
1	3	1

2. Kantendetektion (Sobel-Filter)

-1	0	1	1	2	1
-2	0	2	0	0	0
-1	0	1	-1	-2	-1

3. Bestimmung der Richtung (aus den in 2. berechneten partiellen Ableitungen) und der Stärke der Kante (Intensität)
4. Reduktion der Kanten auf Breite 1 (Pixel)  
Entfernen der Nachbarpixel senkrecht zur Kantenrichtung mit geringerer Intensität (non-maximal suppression)
5. Kantenverfolgung mit zwei Schwellwerten (Hysterese)  
zur Trennung zwischen sicherem / unsicherem / kein Kantenpixel
6. Zuordnung der unsicheren Kantenpunkte durch (Versuch der) Verbindung mit benachbarten sicheren Kantenpunkten  
Ausgangspunkt: sicherer Kantenpunkt  
unsichere 8-Nachbarn in Kante aufnehmen, wenn Verbindung nur durch sichere Nachbarn nicht möglich ist

# Klassen-Einteilung

**Merkmal** (Attribut): definiert durch Name und Wertebereich  
z.B.  $\text{Alter} \in \mathbb{N}$ ,  $\text{Geschlecht} \in \{m, w\}$

Merkmalsraum: kartesisches Produkt der Bereiche

**Objekt** (Individuum): charakterisiert durch Ausprägungen  
(Werte) der Merkmale (Merkmalsvektor)

z.B.  $\text{Alter} = 23$ ,  $\text{Geschlecht} = m$

Punkt (Vektor) im Merkmalsraum, Punktmenge bei  
unbekannten Merkmalsausprägungen

**Klasse** Menge von Objekten mit denselben (ähnlichen)  
Merkmalsausprägungen  
definiert durch

- ▶ Angabe der Merkmale und (Bereichen von)  
Ausprägungen  
z.B. volljährig:  $\text{Alter} \geq 18$ ,  $\text{Geschlecht} \in \{m, w\}$
- ▶ oder Referenzmodell (typischer Vertreter der  
Klasse) und Distanzmaß

Klassen müssen Zerlegung des Merkmalsraumes  
bilden

**Klassifizierung** (eindeutige) Zuordnung der Objektes zu Klassen

# Merkmale von Regionen im Bild

Region  $R \subseteq \text{pos}$  im Bild  $B : \text{pos} \rightarrow \{0, 1\}$   
(bzgl. einer Nachbarschaft zusammenhängend)

Merkmale (Features):

- ▶ topologische Merkmale
- ▶ quantitative Merkmale
- ▶ Formmerkmale
- ▶ geometrische Merkmale

ideal: Merkmale invariant gegenüber Verschiebung, Drehung, Skalierung, Verzerrung

# Topologische Merkmale

**Region:** Menge von Pixeln

**Kontur:** Menge der Randpixel  
(verschiedene Codierungen)

**Skelett:** topologieerhaltendes Bild aus  
Linien von 1 Pixel Dicke

- ▶ Zusammenhangseigenschaften
- ▶ Anzahl von Löchern (Euler-Zahl =  $|\text{Regionen}| - |\text{Löcher}|$ )
- ▶ Eigenschaften von (Ausschnitten des)  
Regionen-Adjazenz-Graphen
- ▶ Symmetrien
- ▶ Kontur der konvexen Hülle

Ziel: invariant gegenüber Verschiebung, Drehung, Skalierung,  
Verzerrung, Helligkeit, Beleuchtung  
z.B. für Zeichen, Gesten

# Quantitative Merkmale

Länge, Breite

Flächeninhalt  $A(R)$ : Anzahl der Positionen in der Region

Umfang der Kontur:  $U(R)$ : Länge, Anzahl der Randpositionen  
(evtl. verschieden gewichtet)

Durchmesser: größter (minimaler Umkreis), kleinster  
(maximaler Inkreis)

