Einführung in Bildverarbeitungsproblematik im Hinsicht auf Fouriertransformation

Dimitri Ovrutskiy

Digitale Bilder

Digitale Bilder (diskrete Signale) besitzen einen diskreten Definitionsbereich (Sampling) und einen diskreten Wertebereich (Quantisierung)

Rauschen

Die Signale werden fast immer gestört. Wir werden uns jetzt einer Art der Störung wenden - dem **multiplikativen Rauschen**.

- Das Rauschen hängt vom Signal ab (im Gegensatz zum additiven Rauschen)
- Oft ist es proportional zum Grauwert: $f_{ij} = g_{ij} + n_{ij}g_{ij} = (1 + n_{ij})g_{ij}$
- Z.B. die durch die Korngröße fotografischer Emulsionen erzeugte Störungen.

Außerdem unterscheidet man noch additives und Impulsrauschen (z.B. Pixelfehler einer Kamera-Matrix).

Rauschen-Maß (Signal-Rausch-Abstand)

Ist g das Ursprungsbild und f das verrauschte Bild, so wird **Signal-Rausch-Abstand** (Dezibel, dB) wie folgt definiert:

$$SNR(f,g) = 10\log_{10}\frac{\sum_{i,j} g_{ij}^2}{\sum_{i,j}(g_{ij}-f_{ij})^2}$$

Manchmal betrachtet man den Peak-Signal-Rausch-Abstand *PSNR*, wobei der Zähler durch

$$N \max_{i,j} g_{i,j}^2$$

ersetzt wird, wobei N die Gesamtzahl der Pixel im Bild ist. Dabei ist das Rauschen mit PSNR gröser als 30 kaum noch wahrnehmbar.

Unschärfe

- neben Rauschen die zweite Störungsquelle
- wird verursacht z.B. durch athmosphärische Störungen, Defokusierung, Unzulänglichleiten des Linsensystems oder Bewegungsunschärfe
- kann interpretiert werden als Grauwertmittelung innerhlab einer Nachbarschaft, Gewichte der Mittelung und Form der Nachbarschaft hängen von der Störungsquelle ab.
- Mittelung läßt sich mathematish als **Faltung** beschreiben

Faltung

Mathematisches Modell der Störungen

1D Faltung

- Faltung zweier diskreten 1D-Signale $f = (f_i)$ und $w = (w_i)$:

$$(f * w)_i := \sum_k f_{i-k} w_k$$

- Einträge von w kann man als Gewichte (gespiegelt), mit denen man f mittelt. Z.B. sei f_i der Börsenkurs am Tag i. Mit $w_i = 1/180, \ i = 1, ..., 180$ und 0 sonst ergibt sich

$$h_i = \sum_k f_{i-k} w_k = \frac{1}{180} \sum_{k=1}^{180} f_{i-k}$$

- kontinuierlicher Fall der Faltung:

$$(f * w)(x) := \int f(x - x') w(x') dx'$$

2D Faltung

- Faltung zweier diskreten Bilder $f = (f_{ij})$ und $w = (w_{ij})$:

$$(f * w)_{ij} := \sum_{k,l} f_{i-k,j-l} w_{k,l}$$

- kontinuierlicher Fall der Faltung:

$$(f * w)(x, y) := \int \int f(x - x', y - y') w(x', y') dx' dy'$$

- Bespiel: ist $w(x, y) = \begin{cases} 1 & f\ddot{u}r\ x^2 + y^2 \leq r^2 \\ 0 & sonst, \end{cases}$ so beschreibt $f * w$ die Glättung eines kontinuierlichen Bildes f durch Mittelung aller Grauwerte innerhalb einer Nachbarschaft mit Radius r.

Die Berechnung im Ortsbereich ist i.A. aufwändig für große Radien.
Die Fouriertransformation bietet eine effiziente Alternative.

Unschärfemodelierung mit Faltungskernen

1. **Defokusierte Kameraoptik**: Zylinderartiger Faltungskern
2. **Atmosphärische Störungen bei Teleskopen**: ungefähr zweidimensionaler Gaußkern
3. **Bewegungsunschärfe**: im Idealfall eindimensionale Boxfunktion längs der Bewegungsrichtung.

Fouriertransformation

Eigenschaften der kontinuierlichen Fouriertransformation

Eigenschaften der FT 1

- Linearität

$$F[af + bg] = aF[f] + bF[g]$$

- Ähnlichkeitssatz

$$F[f(ax, by)](u, v) = \frac{1}{|ab|} F[f] \left(\frac{u}{a}, \frac{v}{b} \right)$$

Dehnung im Ortsbereich bewirkt Stauchung im Fourierbereich.

