

# Zusammenfassung Digitale Bildverarbeitung

## 1. Einführung

### Geschichtliches:

Ab den 1920ern: Bartlane Cable Picture Transmission mit ca. 5 Graustufen, Übertragung eines Bildes von New York nach London in ausreichender Druckqualität in 3 Stunden

(Erst) ab den 1960ern: Wissenschaftliche Disziplin durch Weltraumprogramm

Anwendungen: Medizin (Röntgenbilder), Fotografie (Störungen entfernen), Vermessung der Erde, etc.

### Schritte bei der Bildverarbeitung:

Image Acquisition Bilderfassung mit Sensor. Aktiv (Messung der Reflektionen ausgestrahlter Wellen) oder passiv (ausschließlich Messung von Wellen von Objekten des Sichtbereichs)

Image Enhancement

Image Restoration

Color Image Processing

Wavelets and Multiresolution Processing

Compression

Morphological Processing

Segmentation Einteilen des Bildes in überlappende oder nicht-überlappende Teile

Representation & Description Segmente und Bilder werden weiter manipuliert, um den Inhalt der Bilder so zu charakterisieren, dass Erkennung und Interpretation unterstützt werden

Object Recognition Der eigentlich *Seh*-Teil der Verarbeitung. Herstellung semantischer Informationen.

Wellenlängen des Lichtes:

400 - 420 nm	Violett	575 - 585 nm	Gelb
420 - 490 nm	Blau	585 - 650 nm	Orange
490 - 575 nm	Grün	650 - 750 nm	Rot

**Farbmodelle:** RGB, CMY(K), HIS (Hue-Intensity-Saturation, als Zylinder), HLS (Hue-Luminance-Saturation, als Doppelkegel), HSV (Hue-Saturation-Value, als Kegel)

## Menschlicher

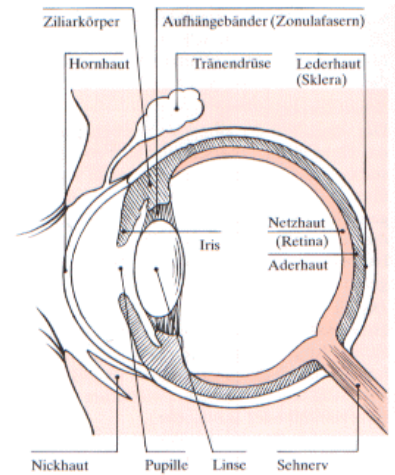
### Sehsinn:

Zäpfchen: Teil des menschlichen Auges zur Wahrnehmung von Farbe

Stäbchen: Wahrnehmung von Helligkeit

### Mach-Band-Effekt:

Sinnesphänomen; an Grenze zwischen heller und dunklen Fläche erscheint auf heller Seite ein schmales Band, welches noch heller ist. Entsprechend auf der dunklen Seite.



## 2. Rasterung und Quantisierung

Darstellung eines Bildes als 2-stellige Funktion  $f : M \times N \rightarrow ]0, \infty[$ , wobei  $M := \{1, \dots, m\}$ ,  $N$  resp. Aufteilung in Beleuchtungs- und Reflektionskomponenten  $I : M \times N \rightarrow ]0, \infty[$ ,  $J : M \times N \rightarrow ]0, 1[$ . Es gelte dabei  $f = I \cdot J$ .

$I$  gibt den Grauwert eines Pixels an.

Aliasing: Allgemeine Definition: Alias-Effekte sind durch die Verletzung des Abtasttheorems (siehe nachfolgenden Absatz) auftretende Fehler beim Abtasten von Signalen.

Auswirkungen bei der Bildverarbeitung: Moiré-Muster, Treppeneffekte an harten Kanten, Verschwinden von Strukturen, die dünner als 1 Pixel sind, etc.

(Nyquist-Shannon'sches) Abtasttheorem: Ein kontinuierliches Signal mit einer Maximalfrequenz  $f_{\max}$  muss mit einer Frequenz größer als  $2 \cdot f_{\max}$  abgetastet werden, damit man aus dem so erhaltenen zeitdiskreten Signal das Ursprungssignal ohne Informationsverlust wieder rekonstruieren kann.

Anti-Aliasing: Entweder feinere Auflösung bei Rasterung wählen: Jede Struktur durch mindestens 2 Bildpunkte repräsentieren oder durch Verwischung das Auge täuschen

Ortsbereich: Bildpunkte  $\leftrightarrow$  Frequenzbereich: enthaltenen Frequenzen (Umwandlung mittels Spektralanalyse)

### 3. Fourier Transformation

Transformationen werden oftmals dazu eingesetzt, Probleme zu vereinfachen – indem etwa die Komplexität der zur Problemlösung verwendeten Algorithmen gesenkt wird. Beispiel: Multiplikation/Division mittels Logarithmentafeln. Transformation des Problems  $f(a, b) = a \cdot b$  auf eine Addition (+ Hin- und Rücktransformation):  $\exp(\ln(a) + \ln(b))$ .

