

Ausarbeitung inoffizieller Fragenkatalog – Grundlagen der digitalen Bildverarbeitung, Prof. Schuster

Autor: Murrel (Murrel.vienna@gmx.at)

Teil 1

1) Beschreiben Sie die 3 Arten von Bildverarbeitung: Computergrafik, Digitale Bildverarbeitung, Mustererkennung (Unterschiede, Vor- und Nachteile beschreiben) (S. 22-45)

Generell gilt: Die Graphische Datenverarbeitung erzeugt ein Bild aus einer Beschreibung (Objektwelt -> geometrisches Modell), die Bildverarbeitung verändert es (Filterung, Transformation, Restauration/Verbesserung, Bewegungsanalyse) und die Mustererkennung erzeugt eine Beschreibung aus einem Bild (Merkmalsextraktion wie Konturen oder Invarianten, Klassifikation, Interpretation).

- Graphische Datenverarbeitung (image synthesis): generiert aus symbolischer Beschreibung (nicht bildhafte Information) Bilder von einem bestimmten Blickwinkel -> Synthese abstrakter geometrischer Elemente zu Bildern. Ziel ist also die Modellierung bzw. bildhafte Darstellung von Objekten (Bsp: Computerspiel, Simulationen, CAD, Architektur, synthetisches Video). Man differenziert Rendering (make a pretty picture) und Visualization (make a useful picture)
- Digitale Bildverarbeitung: Transformation eines Bildes in ein anderes. Erfolgt in 2 Schritten: Segmentierung (Zerlegung in Bereiche, denen inhaltliche Bedeutung zugeordnet werden kann) und Analyse (dem Bild wird Bedeutung zugeordnet). Als klassische Bildverarbeitung versteht man einfache Verbesserungen der Bilddarstellung (in der Medizin sehr wichtig zB bei Hervorheben bestimmter relevanter Aspekte). Man unterscheidet Bildrestauration und Bildverbesserung.
- Mustererkennung: Eingabe eines Bildes bei Ausgabe einer Entscheidung, Klassifikation oder symbolischer Beschreibung durch Merkmalsextraktion, Formdarstellung und Segmentierung (Texturanalyse, Bildvergleich, syntaktische Analyse). Generell wird zunächst ein Bild vorverarbeitet (korrigiert), dann symbolisch abgebildet und beschrieben. Vorteile der Mustererkennung:
 - Bildinhalt in numerischer Form: liefert quantitative Aussagen und ist reproduzierbar
 - erkennen von Eigenschaften, die der Mensch übersieht
 - Trennschärfe der Merkmale kann bestimmt werden
 - entlastet Personal bei Routinearbeiten
 - höhere Effizienz bei Bildfolgen
 - Objektivierung der Analyse

2) Was ist Bildrestauration, was Bildverbesserung? (S. 25-33)

- Bildrestauration: Bei systematischer Verzerrung, ist unabhängig vom Bildinhalt, nur das Abbildungssystem ist wichtig. Es treten systematisch auf:

- radiometrische Fehler (zB atmosphärisch wie Streuung, Dunst etc..., Verschmierung, shading [gleiche Grauwerte entsprechen nicht gleichen Helligkeiten])
- geometrische Fehler (zB Instrumentenfehler, Panoramaverzerrungen, Änderungen der Kameralage, perspektivischer Fehler, Erdrotation)
- 50-Hz-Rauschen

-Bildverbesserung: Bildinhalt wird anders (evtl sogar verfälscht) dargestellt, um Inhalte besser erkennbar zu machen. Ist abhängig vom Bildinhalt und problemorientiert. Typische Methoden:

- Transformation auf bi-level Bilder (Text)
- Kontrastanhebung (Histogramm-modifikation)
- Pseudofarbbilder (-> Colour map)
- Multibilder (-> Hauptachsentransformation, SOMs -> Falschfarbcodierung).