Eigenschaften der FT 2

- Verschiebungssatz

$$F[f(x-x_0,y-y_0)](u,v) = e^{-i2\pi(ux_0+vy_0)}F[f](u,v)$$

  Verschiebung im Ortsbereich bewierkt Phasendrehung im Frequenzbereich. Das *Fourierspektrum* bleibt jedoch unverändert! In diesem Sinn ist die FT translationsinvariant. Zudem ist sie rotationsinvariant.

- Ableitungssatz

$$F\left[\frac{\partial^{n+m}}{\partial x^n \partial y^m} f(x,y)\right](u,v) = (i2\pi u)^n (i2\pi v)^m F[f](u,v)$$

Eigenschaften der FT 3

- Symmetrieeigenschaften

$$Re(\hat{f}(u,v)) = Re(\hat{f}(-u,-v))$$

$$Im(\hat{f}(u,v)) = -Im(\hat{f}(-u,-v))$$

- **Faltungssatz** Das Faltungsintegral läßt sich bequem als Multiplikation im Fourierbereich berechnen:

$$F[f * g] = F[f] \cdot F[g]$$

FT spezieller Funktionen 1

- **FT des Gaußkerns** Die FT eines Gaußkerns ist ein Gaußkern mit rezipoker Varianz:

$$f(x,y) = \exp\left(\frac{-\pi(x^2+y^2)}{\sigma^2}\right) \Rightarrow \hat{f}(u,v) = \exp\left(\frac{-\pi(u^2+v^2)}{\sigma^{-2}}\right)$$

- **FT eines Deltakamms** Die FT eines unendlich ausgedehnten Kamms von Deltaimpulsen mit Peakabstand λ

$$f(x) = \sum_{k=-\infty}^{\infty} \delta(u - k\lambda)$$

ist ein Deltakamm mit rezipokem Peakabstand:

$$\hat{f}(x) = \sum_{k=-\infty}^{\infty} \delta(u - k/\lambda)$$

Wird benötigt beim Samplen von kontinuierlichen Signale

FT spezieller Funktionen 2

FT einer Boxfunktion $f = A$ für $x \in [0, X]$ und $f = 0$ sonst.

$$\hat{f}(u) = \frac{A}{\pi u} e^{-\pi u X} \sin(\pi u X).$$

Mit $|e^{-\pi u X}| = 1$ ist Fourierspekrum $AX|sinc(\pi u X)|$.
f hat endliche Ausdehnung im Ortsbereich, $\hat{f}$ aber unendliche Ausdehnung im Frequenzbereich.

Diskrete FT 1

- Diskretes Analogon zu kontinuierlichen FT
- soll auf gesampelten Signalen mit M Messwerten arbeiten und sie in M Frequenzkomponenten zerlegen.

Diskrete FT 2

Die diskrete Fouriertransformation (DFT) eines 1D-Signals $f = (f_0, ..., f_{M-1})^T$ lautet

$$\hat{f}_p := \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} f_m \exp\left(-\frac{i2\pi pm}{M}\right), \qquad (p = 0, ..., M-1),$$

mit Rücktransformation

$$f_m := \frac{1}{\sqrt{M}} \sum_{p=0}^{M-1} \hat{f}_p \exp\left(\frac{i2\pi pm}{M}\right), \qquad (m = 0, ..., M-1)$$

DFT als Basiswechsel

- Die M Vektoren $(p = 0, ..., M-1)$

$$v_p = \frac{1}{\sqrt{M}} \left(\exp\left(\frac{i2\pi p0}{M}\right), ..., \exp\left(\frac{i2\pi p(M-1)}{M}\right) \right)^T$$

  eine Orthonormalbasis des m-dimensionalen komplexen Raums

- Stellt man ein Vektor $f = (f_i)_{i=0}^{m-1}$ bzgl. dieser Basis dar, so ergibt sich mit dem Skalarprodukt $< f, g >= \sum_{m=0}^{M-1} f_m \bar{g}_m$

$$f = \sum_{p=0}^{M-1} < f, v_p > v_p$$

  die diskrete Fourierdarstellung mit Fourierkoeffizienten $\hat{f}_p =< f, v_p >$.