Masse: Summe der Intensitäten aller Pixel der Region

# Geometrische Merkmale

## Verhältnisse

**Seitenverhältnis:** des kleinsten umschließenden  
(achsenparallelen) Rechtecks (bounding box)

**Kompaktheit:** Verhältnis  $A(R) : U(R)$

**Kreisförmigkeit:**  $4\pi A(R) / U(R)^2$

**fill factor:** Verhältnis  $A(R)$  : Fläche der Bounding-Box

**Dichte:** Verhältnis  $A(R)$  : Fläche der konvexen Hülle der  
Region

**Konvexität:** Verhältnis  $U(R)$  : Umfang der konvexen Hülle der  
Region

## Symmetrien

**Projektionen** (z.B. Linescans)

# Statistische Merkmale

## Schwerpunkt

$$s_x(R) = \frac{1}{A(R)} \sum_{r \in R} r_x$$

$$s_y(R) = \frac{1}{A(R)} \sum_{r \in R} r_y$$

**Intensitäts-Schwerpunkt** von Regionen in Bildern  $B : \text{pos} \rightarrow \text{col}$

$$s'_x(R) = \frac{1}{\sum_{r \in R} B(r)} \sum_{r \in R} r_x B(r)$$

$$s'_y(R) = \frac{1}{\sum_{r \in R} B(r)} \sum_{r \in R} r_y B(r)$$

# Momente

(Region als 2d-Verteilung)

Momente der Region  $R$  im Bild  $B$  : pos  $\rightarrow$  col

$$m_{pq}(R) = \sum_{r \in R} B(r) r_x^p r_y^q$$

Momente der Region  $R \subseteq B$  im Binärbild  $B$  : pos  $\rightarrow$   $\{0, 1\}$

$$m_{pq}(R) = \sum_{r \in R} r_x^p r_y^q$$

Beispiele (für Binärbilder):

- ▶  $m_{00}(R) = \sum_{r \in R} r_x^0 r_y^0 = \sum_{r \in R} 1 = A(R)$
- ▶  $s_x(R) = \frac{1}{A(R)} \sum_{r \in R} r_x^1 r_y^0 = m_{10}(R),$   
 $s_y(R) = \frac{1}{A(R)} \sum_{r \in R} r_x^0 r_y^1 = m_{01}(R)$

## Zentrale Momente

Idee: Verschiebungs-Invarianz um Schwerpunkt

zentrale Momente der Region  $R \subseteq B$ :

$$\mu_{pq}(R) = \sum_{r \in R} B(r) r_{x-s_x(R)}^p r_{y-s_y(R)}^q$$

Verschiebungs-, aber nicht Größen-invariant

Größen-invariante zentrale Momente der Region  $R \subseteq B$ :

$$\mu'_{pq}(R) = \frac{\mu_{pq}(R)}{\mu_{00}(R)^{(p+q+2)/2}}$$

Daraus lassen sich weitere häufig verwendete Merkmale berechnen, z.B. Orientierung, Exzentrizität, Hu-Momente.

# Histograms of Oriented Gradients (HOG)

Idee:

- ▶ lokale Gradientenstruktur ist charakteristisches Merkmal für Regionen und Objekte
- ▶ invariant gegenüber affinen Transformationen und Beleuchtung

Berechnung der HOG (gewichtete)

1. evtl. Glättung
2. Zerlegung des Bildes (bzw. Bildausschnittes) in quadratische Teilbilder (Zellen)  $Z_1, \dots, Z_l$
3. Prewittfilter  $Z_h = (-1, 0, 1)Z$ ,  $Z_v = Z(-1, 0, 1)^T$
4. Bestimmung der in der Zelle vorkommenden Gradienten
  - ▶ lokaler Betrag  $b(p) = \sqrt{Z_h(p)^2 + Z_v(p)^2}$
  - ▶ lokale Richtung  $\theta(p) = \tan^{-1}\left(\frac{Z_v(p)}{Z_h(p)}\right)$  (diskretisiert)
5. Histogramme jeder Zelle  $H_i : \theta \rightarrow \mathbb{R}_{\geq 0}$ ,  $\theta \mapsto |\{p \in Z \mid \theta(p) = \theta\}|$  (evtl. gewichtet durch  $b(p)$  und Abstand zum Zellmittelpunkt)
6. Normierung der Histogramme über Blöcken von Zellen (erntfernt Abhängigkeit von Beleuchtung)