Aus der Analysis ist bekannt, dass eine über  $]-\infty, \infty[$  stückweise stetige, stückweise monotone sowie  $T$ -periodische Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  als Summe von Sinus und Cosinus-Funktionen dargestellt werden kann:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi}{T} kx\right) + b_n \sin\left(\frac{2\pi}{T} kx\right) \quad (1)$$

Durch Umformen und Ausnutzen von Additionstheoremen erhält man, dass sich  $f$  auch in komplexer Form schreiben lässt:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k \exp\left(i \frac{2\pi}{T} kx\right), \text{ mit} \quad (2)$$

$$c_k = \frac{1}{T} \int_{-T/2}^{T/2} f(t) \exp\left(-i \frac{2\pi}{T} kt\right) dt \in \mathbb{C} \quad (3)$$

Definiere nun die sogenannte Fourier-Transformation (FT)  $\mathcal{F}\{f(x)\}$ , sofern sie für ganz  $\mathbb{R}$  existiert. Im Folgenden setzen wir immer  $\lim_{x \rightarrow \pm\infty} = 0$  voraus, da dies hinreichend für die Existenz ist – allerdings nicht unbedingt notwendig).

$$F(u) := \int_{-\infty}^{\infty} f(t) \exp(-i2\pi ut) dt \quad (4)$$

Zur Verallgemeinerung auf den Fall, dass  $f$  nicht-periodisch ist, gehe  $T \rightarrow \infty$ . Damit folgt:

$$T \cdot c_k = F\left(\frac{k}{T}\right) \quad (5)$$

Erweitere nun (2) mit  $T/T$ :

$$f(x) = \sum_{k=-\infty}^{\infty} T c_k \exp\left(i \frac{2\pi}{T} kx\right) \frac{1}{T} \quad (6)$$

Da  $\frac{1}{T}$  für  $T \rightarrow \infty$  infinitesimal klein wird, ersetze im Folgenden:

$$\frac{1}{T} \rightarrow du, \quad \frac{k}{T} \rightarrow u, \quad F\left(\frac{k}{T}\right) \rightarrow F(u) \quad (7)$$

Erhalte durch  $T \rightarrow \infty$ :

$$f(x) = \int_{-\infty}^{\infty} F(u) \exp(i2\pi ux) du \quad (8)$$

Folglich lässt sich jede beliebige über  $]-\infty, \infty[$  stückweise stetige, stückweise monotone Funktion  $f$ , deren FT wie oben beschrieben existiert, als unendliche Summe bzgl. der Basis  $B := \{\exp(-i2\pi ux) \mid u \in \mathbb{R}\}$  darstellen. Dass  $B$  tatsächlich Basis ist, folgt aus der Orthogonalität zweier Elemente  $\exp(-i2\pi u_1 x)$  und  $\exp(-i2\pi u_2 x)$  für  $u_1 \neq u_2$  bzgl. des Skalarprodukts  $\int fg$  auf der Menge aller komplexwertigen Funktionen, d. h.  $f, g : \mathbb{R} \rightarrow \mathbb{C}$ .

**Bemerkungen:** **i)** Die Def. von  $F$  in (4) ist also nichts anderes als das Skalarprodukt zwischen der Funktion  $f$  und dem entsprechenden Basisvektor  $u \rightarrow \exp(-i2\pi ut)$ . (Vgl. den Anschauungsraum, in dem wir beim Basiswechsel für die Koordinatenbestimmung ebenfalls das Skalarprodukt eines Vektors mit den neuen Basisvektoren bilden würden.) **ii)** Das  $u$  in der Definition von  $F$  wird Frequenzvariable genannt.

Da  $F$  im Allgemeinen eine komplexe Funktion ist, teilt man auf:  $F(u) = R(u) + iI(u)$

Als Polarform von  $F$  ergibt sich:

$$F(u) = |F(u)| \exp(i \cdot \Phi(u)), \text{ mit} \quad (9)$$

$$|F| = \sqrt{R^2 + I^2}, \quad \Phi = \tan^{-1} \frac{I}{R}$$

$|F(u)|$  heißt Fourier-Spektrum,  $|F(u)|^2$  Leistungsspektrum von  $f$  und  $\Phi(u)$  der Phasenwinkel.

**Eigenschaften der Fourier-Transformation:**

- Linearität:

$$\mathcal{F}\{af(x) + bg(x)\} = aF(u) + bG(u) \quad (10)$$

- Änderung der Skalierung im Ortsbereich:

$$\mathcal{F}\{f(ax)\} = \frac{1}{a} F\left(\frac{u}{a}\right) \quad (11)$$

Das Copyright für die dieser Zusammenfassung zugrunde liegenden Vorlesungsunterlagen (Skripte, Folien, etc.) liegt beim Dozenten. Darüber hinaus bin ich, Florian Schoppmann, alleiniger Autor dieses Dokuments und der genannte Dozent ist in keiner Weise verantwortlich. Etwaige Inkorrektheiten sind mit sehr großer Wahrscheinlichkeit erst durch meine Zusammenfassung/Interpretation entstanden.