Diskrete 2D FT

Die diskrete Fouriertransformation (DFT) eines 2D-BildesSignals $f = (f_{m,n})$ mit $m = 0, \dots, M-1,\ n = 0, \dots, N-1$ lautet

$$\hat{f}_{p,q} := \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{m,n} \exp\left(-\frac{i2\pi pm}{M}\right) \exp\left(-\frac{i2\pi pn}{N}\right)$$

$(p = 0, \dots, M-1;\ q = 0, \dots, N-1)$ mit Rücktransformation

$$f_{m,n} := \frac{1}{\sqrt{MN}} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \hat{f}_{p,q} \exp\left(\frac{i2\pi pm}{M}\right) \exp\left(\frac{i2\pi pn}{N}\right)$$

$(m = 0, \dots, M-1;\ n = 0, \dots, N-1)$

Eigenschaften der DFT 1

- Höherdimensionale DFT ergeben sich analog.
- Wie die kontinuierliche FT ist auch die DFT separabel.
- Viele Eigenschaften übertragen sich sinngemäß von der kontinuierlichen FT: *Linearität*, *Verschiebungstheorem*, *Faltungstheorem*
- An einem diskreten Gitter nur bedingt erfüllbar: *Skalierungssatz*, *Rotationsinvarianz*

Man verwendet die kontinuierliche FT gern zum Filterdesign, und die DFT zur Filterimplementierung.

Eigenschaften der DFT 2

- Wichtiger Unterschied zu kontinuierlichen FT:
  Signal f hat endliche Ausdehnung $f_0, \dots, f_{M-1}$.
  Wegen Periodizität der komplexen Exponentialfunktion wird das Signal durch die Fourierdarstellung periodisch fortgesetzt.
  Das führt zu unerwünschten Randeffekten:
    - Unstetigkeiten an periodisch fortgesetzten Ränder erzeugen hohfrequente Fourierkomponenten
    - Wraparound-Fehler bei Faltungen: rechter Rand stört Ergebnis am linken Rand.
- Dynamischer Bereich des Fourierspektrums erstreckt sich über viele Zehnerpotenzen. Für Visualisierungszwecken daher logarithmisch transformieren:

$$D_{p,q} = c \log\left(1 + |\hat{f}_{p,q}|\right)$$

Lineare Filter

Grundbegriffe der linearen Systemtheorie

Lineare Systeme

- Lineares System: Ein linearer Filter L erfüllt $\forall \alpha, \beta \in \mathbb{R}$ $\forall$ Bilder $f_1, \; f_2$
$$L(\alpha f_1 + \beta f_2) = \alpha L(f_1) + \beta L(f_2)$$
- Lineares ortsinvariantes System (LSI-System, linear shift-invariant): lineares Signal, dessen Ergebnis nur vom Eingangssignal, aber nicht vom Ort, anhängt:
$$LT_b f = T_b L f$$
für alle Verschiebungen T_b mit $(T_b f)(x) = f(x - b)$

Impulsantwort eines LSI-Systems

- Impulsantwort = Filterergebnis für einen Deltaimpuls
  $h = L\delta_0$ mit $\delta_i(j) = 1$ für $i = j$ und $= 0$ sonst.
- Da jedes diskrete Signal $f = (f_1, ..., f_N)^T$ als Summe von N Impulsen dasrgestellt werden kann, gilt auf Grund der Linearität und Ortsinvarianz

$$Lf = \sum_{i=1}^{N} f_i T_i L\delta_0$$

- Das zeigt: ein lineares ortsinvariantes System ist durch seine Impulsantwort eindeutig festgelegt.

Bedeutung der Faltungen in der linearen Systemtheorie

- Faltungen sind linear und ortsinvariant
- Es gilt sogar die Umkehrung:
  Lineare ortsinvariante Systeme sind stets Faltungen!
- die Faltungsmaske (Faltungskern) ist identisch mit der Impulsantwort.

Bedeutung der Fouriertransformation

- Da sich Faltungen im Ortsbereich in Mutiplikationen im Fourierbereich überführen lassen, spielt die Fouriertheorie eine fundamentale Rolle beim verständnis linearer Filter.
- Faltungen mit großen Masken werden effizienter im Fourierbereich durchgeführt.
- Lineare Filter werden gerne in den Fourierbereich transformiert, um ihr Frequenzverhalten zu studieren.
- Oft entwickelt man Filter direkt im Fourierbereich und transformiert sie danach in den Ortsbereich zurück.