Merkmalsvektor des Bildes (Bildausschnittes):

Verkettung der Histogramme aller Blöcke

# Beispiel Verkehrszeichen-Erkennung

**Farb-basiert** statistische Werte verschiedener Farbkanäle, Farb-Verhältnisse und -Differenzen  
Binarisierung und Regionen-Erkennung in verschiedenen Farbkanälen

- + Störungen, verdeckte Schilder  
Extraktion relevanter Bildbereiche
- abhängig von Beleuchtung, Umgebung, Alter und Farbzustand der Schilder

**Form-basiert** Anwendung Klassifikatoren für die fünf typischen Schilderformen auf Bildbereiche verschiedener Größe und Position

- + unabhängig von Beleuchtung, Alter und Farbzustand der Schilder
- Verzerrung, Störungen, verdeckte, gekippte Schilder

Kombination beider Ansätze

ggf. Normierung der farbbasiert gefundenen Bereiche

# Modellbasierte Merkmalsextraktion

Idee: Verwendung von Kontextinformation über zusammenhängende Regionen

Repräsentation von Kandidaten für den Bildinhalt (Modelle), z.B.

- ▶ Mensch aus (bevorzugt) zusammenhängenden Regionen  
Kopf, Rumpf, Arme, Beine
- ▶ Haus mit Fenstern, Dach, Türen
- ▶ Landschaft mit Strand, Meer, Himmel
- ▶ Achtelnote mit Kopf, Hals, Fähnchen

Modell:

- ▶ Menge von Regionen mit jeweils typischen Eigenschaften (Klassen)
- ▶ Beziehungen (Zusammenhang) zwischen den Regionen

Modell = Graph

Klassifikation von (Regionen in) Bildern =

Vergleich des Modellgraphen mit (Teilgraphen des)

Regionen-Adjazenz-Graphen des Bildes

# Klassifikation von Regionen

Ziel: Zuordnung von (Regionen in) Bildern zu Klassen  
(Bedeutung)

Vorbereitung:

- ▶ Segmentierung des Bildes
- ▶ Extraktion der (signifikanten) Merkmale der Bildregionen
- ▶ evtl. Normierung der Merkmale

Klassifikation bei Klassen-Definition durch

- ▶ Merkmalsausprägungen:  
entsprechend den extrahierten Bild-Merkmalen
- ▶ Referenzmodelle:  
Zuordnung zur Klasse mit dem ähnlichsten Referenzmodell  
(geringster Abstand im gegebenen Distanzmaß)

Klassifikations-Methoden (vorgestellt in anderen Modulen):  
statistisch, Mustererkennung, Clustering, maschinelles Lernen,  
neuronale Netze, Support Vector Machines

# Bildfolgen

Bild  $B : \text{pos} \rightarrow \text{col}$

Folge von Bildern  $B_0, B_1, B_2, \dots$

$$B : \mathbb{N} \rightarrow (\text{pos} \rightarrow \text{col})$$

z.B. Videosequenz

# Operationen auf Bildfolgen

- ▶ Mittelung mehrerer Bilder

$$\forall p \in \text{pos} : B(p) = \frac{1}{n} \sum_{i=1}^n B_i(p)$$

zum Entrauschen, z.B. in Medizintechnik

- ▶ Differenz zwischen aufeinanderfolgenden Bildern

$$\forall p \in \text{pos} : B(p) = |B_{i+1}(p) - B_i(p)|$$

zur Hervorhebung / Erfassung zeitlicher Veränderungen

- ▶ Detektion von Bewegung (z.B. zur Überwachung)
- ▶ Detektion von Intensitätsänderung (z.B. Wirkung von Kontrastmitteln)
- ▶ Videokompression (MPEG)
- ▶ Feststellung korrespondierender Punkte in mehreren Bildern
- ▶ Mosaicing: Zusammensetzen mehrerer Teilbilder zu einem Gesamtbild (z.B. Panoramabild, Karte aus Satellitenaufnahmen)