3) Was versteht man unter der Bildpyramide? (S. 40-41)

Für das computergestützte Auffinden von Objekten ist oftmals eine geringere Auflösung (sowohl räumlich wie numerisch) zweckmäßig¹³ und kann dadurch beträchtlich beschleunigt werden. Auch brauchen bestimmte Analyseschritte verschiedene hohe Auflösungsgrade, um effizient und verlässlich arbeiten zu können. Aus diesem Grund wurden neue Bilddatenstrukturen entwickelt, wie etwa die *Bildpyramiden*. Bildpyramiden sind Datenspeicher, deren unterste Ebene dem Bild mit der größtmöglichen Auflösung entspricht. Die jeweils darüber angeordneten Speicherebenen haben abnehmende räumliche Auflösung, die durch Zusammenfassen von bestimmten Elementen der unteren Ebene entstanden sind. Durch die nach oben hin abnehmende räumliche Auflösung entsteht die Pyramidenstruktur. Auf welche Art und nach welchen Vorschriften Elemente der jeweils unteren Ebene zusammengefasst werden soll, muss je nach Fragestellung frei programmierbar sein. Dies setzt eine vollständige Vernetzung aller Speicherzellen und zwischen den einzelnen Speicherebenen angesiedelte Mikroprozessoren voraus, welche die Transformation (Komprimierung, Extraktion) der Information der unteren Ebene in jene der übergeordneten Ebene durchführen. Ein wesentliches Entwurfskriterium solcher Bildpyramiden ist die dadurch mögliche Kontrolle (Steuerung) der einzelnen Schritte bei der Informationsverdichtung.

Teil 2

4) Was sind Multibilder? (S. 2-5)

Generell wird in der Bildverarbeitung mit natürlichen Bildern gearbeitet, also Bilder, welche eine 2-dimensionale räumliche Verteilung von Helligkeiten darstellen. Sie hängen von verschiedenen Parametern ab:

$$\text{Bild} = f(x, y; \lambda_0, t_0, \rho_0)$$

λ_0 bestimmtes Spektralband
 t_0 bestimmter Zeitpunkt
 ρ_0 Polarisierung

f hängt generell von Sonnenstand und Betrachtungswinkel ab.

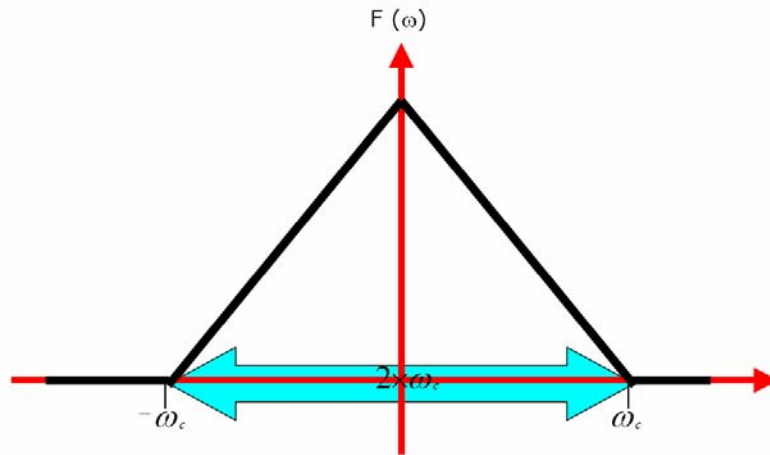
Wird die Szene mit mehreren Parametern (zB von verschiedenen Sensoren) aufgenommen, so werden die einzelnen Bilder f_1, f_2, \dots, f_p zu einer vektorwertigen Funktion zusammengefasst:

$$\text{Diese vektorwertige Funktion } \overrightarrow{f(x, y)} = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \\ \vdots \\ f_p(x, y) \end{pmatrix} \text{ heißt } \underline{\text{Multi-Bild}}.$$

Beispiele hierfür wären multi-spektrale Bilder (Messungen in verschiedenen Spektralbereichen) oder multi-temporale Bilder (Messungen zu verschiedenen Zeitpunkten)

5) Was bedeutet bandbegrenzt, was ist ein Abtastgitter? (S. 28-29, 34-38)

Eine Funktion $f(t)$ heisst bandbegrenzt, wenn es ein ω_c gibt, sodass $F(\omega) = 0$ und $|\omega| > \omega_c$



Da alle technischen Geräte träge sind, können sie nur Frequenzen bis zu einer Maximalfrequenz registrieren, d.h. alle Signale, die über technische Geräte geleitet werden, sind bandbegrenzt.