- Verschiebung im Ortsbereich ändert nur die Phase:

$$\mathcal{F}\{f(x - a)\} = F(u) \exp(i2\pi ua) \quad (12)$$

- Ableitung:

$$\mathcal{F}\left\{\frac{d^n}{dx^n} f(x)\right\} = (i2\pi u)^n F(u) \quad (13)$$

- Integral des Leistungsspektrums:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |F(u)|^2 du \quad (14)$$

- Symmetrieeigenschaften:

$$\begin{aligned} f \text{ reell} &\implies F(-x) = \overline{F(x)} \\ f \text{ imaginär} &\implies F(-x) = -\overline{F(x)} \\ f \text{ gerade, d. h. } f(-x) &= f(x) \\ &\implies F(-x) = F(x) \\ f \text{ ungerade, d. h. } f(-x) &= -f(x) \\ &\implies F(-x) = -F(x) \end{aligned}$$

Für den zweidimensionalen Fall:

$$\mathcal{F}\{f(u, v)\} = F(u, v) \quad (15)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s, t) \exp(-i2\pi(us + vt)) ds dt \quad (16)$$

$$= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(s, t) \exp(-i2\pi us) ds \right] \exp(-i2\pi vt) dt \quad (17)$$

(Separabilität von  $F$ )

gelten die gleichen Eigenschaften wie im eindimensionalen. Zusätzlich: Wird  $f(x, y)$  um den Winkel  $\theta$  gedreht, so auch  $\mathcal{F}\{f(x, y)\} = F(u, v)$ .

Für die praktische Anwendbarkeit benötigen wir die diskrete Fourier-Transformation (DFT). Eine kontinuierliche Funktion  $f(x)$  mit FT  $F(u)$  werde dazu diskretisiert, d. h. auf  $N$  äquidistanten Intervallen  $[0, \Delta_x] = [x_0, x_1], \dots, [x_{N-1}, x_N]$  approximiert durch die Treppenfunktion  $g(x)$  mit FT  $G(u)$ . (OBdA setzen wir hier voraus, dass  $x_0 = 0$ , ggf. wäre ansonsten noch eine Verschiebung notwendig.) Der Kehrwert von  $\Delta_x$  ist dabei die (Abtast-)frequenz.

$$g(x) := \begin{cases} f(x_i) & \text{wenn } x \in [x_i, x_{i+1}[ \\ 0 & \text{sonst} \end{cases} \quad (18)$$

Nach Nyquist'schem Abtasttheorem ist die höchste darstellbare Frequenz bei einer Abtastrate von

$\frac{1}{\Delta_x}$  (also einem Abtastintervall von  $\Delta_x$ ) gerade  $\nu_{\Delta} := \frac{1}{2\Delta_x}$ .

Folglich betrachten wir nur Frequenzen innerhalb des Intervalls  $[-\nu_{\Delta_x}, +\nu_{\Delta_x}]$  mit äquidistanter Zerlegung (also mit Abstand  $\Delta_u = \frac{2\nu_{\Delta_x}}{N} = \frac{1}{N\Delta_x}$ ):  $u_0 < \dots < u_N$ .

Es kann nun approximiert werden:

$$\begin{aligned} F(u_n) &= \int_{-\infty}^{\infty} f(t) \exp(-i2\pi u_n t) dt \\ &\approx G(u_n) = \Delta_x \sum_{k=0}^{N-1} g(k\Delta_x) \exp(-i2\pi u_n k\Delta_x) \\ &= \Delta_x \sum_{k=0}^{N-1} g(k\Delta_x) \exp(-i2\pi k(\frac{n}{N} - \frac{1}{2})) \\ &\quad \text{da } u_n \Delta_x = (-\frac{1}{2\Delta_x} + \frac{n}{N\Delta_x}) \Delta_x \end{aligned} \quad (19)$$

Da die Abbildung  $n \rightarrow G(u_n)$  (bei entsprechender Fortsetzung der  $u_n$ )  $N$ -periodisch ist, wird aus praktischen Gründen die DFT  $H$  von  $g(x)$  jedoch zwischen 0 und  $\nu_{\Delta_x}$  als  $n$ -elementige Folge definiert. Zudem wird  $\Delta_x = 1$  gewählt. (Bei einer beliebigen Funktion würden wir die DFT auf einer Skalierung vornehmen, welche die Abtastpunkte auf einen Abstand von 1 abbildet.) Zusammen mit etwas geänderter Skalierung ( $\frac{1}{N}$  steht nun bei der Hintransformation) und  $g_n := g(n\Delta_x) = g(n)$  folgt:

$$G_n := \frac{1}{N} \sum_{k=0}^{N-1} g_k \exp(-i2\pi k \frac{n}{N}) \quad (20)$$

$$g_n = \sum_{k=0}^{N-1} G_k \exp(i2\pi k \frac{n}{N}) \quad (21)$$

**Bemerkungen:** Einsetzen von (21) in (20) ergibt

$$\begin{aligned} G_n &= \frac{1}{N} \sum_{k=0}^{N-1} \left[ \sum_{l=0}^{N-1} G_l \exp(i2\pi l \frac{k}{N}) \exp(-i2\pi k \frac{n}{N}) \right] \\ &= \frac{1}{N} \sum_{l=0}^{N-1} G_l \underbrace{\sum_{k=0}^{N-1} \exp(i2\pi(l-n) \frac{k}{N})}_{\substack{N \text{ wenn } l=n, \\ \text{sonst } 0 \text{ wg. (23)}}} = G_n \end{aligned} \quad (22)$$

Man sieht also, dass das Produkt der konstanten Vorfaktoren  $\frac{1}{N}$  ergeben muss. Die Aufteilung ist prinzipiell willkürlich möglich.

Sei  $w := \exp(i2\pi(l-n)\frac{1}{N})$ . Es gilt:  $w$  ist  $N$ -te Einheitswurzel (Gruppeneigenschaft), und  $\exp(i2\pi(l-n)\frac{k}{N}) = w^k$ , da  $k \in \mathbb{N}$ . Folglich:

$$\sum_{k=0}^{N-1} w^k = \frac{w^N - 1}{w - 1} = 0 \quad (23)$$

Die Erweiterung auf den 2-dimensionalen Fall erfolgt analog zur kontinuierlichen FT. Es gelten die für die kontinuierliche FT oben genannten Eigenschaften.

Durchschnitt  $\bar{f}$  einer 2-dimensionalen diskreten Funktion  $f(x, y)$ :

$$\begin{aligned} \bar{f}(x, y) &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \\ &= F(0, 0) \quad (\text{durch Einsetzen}) \end{aligned} \quad (24)$$

Die Faltung zweier Funktionen  $f, g$  ist definiert im Eindimensionalen (25) und 2-Dimensionalen (26):

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt \quad (25)$$

$$(f * g)(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s, t)(x-s, y-t)dsdt \quad (26)$$

Faltungstheorem (Convolution Theorem):

$$\mathcal{F}\{(f * g)(x)\} = F(u) \cdot G(u) \quad (27)$$

$$\mathcal{F}\{f(x) \cdot g(x)\} = (F * G)(u) \quad (28)$$

Faltung im diskreten Fall,  $f, g : \mathbb{Z} \rightarrow \mathbb{R}$ , beide  $N$ -periodisch (2-dimensionaler Fall analog):

$$(f * g)(x) = \sum_{i=0}^{N-1} f(i)g(x-i) \quad (29)$$

Der Dirac-Impuls  $\delta(x-a)$  eine Distribution, die über ihr Integral definiert ist. Für eine beliebige Funktion  $f(x)$  muss gelten:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a) \quad (30)$$

$$\int_{-\infty}^{\infty} \delta(x-a)dx = \lim_{t \rightarrow 0} \int_{a-t}^{a+t} \delta(x-a)dx = 1 \quad (31)$$

Im 2-Dimensionalen analog.

Die Fast Fourier-Transformation (FFT). Im Folgenden sei  $W_N = \exp(-i2\pi/N)$  primitive  $N$ -te Einheitswurzel.  $N$  sei der Form  $N = 2^m = 2M$ ,  $m, M \in \mathbb{N}$ . Nach (20) folgt ( $(f_n)$  sei Folge von Funktionswerten im Orts- ( $F_n$ ) im Frequenzbereich):

$$\begin{aligned} F_n &= \frac{1}{2M} \sum_{k=0}^{2M-1} f_k W_{2M}^{nk} \\ &= \frac{1}{2} \left[ \frac{1}{M} \sum_{k=0}^{M-1} f_{2k} W_{2M}^{n(2k)} \right. \\ &\quad \left. + \frac{1}{M} \sum_{k=0}^{M-1} f_{2k+1} W_{2M}^{n(2k+1)} \right] \end{aligned} \quad (32)$$

$$\begin{aligned} &= \frac{1}{2} \left[ \underbrace{\frac{1}{M} \sum_{k=0}^{M-1} f_{2k} W_M^{nk}}_{=: F_n^{(\text{even})}} \right. \\ &\quad \left. + \frac{1}{M} \sum_{k=0}^{M-1} f_{2k+1} W_M^{nk} W_{2M}^n \right] \end{aligned} \quad (33)$$

$$\begin{aligned} & \quad (\text{da } W_{2M}^{2k} = W_M^{nk}) \\ &= \frac{1}{2} (F_n^{(\text{even})} + F_n^{(\text{odd})} W_{2M}^n) \end{aligned} \quad (34)$$

Das Divide & Conquer-Prinzip werde entsprechend fortgeführt... Laufzeitverbesserung von  $N^2$  (mit „normalem“ naivem Ansatz) auf  $N \log_2 N$ .