Grundtypen LSI-Filter

- *Tiefpass*: lässt tiefe Frequenzen besser passieren als hohe
- *Hochpass*: lässt hohe Frequenzen besser passieren als tiefe
- *Bandpass*: lässt hauptsächlich Strukturen eines vorgegebenen Frequenzbandes passieren

Tiefpassfilter 1

- Glätten eines Bildes durch elimination von Rauschen und unwichtigen feinskaligen Details
- Ortsbereich: Faltung mit Mittelungsmasken
- Frequenzbereich: Dämpfen hoher Frequenzen

Tiefpassfilter 2

- Boxfilter im Ortsbereich:
  - $(2m+1)(2m+1)$-große Faltungsmaske mit Gewichten $\frac{1}{(2m+1)^2}$
  - Auch bei großen Masken effizient im Ortsbereich realisierbar, mit linearem Aufwand
  - Ergebnis oft nicht sehr gut: bevorzugt horizontale & vertikale Sturkturen.
  - Kontinuierliche FT der Boxfunktion ist $sinc-$artig
  - Unbefriediegendes Frequenzverhalten: mehrere Maxima, Dämpfung hoher Frequenzen nur mit $1/|u|$
- Optimalität im Frequenzbereich, der ideale Tiefpass:
  - sämtliche Frequenzen (u,v) mit $u^2+v^2>T^2$ werden 0 gesetzt
  - ebenfalls unbefriedigend: *sinc*-artige rotationsinvariante Funktion im Ortsbereich mit sichtbaren Ring-Artefakten

Gaußkerne

- *m*-Dimensionaler Gaußkern:

$$K_\sigma(x) = \frac{1}{(2\pi\sigma^2)^{m/2}} \exp\left(-\frac{|x|^2}{2\sigma^2}\right)$$

"Breite" σ - *Standardabweichung*, σ^2 - die *Varianz*

- Frequenzbereich: Wieder Gaußkern mit rezipoker Varianz
- guter Kompromiss: jeweils nur ein Maximum im Orts- und Frequenzbereich
- einziger Faltungskern, der separabel und rotationsinvariant ist
- bei iteriereten Gaußkernen addieren sich die Varianzen
- zahlreiche Möglichkeiten zur effizienter Realisierung einer Faltung mit einem Gaußkern

Approximationsmöglichkeiten für einen Gaußkern 1

- Abtastung des Gaußkerns im Ortsbereich:
  - Separabilität und Symmetrie werden ausgenutzt. Abtastintervall $[-k\sigma, k\sigma]$ liefert sehr gute Genauigkeit für $k \geq 3$. Koeffizientensumme auf dem Intervall auf 1 renormieren.
  - **Pro's:** einfach und flexibel (σ kontin. einstellbar). **Contra's:** Komplexität wächst mit σ
- Multiplikation im Fourierbereich:
  - **Pro's:** Komplexität fast linear: $O(N \log N)$
  - **Contra's:** Randeffekte durch periodische Fortsetzung; Standard-FFT an bestimte Bildgrößen gebunden
- Frequenzbereich: Dämpfen hoher Frequenzen

Approximationsmöglichkeiten für einen Gaußkern 2

- Binomialkerne (Tabelle):
  - Grenzwertsatz von Moivre-Laplace: Binomialkerne approximieren Gaußkerne
  - Separabilität und Symmetrie ausnutzen. Iterierte Binomialkerne ergeben erneut Binomialkerne.
  - **Pro's:** prinzipiell in Integerarithmetik durchführbar (Division durch Zweierpotenzen als Shiftoperation)
    Contra's: Komplexität wächst mit σ; σ nicht kont. einstellbar
- Iterierte Boxfilter:
  - Zentraler Grenzwertsatz der Statistik: Iterierte Mittelungskerne konvergieren gegen eine Gaußverteilung
  - Dreifach iterierter Boxfilter ist bereits gute Approximation
  - **Pro's:** lineare Komplexität, unabhängig von σ
    Contra's: σ nicht kontinuierlich einstellbar

Hochpass- und Bandpassfilter

- können als Differenz von Tiefpass-gefiltertem und dem Originalbild (Hochpass) und als Differenz zweier Tiefpässe mit unterschidlichen Parameter (Bandpass) designet werden
- In diesem Fall werden genau so wie Tiefpässe behandelt
- Hochpässe entfernen tieffrequente Hintergrundstörungen und schärfen verschwommene Bildstrukturen;
  konnen destabilisierend wirken
- Es gibt viel kompliziertere, aber auch viel effizientere Hochpässe, z.B. die Ableitungsfilter, die hier nicht besprochen werden
- Bandpässe dienen der Extraktion interessanter Bildstrukturen auf bestimmten Skalen