Abtastgitter und reziproke Abtastgitter ermöglichen es, die Theorie des Abtastens (Funktion wird durch äquidistante Abtastwerte dargestellt) auf 2D-Funktionen zu erweitern:

Eine periodische Anordnung von Punkten in der x-y-Ebene, bei der die Punkte durch die Positionsvektoren

$$\vec{r}_{m,n} = m \times \vec{r}_1 + n \times \vec{r}_2 \quad m \in \mathbb{Z} \quad n \in \mathbb{Z}$$

definiert sind, heisst Abtast-Gitter

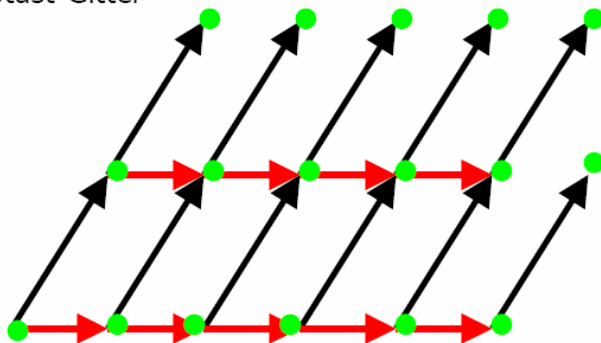
Eine periodische Anordnung von Punkten in der u/v-Ebene, deren Position durch die Vektoren

$$\vec{\omega}_{m,n} = m \times \vec{\omega}_1 + n \times \vec{\omega}_2 \quad m \in \mathbb{Z} \quad n \in \mathbb{Z}$$

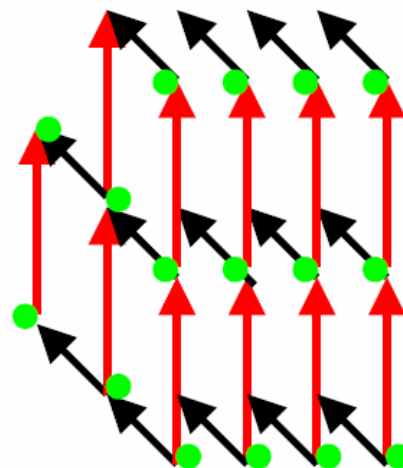
definiert ist, heisst reziprokes Abtast-Gitter.

(wobei ω_1 und ω_2 ein reziprokes Gitter in der u/v Ebene definieren)

Abtast-Gitter



reziprokes Abtast-Gitter



Hiermit kann eine Funktion aus ihren Abtastwerten rekonstruiert werden.

6) Was ist Quantisierung? (S. 42-46 [+47-50])

Bei der Quantisierung wird der Wertebereich der möglichen Funktionswerte in G Intervalle geteilt und alle Werte innerhalb eines Intervalls werden durch einen einzigen Wert q_l dargestellt. Ändert sich der Funktionswert nur langsam, ist eine feine Quantisierung wichtig, denn sonst können leicht falsche Konturen erzeugt werden.

Einfacher ist es, wenn die Quantisierungsgrenzen äquidistant sind, allerdings ist es optimaler, wenn sie in den Teilbereichen, in denen viele Werte liegen, enger beisammen und in den anderen weiter auseinander liegen. Diese Methode nennt man getaperte Quantisierung.

7) Was für Pixelrelationen (Nachbarschaftsbeziehungen) gibt es? (S. 60-62)

- $N_4(p)$: Die Menge der 4 direkten Nachbarpixel heißt 4-Nachbarn von p und wird $N_4(p)$ geschrieben

- $N_D(p)$: Die Menge der 4 indirekten Nachbarpixel (die diagonal anliegenden Pixel) schreibt man $N_D(p)$

- $N_8(p)$: Direkte und indirekte Nachbarn zusammen nennt man die 8-Nachbarn von p und schreibt man $N_8(p)$.