#### 4. Punktoperationen

Punktoperationen verändern die Pixel unabhängig von den Nachbarpixeln: Operation  $T$  auf einen Pixel an Position  $(x, y)$  und altem Grauwert  $r$ . Neuer Grauwert  $s = T(g, x, y)$ , falls nur von Grauwert abhängig (oft), dann  $s = T(r)$ .

(Nachbarschaftsoperationen – siehe nächstes Kapitel – berücksichtigen umliegende Pixel.)

Beispiele für Punktoperationen:

- Negativ
- Exponential-Transformation („Gamma Korrektur“):  
 $T(r) = cr^\gamma$  oder  $T(r) = c(r + \epsilon)^\gamma$
- Operationen über mehrere Bilder: Bspw. Durchschnittbildung

**Maße für Qualität eines Bildes:**

Mittlerer Grauwert

$$\Gamma := \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \quad (35)$$

Varianz Mittlere quadratische Abweichung:

$$\Lambda := \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x, y) - \Gamma)^2 \quad (36)$$

Histogramm Bildet einen Grauwert  $g$  auf die Anzahl Pixel mit diesem Grauwert  $r$  ab:  $r \rightarrow |\{(x, y) \mid f(x, y) = r\}|$ . Wenn normiert (also durch  $MN$  geteilt), entspricht das Histogramm einer Wahrscheinlichkeitsverteilung über den Grauwert eines zufällig gewählten Pixels.

Grauwertprofil Grauwerte entlang einer Linie, d. h. Zuordnung Position  $\rightarrow$  Grauwert

Histogramm Einebnung: Betrachte eine stetige (streng) monoton steigende (und damit bijektive) Punktoperation  $s = T(r)$ , deren Inverse  $T^{-1}$  ebenfalls monoton steigt. Zur Einfachheit sei  $T : [0, 1] \rightarrow [0, 1]$ . Mit  $p_r, p_s$  seien im Folgenden die Wahrscheinlichkeitsdichten bezeichnet für die Zufallsvariable  $R, S = T(R)$ , die den Grauwert eines zufällig ausgewählten Pixels vor bzw. nach der Transformation  $T$  angeben. Es folgt:

$$\begin{aligned} \Pr(T(a) \leq S < T(b)) &= \int_a^b p_r(r) dr \\ &= \int_{T(a)}^{T(b)} p_r(T^{-1}(s))(T^{-1})'(s) ds \quad (\text{Subst.}) \\ \implies p_s(s) &= \frac{d}{ds} \Pr(S < s) \quad (\text{Def. der Dichte}) \\ &= p_r(T^{-1}(s))(T^{-1})'(s) = p_r(r) \frac{dr}{ds} \end{aligned} \quad (37)$$

Für die Einebnung gewünscht ist  $p_s(s) = 1$  für  $0 \leq s \leq 1$  (und 0 sonst). Definiere:

$$s = T(r) := \int_0^r p_r(t) dt \quad (0 \leq r \leq 1) \quad (38)$$

Hiermit folgt:

$$p_r(r) = \frac{ds}{dr} = \left(\frac{dr}{ds}\right)^{-1} \quad (39)$$

$$p_s(s) = p_r(r) \frac{1}{p_r(r)} = 1 \quad (40)$$

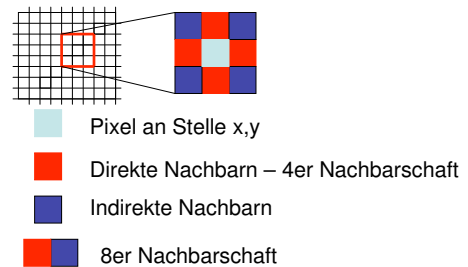
Die Einebnung im diskreten Fall (es stehen  $k$  Graustufen  $r_0, \dots, r_k$  zur Verfügung):  $n_j$  bezeichne die Anzahl der Pixel mit Grauwert  $r_j$ ,  $N$  die Gesamt-pixelzahl. Es gelten (analog zu (38)) folgende Definitionen:

$$p_r(r_j) := \frac{n_j}{N} \quad (41)$$

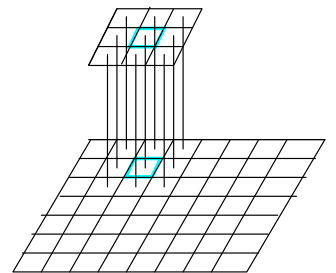
$$s_k = T(r_k) := \sum_{j=0}^k p_r(r_j) \quad (42)$$

**5. Filteroperationen**

Darstellung Nachbarschaft:



Grafische Darstellung einer Faltung: Multiplizieren, Aufsummieren, Zentrumspixel ersetzen



Einige Filter:

Boxcar:  $\frac{1}{9} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Gauss:  $\frac{1}{8} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

Median: Medianwert in Nachbarschaft nehmen (ohne Entsprechung im Frequenzbereich!)