Entfernung strukturierter Störungen

- Periodische Störungen treten z.B. als Gitterraster im Zeitungsdruck auf
- Liniertes oder kariertes Papier erzeugt ebenfalls periodische Störungen
- Periodische Störungen lassen sich im Fourierbereich gut lokalisieren und entfernen

Homomorphe Filterung 1

- Aufbereitung eines unter unregelmäßigen Beleuchtungsverhältnissen aufgenommenen Bildes
- Grauwert eines Bildes ist Produkt aus Beleuchtungsintensität $I(x, y)$ und Reflektionsvermögen $R(x, y)$:

$$f(x, y) = I(x, y)R(x, y)$$

  - Die Beleuchtungsintensität ändert sich meist nur langsam, ist also eine tieffrequente Komponente
  - Das Reflektionsvermögen kann sich abrupt ändern, z.b. an Kanten, daher viele Hochfrequenzanteile

Homomorphe Filterung 2

- Berechene den Logarithmus aller Grauwerte

$$\log f(x,y) = \log I(x,y) + \log R(x,y)$$

  Wegen der Summe ist jetzt die FT einsetzbar.
- Bilde die FT des entstandenen Bildes
- Multipliziere im Fourierbereich mit einer Funktion, die tiefe Frequenzen dämpft und hohe anhebt
- Berechne die inverse FT
- Bilde die Exponentialfunktion aller Grauwerte

(Pseudo-)Inverse Filterung 1

- **Model einer Bilddegradation**

$$f(x,y) = (h * u)(x,y) + n(x,y)$$

  Hierbei: f das vorliegende degradierte Bild; u das unbekannte ideale Bild; h ein verschiebungsinvarianter bekannter Faltungskern; n das Rauschen (durch Sensor, Übertragung, Quantisierung,...)

- Der **Der Fahltungskern** h **modeliert** z.B. **Defokusierung** der Kameraoptik (zylinderartig); **Bewegungunschärfe** (orientierte Boxfunktion); **Atmosphärische Störungen** bei Teleskopen (ungefär gaußarig); **Pfusch bei der Produktion** beim Hubble-Teleskop (gaußartig)

(Pseudo-)Inverse Filterung 2

- Falls (!) man das Rauschen vernachlssigen kann, ergibt sich im Fourierbereich
$$\hat{f} = \hat{h} \cdot \hat{u}$$
- Falls (!) $\hat{f}$ nicht verschwindet, rekonstruiert man u aus der inversen FT von
$$\hat{u} = \frac{\hat{f}}{\hat{h}} \qquad (\textit{inverse Filterung})$$

(Pseudo-)Inverse Filterung 3

- h typischerweise ein Glättungskern (Tiefpass)
- Für hohe Frequenzen liegt $\hat{h}$ meist dicht bei 0
- einfachste Stabilisierung (*pseudoinverse Filterung*):

$$\hat{u} = \begin{cases} \frac{\hat{f}}{\hat{h}} & \textit{falls } \hat{h} > \epsilon \\ 0 & \textit{sonst} \end{cases}$$

Wiener-Filterung (N. Wiener (1894-1964)) 1

- stabilisierte Form der inversen Filterung
- approximiert $\hat{u}$ durch

$$\hat{u} \approx W_K(\hat{h}, \hat{f}) := \left(\frac{1}{\hat{h}} \frac{|\hat{h}|^2}{|\hat{h}|^2 + K} \right) \hat{f}$$

  mit einem positiven Parameter K
- kann nicht mehr divergieren für $\hat{h} \to 0$: $W_K(\hat{h}, \hat{f}) \to 0$
- verhalt sich wie ein Bandpass: *Hochpasseigenschaft* durch inverse Filterung; *Tiefpasseigenschaft* durch K
- K sollte von der geschätzten Rauschvarianz σ^2 abhängen:

$$K := 2\sigma^2$$

Wiener-Filterung (N. Wiener (1894-1964)) 2

- wesentlich robuster gegenüber Rauschen als inverse Filterung
- gilt als das beste lineare Verfahren für Dekonvolutionsprobleme mit additivem Gauß-Rauschen
- gleichzeitiges Schatzen des verschiebungsinvarianten Faltungskerns und des ungestörten Bildes (blinde Dekonvolution) erfordert wesentlich tiefergehende Mathematik (Variationsansätze)