Zusammenhängend sind die Pixel aber erst, wenn sie Nachbarn (je nach verwendeter Beziehung) sind und gleichzeitig gleiche Grauwerte besitzen. Im Allgemeinen definiert man zur Modellierung der Ähnlichkeit eine Menge V von Grauwerten, und nennt die Pixel ähnlich, wenn beide Grauwerte aus V besitzen. Zwei Pixel p und q mit Grauwerten aus V sind:

- 4-zusammenhängend, wenn $q \in N_4(p)$

- 8-zusammenhängend, wenn $q \in N_8(p)$

Man nennt dementsprechend eine Menge von Pixeln 4-Weg, wenn sie alle 4-zusammenhängend sind, und 4-Region, wenn es für alle Pixelpaare aus R einen 4-Weg gibt, der in ganz R liegt (Analog für 8-zusammenhängend).

Mittels des Connected Components Labeling lassen sich über zusammenhängende Pixel je nach Beziehung zusammenhängende Regionen markieren.

Teil 3

8) Was ist die LUT (Lookup Table)? / Wie werden Werte in Helligkeiten umgerechnet? *2 (S. 2-4)

Die LookUp Table (LUT) ist eine hardwaremäßig geschaltete (zwischen Datenspeicher und Darstellungsgerät) Liste von Wertepaaren, die jedem Grauwert r einen neuen Grauwert $s = T(r)$ zuordnet. Sie wird daher bei Punktoperatoren (Grauwertabbildungen) genutzt. (s. auch nächste Frage)

Man kann Werte auch anders in Helligkeiten umrechnen:

- lineare Zuordnung über den gesamten Wertebereich: monoton wachsend oder fallend (Bildnegativ, notwendig da sich im Negativ manche Bilder besser erkennen lassen)

- lineare Zuordnung außerhalb der Extrembereiche: ein Teil der extremen Helligkeitswerte wird auf das Maximum bzw. Minimum abgebildet, wird gewählt wenn nur wenige Pixel außerhalb der Darstellungsgrenzen A und B liegen.

9) Wie erfolgt die Umsetzung von Modifikationen in der Bilddarstellung? (S. 20)

Es gibt 2 Möglichkeiten:

- herkömmlich: Zunächst wird die Zuordnungsfunktion $p(i)$ für den aktuellen Bildausschnitt berechnet, dann nach und nach einzeln jeder Pixel eingelesen, sein neuer Wert berechnet und der neue Wert in die Bildmatrix eingetragen. Dies hat jedoch die Nachteile, dass die Berechnung für jeden Pixel neu durchgeführt werden muss und dass Originaldaten teilweise irreversibel zerstört werden.

- Übersetzungstabellen: Die Ergebnisse der Transformation werden in einer Übersetzungstabelle (LookupTable [LUT] oder Video LookUpTable [VLT]) gespeichert, es fallen nur so viele Berechnungen an wie es Quantisierungsstufen gibt. Diese Übersetzungstabelle ist hardwaremäßig zwischen Datenspeicher und Darstellungsgerät eingeschaltet: Werte im Datenspeicher werden nicht direkt am Bildschirm dargestellt, sondern dienen vorerst nur als Adresse zur Übersetzungstabelle. Dies ermöglicht es, die Darstellung des Bildes rasch und korrigierbar (reversibel) zu verändern.

10) Wie erfolgt Kontrastspreizung? (S. 4 und S. 21-22)

Normalerweise treten in einem Bild Grauwerte eines bestimmten Bereiches R sehr häufig auf und alle Werte außerhalb von R eher selten. In solchen Fällen kann man R aufspreizen, und damit wird der Kontrast für einen großen Teil des Bildes erhöht. Generell reichen hierfür 1-2 Knickpunkte in einem stückweisen linearen Punktoperator.

Diese Bereichsspreizung wird angewendet, wenn der gewünschte Bereich der Graustufen [A,B] ist, aber die aufgenommenen Grauwerte im Bereich [a, b] mit $A \leq a < b \leq B$ liegen. Man transformiert den neuen Bereich mittels der Formel:

$$g_{neu} = \frac{(B - A) \times (g_{alt} - a)}{b - a}$$

Man kann hierbei auf 2 Arten vorgehen:

- Lokale Intervallspreizung: um jeden Bildpunkt wird eine n*n (zB 25*25) Umgebung ausgelesen, von diese Umgebung wird das lokale Minimum und Maximum ermittelt. Liegt ein Funktionswert der Umgebung nahe am lokalen Maximum, wird er heller, liegt er nahe am lokalen Minimum, wird er dunkler. Somit werden die Differenzen zwischen hellen und dunklen Punkten erhöht, wodurch der Kontrast erhöht wird.
 - Prozentuelle lokale Intervallspreizung: Bei der lokalen Intervallspreizung werden Kontraste oft zu stark angehoben. Im Gegensatz zu dieser 100%-Spreizung wird bei der prozentuellen lokalen Intervallspreizung nur um α % gespreizt, was eine geringe Kontrasterhöhung bewirkt.
- Bemerkung: Es ist NICHT sinnvoll, homogene Bereiche zu spreizen, diese werden über max-min erkannt und nicht gespreizt.

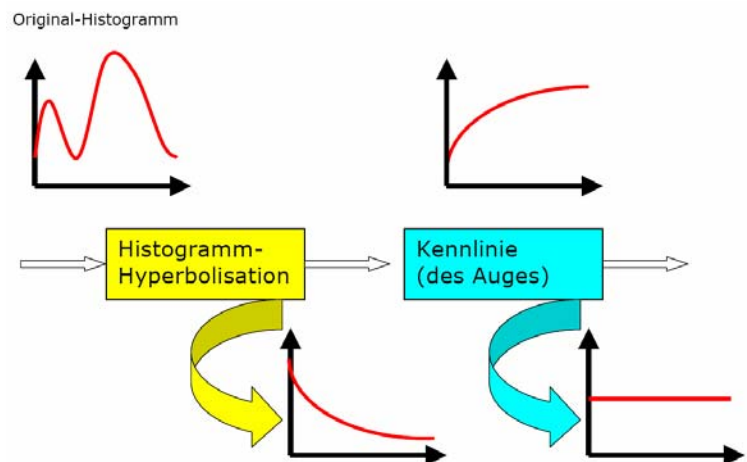
11) Wodurch charakterisiert sich Histogramm-Einebnung, wodurch Histogramm-Hyperbolisierung? Wieso spricht man davon, dass das Auge logarithmisch sieht?

***3 (S. 29-36)**

Die Häufigkeitsverteilung, mit der jeder Grauwert i im Bild vorkommt, heißt Histogramm.

Das Ziel der Histogramm-Einebnung ist, dass jeder Grauwert etwa gleich oft im Bild vorkommt. (Motivation: Informationstheorie – die beste Codierung ist jene, bei der alle Werte gleich oft vorkommen). In der Praxis ist eine genaue Gleichverteilung nicht möglich, durch Anwendung der Einebnungsformel findet sich vor Allem in der Mitte und an den Extremstellen des Grauwertbereichs eine geringere Verteilung wieder. Dies ergibt sich durch die diskreten Werte.

Ziel der Histogramm-Hyperbolisierung ist es, dass das Histogramm des Ausgangsbildes hyperbolisch ist. (Motivation: Das menschliche optische System reagiert auf Helligkeiten nicht linear, sondern annähernd logarithmisch [Gesetz von Weber])



Teil 4

12) Wodurch charakterisieren sich: Pseudofarbbilder, Falschfarbbilder, Differenzbilder, Quotientenbilder? (S. 16-21)

Bei Pseudofarbbildern werden die Grauwerte eines Bildes in eine beliebige Farbskala projiziert. Die Motivation dahinter ist, dass das menschliche optische System wesentlich mehr Farbtöne als Graustufen unterscheiden kann. Ist ein großer dynamischer Bereich im Originalbild vorhanden, kann durch Pseudofarbgebung mehr Information sichtbar gemacht werden.

Die folgenden Bilder sind Multi-Bild-Verbesserungen:

Beim Falschfarbbild werden die spektralen Informationen genauer dargestellt (verschiedene spektrale Eigenschaften -> andere Farbe).

Differenzbilder kommen zum Einsatz, wenn man spektrale oder temporale Differenzen besser visualisieren will. (Da Differenzen positiv oder negativ sein können, ist hier eine Skalierung notwendig, um nicht aus dem dynamischen Bereich zu fallen). Sie akzentuieren also die Variationen zwischen den verschiedenen Bildkomponenten (zB spektralen Bändern eines Multispektralbildes) und eliminieren den gemeinsamen Luminance-Bias.