Wallis' Operator: Bezeichne  $\sigma_{ij}$  die lokale Varianz des Pixels an Pos.  $(i, j)$ ,  $\sigma_d$  die gewünschte Varianz,  $\mu_{ij}$  das lokale Mittel,  $\mu_d$  das gewünschte Mittel:

$$s = T(r) = \mu_d + \frac{\sigma_d}{\sigma_{ij}}(r - \mu_{ij}) \quad (43)$$

Lee Filter: Einfachere Version des Operators für selbst zu bestimmende Konstante  $k$ :

$$s = T(r) = \mu_d + k(r - \mu_{ij}) \quad (44)$$

Kantenerkennung:

Roberts:  $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$  oder  $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

Das Copyright für die dieser Zusammenfassung zugrunde liegenden Vorlesungsunterlagen (Skripte, Folien, etc.) liegt beim Dozenten. Darüber hinaus bin ich, Florian Schoppmann, alleiniger Autor dieses Dokuments und der genannte Dozent ist in keiner Weise verantwortlich. Etwas Inkorrektheiten sind mit sehr großer Wahrscheinlichkeit erst durch meine Zusammenfassung/Interpretation entstanden.

Sobel: 

-1	-2	-1
0	0	0
1	2	1

Laplace: 

0	1	0
1	-4	1
0	1	0

 oder 

1	1	1
1	-8	1
1	1	1

 (mit Diag.)

Erste partielle Ableitungen eines Bildes  $f(x, y)$  (Gradient):

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} f(x+1, y) - f(x, y) \\ f(x, y+1) - f(x, y) \end{pmatrix} \quad (45)$$

Der Gradient zeigt in die Richtung des maximalen Anstiegs von  $f$ .

Zweite partielle Ableitungen ( $\frac{\partial^2 f}{\partial y^2}$  analog):

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= [f(x+1, y) - f(x, y)] - \\ &\quad [f(x, y) - f(x-1, y)] \\ &= f(x+1, y) - 2f(x, y) + f(x-1, y) \end{aligned} \quad (46)$$

Dies entspricht dem Laplace-Filter.

Laplace mit 

1	1	1
1	-8	1
1	1	1

 entspricht ebenfalls einem

Unschärf Maskieren, d. h. Kantenverstärktest Bild =  $c \cdot$  Originalbild - geglättetes Bild.

Der Filterkern summiert sich auf 1 bei einem Glättungsfilter, um die Skalierung der Grauwerte zu behalten, und summiert sich auf 0 bei einer Kantenerkennung, so dass bei Filterung über ein homogenes Bild das Ergebnis 0 bleibt (keine Kanten).

Ergebnisse: 1. Ableitung erzeugt dickere Kanten, 2. Ableitung dünnere. Insb. die 2. Ableitung verstärkt Rauschen stark durch „ringing“, „double edges“

Ein Tiefpassfilter lässt nur „niedrige“ Frequenzen durch, um so Störungen mit hohen Frequenzen zu eliminieren. Dies ist intuitiv im Frequenzbereich möglich: Sei  $F(u, v)$  die Transformation eines Bildes  $f(x, y)$ . Gesucht ist nun ein  $H(u, v)$ , so dass die Rücktransformation  $g(x, y)$  von  $G(u, v) = H(u, v)F(u, v)$  ein entstörtes (weichgezeichnetes) Bild ergibt. Ideal wäre also (im kontinuierlichen Fall), wenn sämtliche Frequenzpaare  $(u, v)$ , deren

(absolutes) geometrisches Mittel größer einer bestimmten Frequenz  $D_0$  ist, ausgefiltert werden:

$$D(u, v) := \sqrt{u^2 + v^2} \quad (47)$$

$$H(u, v) = \begin{cases} 1 & \text{wenn } D(u, v) \leq D_0 \\ 0 & \text{sonst} \end{cases} \quad (48)$$

Der Butterworth Filter hat im Unterschied keinen scharfen Abschnitt:

$$H(u, v) = \frac{1}{1 + \left(\frac{D(u, v)}{D_0}\right)^{2n}} \quad (49)$$

Hochpassfilter ist analog definiert – im idealen Fall bspw.  $H'(u, v) := 1 - H(u, v)$ .