# Korrespondenzen in Bildern

Ziel: Bestimmung von Punkten mit derselben Bedeutung in mehreren Bildern

Anwendung z.B. bei

- ▶ Kombination der Informationen mehrerer mit verschiedenen Kameras aufgenommener Bildkanäle
- ▶ Stereobildverarbeitung
- ▶ Suche in Bilddatenbanken
- ▶ Zusammensetzen von Bildern aus überlappenden Teilbildern

z.B. durch descriptives Matching

# Hervorhebung von Bewegungen

Annahme:

- ▶ statischer Hintergrund
- ▶ bewegte Objekte
  
- ▶ Mittelung mehrerer aufeinanderfolgender Bilder:  
bewegtes Objekt „verschwimmt“ mit Hintergrund
- ▶ (Betrag der) Differenz aufeinanderfolgender Bilder ist bei  
bewegten Objekten hoch,  
gering beim statischen Hintergrund

# Modul Digitale Bildverarbeitung SS19

Bestandteile der Lehrveranstaltung und Prüfung:

- ▶ Vorlesungen
- ▶ Übungsserien
- ▶ Praktika (ImageJ)
- ▶ Projekt

Prüfungsvorleistungen:

- ▶ ImageJ-Praktikum (4 Serien)
- ▶ Projektbearbeitung, -dokumentation, -präsentation

(alles in OPAL eintragen)

Prüfung Klausur (120 min)

am Montag, 22. Juli 2019 um 13:00 - 15:00 Uhr in LNW006

ausschließlich zugelassenes Hilfsmittel:

ein handgeschriebenes A4-Blatt

# Vorlesung – Themen

- ▶ Digitale Bilder, Eigenschaften
- ▶ statistische Merkmale
- ▶ Punktoperationen
- ▶ geometrische Operationen
- ▶ digitale Signale (1d, 2d)
- ▶ Fourier-Transformation, -Analyse
- ▶ Faltung, Filter
- ▶ morphologische Operationen
- ▶ Merkmale von Bildregionen
- ▶ Segmentierung

# Digitale Bilder

digitales Bild:

Zuordnung  $B : \text{pos} \rightarrow \text{col}$

von Farben (aus der Menge col)

zu Positionen (aus der Menge pos)

**pos** Menge der Positionen

**col** Menge der Farben

- ▶ Binärbild:  $\text{col} = \{0, 1\}$  (  $\{0, 255\}$  )
- ▶ Grauwertbild (Intensitäten):  
 $\text{col} = \{0, \dots, m\}$  (real  $[0, 1] \subset \mathbb{R}$ )
- ▶ Farbbild mit mehreren Farbkanälen  
( $\text{col}_1, \dots, \text{col}_n$ ):  $\text{col} = \text{col}_1 \times \dots \times \text{col}_n$

für rechteckige Positionsmengen  $\text{pos} = [x_{\min}, x_{\max}] \times [y_{\min}, \dots, y_{\max}]$

mit  $[x_{\min}, x_{\max}] \subset \mathbb{N}$  und  $[y_{\min}, y_{\max}] \subset \mathbb{N}$

(real  $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \subseteq \mathbb{R}^2$ )

Bild = Matrix  $B \in \text{col}^{\text{pos}}$

z.B. für  $\text{pos} = \{0, \dots, m-1\} \times \{0, \dots, n-1\}$  und  $\text{col} = \mathbb{N}$ :

Bild  $B \in \mathbb{N}^{\{0, \dots, m-1\} \times \{0, \dots, n-1\}} = \mathbb{N}^{m \times n}$

# Statistische Merkmale

(für Grauwert-Bilder)

- ▶ Auflösungen:
  - ▶ Ortsauflösung
  - ▶ Intensitätsauflösung
- ▶ Extrema  
(minimale, maximale Intensität)
- ▶ Mittelwerte  
(arithmetisches Mittel, Median)
- ▶ mittlere quadratische Abweichung
- ▶ Intensitätsprofile (z.B. Linienprofil)
- ▶ Histogramm (ausgeglichen, bimodal)  
kumulatives Histogramm
- ▶ Co-occurrence-Matrix
- ▶ Entropie