Quotientenbilder oder auch Verhältnis-Bilder basieren auf dem Quotienten aus zwei Bildkomponenten. Dies unterdrückt die Helligkeitsvariation durch das topographische Relief und verstärkt so die spektrale Farbvariation. Somit wird der Beleuchtungsfaktor b eliminiert. ($f=r*b$, r =Reflexivität). Sie verstärken jedoch auch zufälliges Rauschen und kohärentes Rauschen, welches nicht mit den beiden Bildkomponenten korreliert. Andere Probleme sind, dass der Quotient stark schwanken kann und man nicht immer weiss, welche Quotienten und welche Farbzuteilungen sinnvoll sind (und es sind viele möglich).

13) Glätten im Ortsraum, Charakteristika von: Mittelwertfilter, selektive Mittelwertbildung, Makota-Verfahren, Binomialfilter, Medianfilter *5 (S. 33-44)

Glätten im Ortsraum dient generell dem Entfernen von weißem Rauschen, welches über alle Frequenzen gleichverteilt ist. Zur Glättung des Rauschens sollte der Filter eine Tiefpassform haben.

Ein Mittelwertfilter führt Mittelung über Umgebungen durch, d.h. im gesamten Bild werden die Pixel auf den arithmetischen Mittelwert ihrer Nachbarpixel (in oder exklusive des Punktes selbst) gesetzt. Je nach Radius können dies bei $r = 1$ die 4-Nachbarn, bei $r = \sqrt{2}$ die 8-Nachbarn oder bei mehr noch mehr Pixel in der Nähe sein. Generell wird r zwischen 1 und 5 gewählt, damit arbeitet man mit 3 bis 11 Pixeln. Diese Mittelwertbildung beseitigt hochfrequente Störungen, hat aber den Nachteil, dass echte Grauwertkanten unschärfer werden. Man kann diesem Verschmieren jedoch damit begegnen, indem man für jeden Pixel prüft, ob er auf einer Kante liegt, und in diesem Bereich nur die Kantenpunkte mittelt. Alternativ dazu kann man auch nur die Nachbarpixel mitteln, die sich ähnlich sind, oder das Makota-Verfahren anwenden (s.u.).

Die selektive Mittelwertbildung dient ebenfalls zur Rauschbeseitigung. Hier werden nur die Pixel verändert, welche einen sehr großen Wertunterschied zwischen ihnen und ihren Nachbarn aufweisen (Grenze S wird angegeben). Ihr neuer Wert wird nur aus dem Mittelwert ihrer Nachbarpixel, nicht aus ihrem eigenen, bestimmt.

Das Verfahren von Makota zielt darauf ab, Objektkanten bei der Mittelwertfilterung zu erhalten. In einer zB $5*5$ großen Matrix wird ein Bereich von 7 Pixel um das Zentralpixel rotiert, von den 8 anderen möglichen Umgebungen wird jeweils die Varianz (maximale Grauwertdifferenz usw...) berechnet. Von dem Bereich mit der kleinsten Varianz wird dann der Mittelwert errechnet. Somit werden nur ähnliche Bereiche gemittelt.

Der Binomialfilter hat im Grunde genommen dieselbe Funktion wie der Mittelwertfilter, besitzt jedoch andere Koeffizienten, diese entsprechen den Funktionswerten einer

diskretisierten Gaussglocke. Somit ist der Mittelpunkt am stärksten gewichtet, die 4-Nachbarn weniger stark und die 8-Nachbarn eher schwach.

Beim Medianfilter wird im Gegensatz zum Mittelwertfilter der Median der Nachbarpunkte genommen (d.h. der vorkommende Wert, für den im Bereich genau gleichviele größere und kleinere Werte existieren). Hierbei wird immer mit einer ungeraden Anzahl Werte gearbeitet. Er behandelt dünne Kurven und Linien wie Rauschen, rundet aber scharfe Ecken ab. Mittels einer einfachen Modifikation lassen sich mithilfe dieses Filters auch Verkleinerungen oder Vergrößerungen von Bildmustern erreichen.