Homomorphic Filtering: Ein Bild  $f(x, y)$  kann aus zwei Komponenten zusammengesetzt gedacht werden:  $f(x, y) = \iota(x, y) \cdot \rho(x, y)$ , wobei  $0 < \iota < \infty$  die Beleuchtung angibt,  $0 < \rho < 1$  die Reflektion. Da die FT bzgl. der Multiplikation nicht strukturverträglich ist (allerdings bzgl. der Addition) und es manchmal wünschenswert ist, auf beiden Komponenten separat zu arbeiten, kann das Filterergebnis  $g(x, y)$  wie folgt gewonnen werden:

$$\begin{aligned} z(x, y) &:= \ln f(x, y) \\ &= \ln \iota(x, y) + \ln \rho(x, y) \end{aligned} \quad (50)$$

$$\underbrace{\mathcal{F}\{f'(x, y)\}}_{:=Z(u, v)} = \underbrace{\mathcal{F}\{\iota(x, y)\}}_{:=I(u, v)} + \underbrace{\mathcal{F}\{\rho(x, y)\}}_{:=R(u, v)} \quad (51)$$

Es werde nun der Filter  $H(u, v)$  angewandt:

$$\begin{aligned} g(x, y) &= \exp(\mathcal{F}^{-1}\{H(u, v)Z(u, v)\}) \\ &= \exp(\mathcal{F}^{-1}\{H(u, v)I(u, v)\}) + \\ &\quad \mathcal{F}^{-1}\{H(u, v)R(u, v)\} \\ &= \exp(\mathcal{F}^{-1}\{H(u, v)Z(u, v)\}) \cdot \\ &\quad \exp(\mathcal{F}^{-1}\{H(u, v)R(u, v)\}) \end{aligned} \quad (52)$$

Zusammenfassend:  $f(x, y) \rightarrow \boxed{\ln} \rightarrow \boxed{\text{FFT}} \rightarrow \boxed{H(u, v) \cdot \star} \rightarrow \boxed{\text{FFT}^{-1}} \rightarrow \boxed{\exp} \rightarrow g(x, y)$

## 6. Restauration

### Restaurierung

Modell im Orts- und Fourier-Bereich:

$$g(x, y) = \underbrace{f(x, y)}_{\text{Originalbild}} * \underbrace{h(x, y)}_{\text{Degrädierung}} + \underbrace{n(x, y)}_{\text{„noise term“}} \quad (53)$$

$$G(u, v) = F(u, v)H(u, v) + N(u, v) \quad (54)$$

Das Copyright für die dieser Zusammenfassung zugrunde liegenden Vorlesungsunterlagen (Skripte, Folien, etc.) liegt beim Dozenten. Darüber hinaus bin ich, Florian Schoppmann, alleiniger Autor dieses Dokuments und der genannte Dozent ist in keiner Weise verantwortlich. Etwaige Inkorrektheiten sind mit sehr großer Wahrscheinlichkeit erst durch meine Zusammenfassung/Interpretation entstanden.

Inverse Filtering, Näherung für Original:

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)} - \frac{N(u, v)}{H(u, v)} \quad (55)$$

Problematisch beim Inverse Filtering sind Fälle, in denen  $H(u, v)$  sehr kleine Werte annimmt (oder – schlimmer noch – sogar ganz verschwindet).

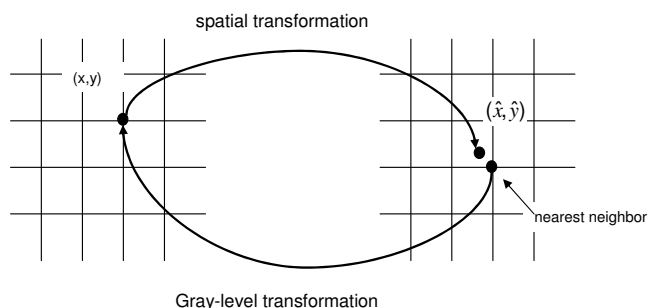
Der Least Mean Square (LMS) Filter, wobei  $S_f(u, v)$  und  $S_n(u, v)$  das Leistungsspektrum von  $f_e$  bzw.  $n_e$  darstellen (den für die diskrete Faltung mit 0en erweiterten Funktionen  $f$  und  $n$ ):

$$\hat{F}(u, v) = \frac{\overline{H(u, v)}}{|H(u, v)|^2 + \underbrace{\gamma \left( \frac{S_n(u, v)}{S_f(u, v)} \right)}_{=:k}} \cdot G(u, v) \quad (56)$$

Wenn  $S_n$  und  $S_f$  unbekannt, so muss  $k$  selbst bestimmt werden.

Der LMS für  $\gamma = 1$  heißt Wiener Filter. Für ihn gilt:  $E[(f(x, y) - \hat{f}(x, y))^2]$  wird minimal.

**Geometrische Verzerrung**



Neue Grauwertberechnung: Nearest Neighbor oder bilineare Interpolation. Faustregel: „Backward mapping“ sinnvoller als „forward mapping“.

**Segmentierung**

Erkennung von Objekten auf Bildern, Unterscheidung von Bildern, Suche nach Bildern

Verfahren: Punktoperationen zur Trennung von Objekten und Hintergrund (Schwellwert, etc.), Klassifikation durch Speichern bestimmter Merkmale wie Grauwerte, lokaler Kontrast, Form/Lage bestimmter Pixel, etc.