# Punktoperationen

Farbtransformation  $col \rightarrow col'$

Anwendung auf jede Position im Bild

(ohne Zugriff auf andere Positionen im Bild)

Operationen auf einem Bild:

- ▶ Zerlegung / Kombination der Farbkanäle
- ▶ Binarisierung (mit Schwellwert)
- ▶ Invertierung
- ▶ Graustufen-Zuordnung
- ▶ Histogrammspreizung

Operationen auf mehreren Bildern:

- ▶ logische Operationen auf Binärbildern ( $\vee, \wedge, \text{XOR}, \dots$ )  
entsprechen Mengenoperationen auf dem Vordergrund  
(weiß)
- ▶ arithmetische Operationen auf Grauwertbildern ( $+, -, \cdot, \dots$ )
- ▶ Überlagerung

# Geometrische Operationen

- ▶ homogene Koordinaten
- ▶ lineare Transformationen
  - ▶ Verschiebung
  - ▶ Drehung
  - ▶ Skalierung
- ▶ Zusammensetzung durch Matrixmultiplikation
- ▶ nichtlineare Transformationen
  - ▶ perspektivische Ver- und Entzerrung
- ▶ Invertierung der Operationen
- ▶ Interpolations-Funktionen
  - ▶ nächster Bildpunkt
  - ▶ Mittelwerte der nächsten Bildpunkte
  - ▶ lineare Interpolation
  - ▶ bilineare, kubische Interpolation

# Fourier-Analyse

- ▶ Signale (analog / digital, 1d / 2d)
  - ▶ Abtastung (Dirac-Impuls, Dirac-Kamm)
  - ▶ Nyquist-Shannon-Abtasttheorem
  - ▶ Alias-Effekte
  - ▶ Faltung von Funktionen (Signalen)
- 
- ▶ Fourier-Transformation
  - ▶ Fourier-Spektrum (Frequenzraum)
  - ▶ Bildeffekte durch Transformationen im Frequenzraum
  - ▶ Faltung im Ortsraum entspricht Multiplikation im Frequenzraum

# Filter

- ▶ Tiefpass-, Hochpass-, Bandpassfilter
- ▶ Ersetzung der Operationenfolge:
  1. Fourier-Transformation
  2. Transformation im Frequenzraum
  3. inverse Fourier-Transformationdurch Faltung mit Matrix (Filterkern)
- ▶ lineare Filter:  
Glättung, Kanten-Hervorhebung
- ▶ nichtlineare Filter:  
lokale Extrema, Median
- ▶ Spezialfälle zum Filtern am Bildrand:  
Rand nicht filtern, spezielle Filtermatrizen,  
schwarz, konstant, gespiegelt, periodisch

# Morphologische Operationen

- ▶ Strukturelemente (Masken)
- ▶ Erosion
- ▶ Dilatation
- ▶ Opening: erst Erosion, dann Dilatation
- ▶ Closing: erst Dilatation, dann Erosion
- ▶ Hit-or-Miss-Transformation,  
Hit-or-Miss-Masken
- ▶ Konturbestimmung durch morphologische Operationen
- ▶ Distanztransformation
- ▶ Skelettierung, Thinning
- ▶ Thickening, approximierte konvexe Hülle

# Regionen in Bildern

Merkmale von Regionen:

- ▶ Farbwert(-Bereich)
- ▶ Textur: typische Farbwert-Verteilung
- ▶ quantitativ: Ausdehnungen, Volumen, Umfang
- ▶ Formmerkmale: Verhältnisse, Momente, fill factor
- ▶ topologische Merkmale:  
Zusammenhang, Skelett, Kontur

# Segmentierung

Multiskalen-Darstellung (Pyramiden)

Bestimmung von Regionen:

- ▶ Homogenitätskriterien
- ▶ Region Growing
- ▶ Region Labeling
- ▶ Region Merging, Regionen-Adjazenz-Graph
- ▶ Split and Merge (Quad-Trees)

Modell-basiert:

- ▶ Hough-Transformation für Geraden, Kreise, ...

Kontur-basiert:

- ▶ Darstellung durch Kettencodes
- ▶ Edge-Linking
- ▶ Canny-Edge-Detector