14) Wie eliminiert man Salz und Pfeffer Rauschen? (S. 48-49)

Zwei lokale Operatoren pixmax, pixmin werden so definiert, dass sie einen Bildpunkt auf den größten (kleinsten) Wert seiner Umgebung (mit ihm als Mittelpunkt) setzen. Wird zuerst der eine Operator n-mal hintereinander angewandt und dann der andere, also

$P = \text{pixmax}^n [\text{pixmin}^n (\text{Bild})]$ (Elimination von schmalen Gipfeln)

$T = \text{pixmin}^n [\text{pixmax}^n (\text{Bild})]$ (Elimination von engen Tälern)

dann erzeugt das ein Bild, in dem alle Spitzen/Täler, die weniger als $2n$ breit waren, abgeflacht bzw. aufgefüllt worden sind. Diese Methode wird vor Allem angewendet, wenn das Rauschen aus isolierten Punkten besteht, vorausgesetzt, solche Punkte treten nicht als Bildinformation auf.

15) Was ist Highlight/Lowlight Enhancement? (S. 49)

Durch Kombination der Operatoren

$P = \text{pixmax}^n [\text{pixmin}^n (\text{Bild})]$

$T = \text{pixmin}^n [\text{pixmax}^n (\text{Bild})]$

kann der Kontrast von kleinen Spitzen und Tälern verstärkt werden, indem die Täler vom Originalbild abgezogen und die Spitzen zum Originalbild addiert werden. Dies nennt man highlight/lowlight enhancement.

16) Wodurch charakterisiert sich ein Tiefpassfilter? (S. 57-60)

Ein Filter, der hohe Frequenzen stark in ihrem Anteil reduziert oder ganz eliminiert (durch einen Faktor 0) wird Tiefpassfilter (=low-pass-filtering) genannt. Ein solcher Filter hat eine glättende Wirkung und entspricht einem linearen Glättungsfilter im Ortsraum. Man unterscheidet verschiedene Tiefpassfilter:

Bei einem idealen Tiefpass-filter werden also alle Frequenzen $\leq \omega_0$ ohne Abschwächung übernommen und die restlichen vollkommen unterdrückt. Dieser Wert ω_0 heißt Cut-off-frequenz und ermöglicht es, verschiedene Tiefpassfilter zu vergleichen. Nachteile des Filters sind Verschmierung und ringing (= Überschwingen an scharfen Kanten). Ausserdem muss ω_0 groß gewählt werden, da in den hohen Frequenzen auch die Kanteninformation liegt.

Der trapezoidale Tiefpassfilter hat einen weniger starken Übergang zwischen Nichtunterdrückung und Unterdrückung. Nachteil daran ist das ringing (wegen Unstetigkeitsstelle, ist aber nicht so arg wie beim idealen TPF).

Ein Butterworth-Tiefpassfilter der Ordnung n hat keine Unstetigkeitsstelle, seine Cut-off-frequenz liegt bei 50% des Maximalwertes.

Der exponentielle Tiefpassfilter bedient sich der Exponentialfunktion.

An Hochpassfiltern existieren analog dieselben Filter.

Ein idealer Bandpassfilter ist die Kombination eines Tiefpass- und Hochpassfilters, ein Notch Filter ein inverser Bandpass-filter.

17) Kanten: Was sind Gradientenverfahren, was ein Gradient? Was ist ein anisotroper / isotroper Filter? Geben Sie zu letzteren beiden je einen Filter als Beispiel an. Wodurch charakterisieren sich Laplace-Masken? *3 (S. 67-70)

Gradientenfilter beruhen auf dem Prinzip, für die Kantendetektion des Bildes dessen 1. partielle Ableitung zu betrachten:

Im Fall von linearen Gradienten wird zur Grauwertermittlung die Grauwertdifferenz zwischen benachbarten Bildpunkten ermittelt, was der Differenziation des Bildes gleichkommt, wobei zu berücksichtigen ist, dass Kanten nicht nur in x- und y-Richtung auftreten können, sondern in jeder beliebigen Richtung.