Erkennung des Objektes nach der Segmentierung durch Pattern Recognition, Computer Vision, Experten

**Pyramiden / Multi Resolution Processing**

Ein Wavelet ist Basisfunktion der Wavelet-Transformation. Im Gegensatz zu Sinus und Cosinus besitzt ein Wavelet nicht nur Lokalität im Frequenzspektrum sondern auch im Zeitbereich. Wavelets haben daher die Form von nach außen hin auslaufenden (kleiner werdenden) Wellen (also „Wellchen“ = Wavelet).

Pyramiden sind Bilder in unterschiedlichen Auflösungen. Von voller Auflösung  $2^n \times 2^n$  kommt man durch Downsampling (bspw. Durchschnittsbildung, Gaussfilter, Nearest Neighborhood) auf kleinere Bilder  $2^j \times 2^j$ ,  $0 \leq j < n$ .

Implementierung mit Residual Pyramiden in jedem Level  $2^j \times 2^j$ ,  $0 < j \leq n$ , die nur die Differenz zwischen Upsampling des vorhergehenden Levels und aktuellem Level speichern. Der Speicherbedarf einer vollen Pyramide für ein Bild der Größe  $2^n \times 2^n$  ist beschränkt durch:

$$\sum_{j=0}^n 2^j 2^j = \sum_{j=0}^n 4^j = \frac{4^{n+1} - 1}{4 - 1} = \frac{4}{3} (4^n - \frac{1}{4}) < \frac{4}{3} (2^n 2^n) \quad (57)$$

Beim Subband Coding wird das Bild mittels Band Pass Filtern in einige wenige Frequenzbänder zerlegt, die in Summe das Originalbild errechnen lassen. Sie sind oft einfacher zu komprimieren als das Originalbild.

**7. Kompression**

Einige grundlegende Speicherungsformate: Pixelbild, Lauflängen-Codierung, Baumstruktur, RGB-Farbmodell oder bspw. YIC (Y: Helligkeit, I und C: Chrominanz) mit nur zwei Farbkomponenten, Speicherung des Orts- oder Frequenzbereichs

Mögliche Baumstruktur (Quadtrees): Wenn das Bild nicht aus einem konstanten Wert besteht, unterteile es in vier gleiche Teile. Setze rekursiv fort.

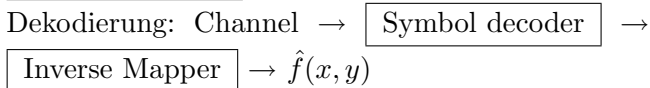
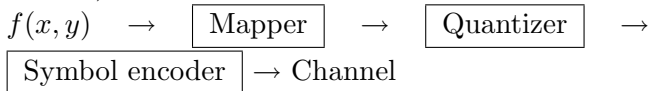
Baumstruktur als Pyramiden: Speichere in jeder Ebene den Durchschnittswert der *im Originalbild* überdeckten Pixel.

Die Cosinus-Transformierte kann aus der Fourier-Transformierten berechnet werden und eignet sich zur Speicherung, da sie nur reellwertig und kompakter als die FT ist.

Beispiele für verlustfreie (lossless) Kodierungen: Lauflängen, Huffman, Quadtrees Verlustbehaftete

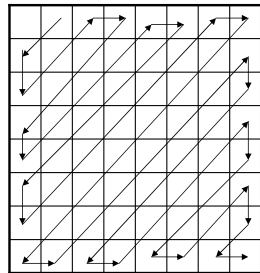
(lossy) Kodierungen: Pyramidenebenen weglassen, im YIC-Farbmodell nur die Helligkeit mit voller Auflösung, im RGB Würfel Farbwerte neu quantisieren, bei der Cosinus-Transformation nur die „wichtigen“ Werte speichern

Schema für Kodierung (Verlust tritt beim Quantizer auf):



Schema beim JPEG: Mapper transformiert von RGB in Helligkeitsmodell (hier *kann* Verlust auftreten) und nimmt diskrete Cosinus-Transformation (DCT) vor, Quantizer *kann* Auflösung der DCT-Werte vergrößern, Symbol encoder benutzt Huffman und Lauffängen-Codierung

Rechts: DCT Komponenten geordnet nach Frequenz. Bei JPEG werden die Werte nach „Wichtigkeit“ quantisiert. Die Folge der Komponenten links oben wird für alle  $8 \times 8$ -Matrizen Huffman kodiert. Die jeweils anderen Komponenten werden geordnet nach Frequenzen (also in Pfeilrichtung) abgespeichert und Lauffängen-kodiert (es treten viele 0en auf).



Wavelet-Komprimierung: Zerlegung des Bildes in eine Skalierung und 3 Wavelets (horizontal, vertikal, diagonal). Verlustfreie Rekonstruktion möglich, wenn keine Quantisierung vorgenommen wurde. Eignung für Pyramiden:

