

1 Bilder

Ein digitales Bild entsteht durch die Diskretisierung einer kontinuierlichen Szene.

- **Lochkameramodell:** Projektion der 3D-Welt auf eine 2D-Ebene.
- **Rasterisierung:** Umwandlung des kontinuierlichen Signals in ein digitales Gitter (Pixel).
- **Repräsentation:** Ein Bild wird als Matrix von Intensitätswerten (Grauwerten oder Farbkanälen) dargestellt.

1.1 Digitale Bildverarbeitung im Ortsraum

Im Ortsraum (Spatial Domain) werden Operationen direkt auf den Pixelwerten des Bildes durchgeführt.

1.1.1 Pixeloperationen

Pixeloperationen manipulieren einen Pixelwert unabhängig von seiner Nachbarschaft (globaler Kontext, aber lokale Anwendung).

Grauwert-Abbildung (Mapping)

Die Transformation eines Pixels an der Stelle (m, n) wird beschrieben durch:

$$g[m, n] = T(f[m, n])$$

wobei f das Eingabebild, g das Ausgabebild und T die Transformationsfunktion ist.

Wichtige Pixeloperationen:

- **Negativ:** $g[m, n] = f_{max} - f[m, n]$ (Invertierung der Helligkeit).
- **Binärisierung (Thresholding):**

$$g[m, n] = \begin{cases} f_{max} & \text{falls } f[m, n] > \tau \\ f_{min} & \text{falls } f[m, n] \leq \tau \end{cases}$$

Dient zur Trennung von Vorder- und Hintergrund.

- **Grauwertfensterung:** Hervorheben eines bestimmten Intensitätsintervalls (z.B. in der Medizintechnik/CT).
- **Kontrastspreizung:** Abbildung des genutzten Grauwertbereichs auf die volle Skala (z.B. 0-255), um den visuellen Kontrast zu erhöhen.
- **Mittelung:** Unterdrückung von unkorreliertem Rauschen durch Mittelung über mehrere Aufnahmen derselben Szene ($g = \frac{1}{k} \sum f_i$).

1.1.2 Histogramme

Ein Histogramm ist die graphische Darstellung der Häufigkeitsverteilung der Grauwerte in einem Bild.

- **Bildhelligkeit:** Entspricht dem Mittelwert aller Grauwerte.
- **Bildkontrast:** Entspricht der Varianz aller Grauwerte.

Histogrammausgleich (Equalization): Ziel ist es, die Grauwerte so umzuverteilen, dass sie gleichmäßig über den gesamten Bereich verteilt sind (Maximierung des Kontrasts). Dies geschieht über die **Summenwahrscheinlichkeit** (kumulative Verteilungsfunktion).

$$p(g) = \max(\text{Intensität}) \cdot \sum_{i=0}^g p(i)$$

Dies ist eine verlustbehaftete Operation und nicht umkehrbar.

1.1.3 Filterung im Ortsraum (Konvolution)

Hierbei wird der neue Wert eines Pixels unter Berücksichtigung seiner Nachbarschaft berechnet (lokaler Kontext).

Faltung (Convolution)

Die Faltung eines Bildes f mit einer Filtermaske w (Kernel) ist definiert als:

$$(f * w)(m, n) = \sum_{i=-k/2}^{k/2} \sum_{j=-l/2}^{l/2} w(i, j) \cdot f(m + i, n + j)$$

Tiefpass-Filter (Glättung) Dienen der Rauschunterdrückung und Weichzeichnung (Blurring).

- **Eigenschaften:** Koeffizienten sind positiv und summieren sich meist zu 1 (Normalisierung), um die Helligkeit zu erhalten.
- **Box-Filter (Mittelwert):** Alle Koeffizienten sind gleich (z.B. alle $\frac{1}{9}$ bei einer 3×3 Matrix). Nachteil: Erzeugt Artefakte (rechteckige Unschärfe).
- **Gauß-Filter:** Approximation der Gauß-Glocke.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Gewichtet zentrale Pixel stärker als äußere. Besser als Box-Filter, da natürliche Glättung. Kann durch Binomialfilter approximiert werden.

Nicht-Lineare Filter

- **Median-Filter:** Ersetzt den Pixelwert durch den **Median** der Nachbarschaft.
- **Vorteil:** Entfernt "Salt-and-Pepper"-Rauschen (Ausreißer) extrem effektiv und **erhält Kanten** (im Gegensatz zu Tiefpassfiltern, die Kanten verschmieren).

Hochpass-Filter (Kantenextraktion) Dienen dem Finden von Kanten (Hervorhebung von hohen Frequenzen/Änderungen).

- **Eigenschaften:** Koeffizienten sind positiv und negativ, Summe ist meist 0. Ergebnis kann negative Werte enthalten.
- **Gradienten (1. Ableitung):** Reagieren auf Rampen und Stufen (Kanten).
- **Laplace-Filter (2. Ableitung):** Isotrop (richtungsunabhängig).