Mittelwert-Differenz-Operatoren berechnen die Steigung des Grauwertes in einem Punkt, indem die Grauwertdifferenzen zu seinen Nachbarn ermittelt werden und die Steigung als arithmetischer Mittelwert der beiden linearen Gradienten von x und dem linken und x und dem rechten Nachbarn angenommen wird.

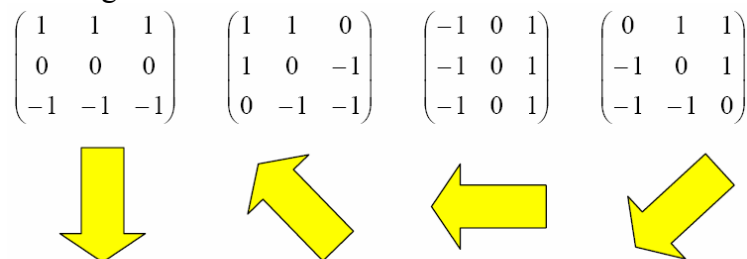


Abbildung 5.5: Mittelwert-Differenz-Operatoren

Eine horizontale/vertikale Grauwertkante lässt sich gut durch den symmetrischen x-Gradientenfilter/y-Gradientenfilter errechnen. Beide zusammen nennt man den Prewitt-Operator:

$$G_x^1 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ x-Gradient} \quad G_y^1 = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \text{ y-Gradient}$$

3x3 Masken der Gradientenfilter

Diese Filter erkennen horizontale oder vertikale Kanten, Diagonalkanten jedoch nur unzureichend. Man nennt sie daher anisotrop. Ein weiteres Problem des Prewitt-Operators ist, dass er zwei Bilder erzeugt, während man an einem einzigen interessiert ist, das alle Kanten zeigt. Dieses Problem kann umgangen werden, indem man die beiden Ableitungen zu einem Gradientenvektor kombiniert und dessen Betrag errechnet (in der Praxis wird jedoch meist einfach nur der arithmetische Mittelwert der beiden genommen). Auch diese Kombination bleibt jedoch anisotrop.

Laplace-Masken ermöglichen das Entdecken von Kanten ohne Bevorzugung bestimmter Richtungen (sie sind daher isotrop). Die Anisotropie wird dadurch gemildert, dass mit der 2ten Ableitung der Bildfunktion statt mit der ersten gearbeitet wird.

Der Laplace-Operator (für Radius $r = 1$)

Laplace-Operator (Radius $r = 1$)		
$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & -2 & -1 \\ -2 & 12 & -2 \\ -1 & -2 & -1 \end{pmatrix}$
nur unmittelbare Nachbarn (N_4)	alle benachbarten Pixel (ungewichtet)	alle benachbarten Pixel (N_8) (gewichtet)

ist eine einfache Approximation für die 2. Ableitung.

Dieser Filter reagiert auf nichtlineare Krümmungen im Grauwertbereich unabhängig von ihrer Richtung. Großer Nachteil des Laplace-Operators ist jedoch, dass er auf Rauschen ebenso empfindlich wie auf Kanten reagiert. Aus diesem Grund schaltet man vor seiner Anwendung

oft eine Glättung des Bildes vor. Man kann die Glättung auch gleich in den Laplace-Operator einbauen, indem man die Masken kombiniert (d.h. Binomial-Operator* Laplace Operator ergibt glättenden Kantenoperator).

Teil 5

18) Mittels welcher Mechanismen erfolgt Kantendetektion? (S. 4-6)

Es existieren folgende Kantendetektoren auf Basis eines Gauss-Filters:

Der Marr-Hildreth-Edge-detector ist einer der komplexeren Kantendetektoren, da er zusätzlich zur Kantendetektion eine Glättung mit Tiefpassfilter durchführt. Sein Hauptproblem ist, dass es bei seiner Anwendung im Bereich von nichtlinearen Kanten zu „displacement“ zwischen Grauwertsprung und detektierter Kante kommt.

Der Canny-edge-detector ist ein komplexes Segmentierungsverfahren. Es ist optimal in den folgenden Bereichen: Es erkennt Kanten gut und erkennt keine falschen Kanten, diese Kanten werden an der Mitte ihres richtigen Auftretens markiert und jede Kante wird nur einmal erkannt.