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Reagiert stark auf feine Details und Rauschen ("Double Response" an Kanten: Nulldurchgang).

- **Laplacian of Gaussian (LoG):** Auch "Mexican Hat" oder Sombrero-Filter. Kombination aus Glättung (Gauß) zur Rauschunterdrückung und Laplace zur Kantenfindung.

1.2 Digitale Bildverarbeitung im Frequenzraum

Bilder können durch Transformationen (z.B. Fourier) in Frequenzanteile zerlegt werden.

- **Tiefe Frequenzen:** Grobe Strukturen, Flächen, homogener Hintergrund. (Im Zentrum des Spektrums).
- **Hohe Frequenzen:** Feine Details, Kanten, Rauschen. (Außen im Spektrum).

Faltungssatz

Eine Faltung im Ortsraum entspricht einer Multiplikation im Frequenzraum (und umgekehrt).

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

1.2.1 Ablauf der Frequenzfilterung

1. Transformation des Eingabebildes f mittels FFT (Fast Fourier Transform) $\rightarrow F(u, v)$. 2. Multiplikation mit der Filterfunktion $H(u, v)$. 3. Rücktransformation mittels inverser FFT \rightarrow gefiltertes Bild $g(x, y)$.

1.2.2 Filtertypen im Frequenzraum

- **Idealer Tiefpass:** Schneidet alle Frequenzen oberhalb D_0 ab (Rechteckfunktion). Führt im Ortsraum zu **Ringing-Artefakten** (Wellen um Kanten), da die FT einer Rechteckfunktion eine Sinc-Funktion ist.
- **Gauß-Tiefpass:** Weicher Übergang. Kein Ringing, da FT einer Gauß-Funktion wieder eine Gauß-Funktion ist.
- **Hochpass:** Kehrt das Prinzip um (Zentrum wird auf 0 gesetzt, hohe Frequenzen bleiben).

Aliasing: Tritt auf, wenn Abtastrate zu gering (Unterabtastung) oder im Frequenzraum Frequenzen abgeschnitten werden ("Leakage").

1.3 Bildkompression

Ziel: Reduktion der Datenmenge durch Eliminierung von Redundanzen (räumlich, zeitlich, spektral) und Irrelevanzen (nicht wahrnehmbare Informationen).

1.3.1 Klassifikation

- **Verlustfrei (Lossless):** Original kann exakt wiederhergestellt werden. (z.B. Huffman, RLE, LZW \rightarrow PNG, GIF, TIFF).
- **Verlustbehaftet (Lossy):** Informationen gehen verloren, höhere Kompressionsraten. Nutzt Wahrnehmungsmodelle des Menschen (z.B. JPEG, MPEG).

1.3.2 JPEG (Joint Photographic Experts Group)

Ein Standard für verlustbehaftete Kompression, optimiert für natürliche Fotos (fließende Übergänge).

Die 5 Schritte des JPEG-Baseline-Codecs:

1. **Farbraum-Transformation:** Umwandlung von RGB in $Y C_b C_r$.
 - Y : Luminanz (Helligkeit).
 - C_b, C_r : Chrominanz (Farbdifferenzen Blau/Rot).
2. **Farb-Subsampling (Downsampling):** Da das menschliche Auge Helligkeit besser auflöst als Farbe, werden C_b und C_r räumlich reduziert (z.B. 4:2:0 oder 2x2 Mittelung). Y bleibt voll erhalten.
3. **Block-Aufteilung & DCT (Diskrete Kosinustransformation):** Das Bild wird in 8×8 Blöcke unterteilt. Jeder Block wird mittels DCT in den Frequenzraum transformiert.
 - Ergebnis: 1 DC-Koeffizient (Gleichanteil/Grundhelligkeit) und 63 AC-Koeffizienten (Wechselanteile).
 - Energiekonzentration: Meist sind nur wenige Koeffizienten (niedrige Frequenzen) signifikant groß.
4. **Quantisierung (Der eigentliche Verlust):**

$$F^Q = \text{Round}(F_{ij}/Q_{ij})$$

Die DCT-Koeffizienten werden durch eine Quantisierungsmatrix Q geteilt und gerundet.

- Hohe Frequenzen werden stark quantisiert (durch große Werte in Q), oft zu Null.
- Steuert die Qualität vs. Kompressionsrate.

5. Kodierung (Entropy Coding):

- **Zig-Zag-Scan:** Sortiert die Koeffizienten von niedrigen zu hohen Frequenzen, um lange Ketten von Nullen zu erzeugen.
- **RLE (Run-Length Encoding):** Kodiert die Nullen-Ketten effizient.
- **Huffman-Kodierung:** Weist häufigen Werten kurze Bitfolgen zu.