Projektpartner: Universität Karlsruhe - Institut für Photogrammetrie und Fernerkundung
Datum: 15.11.2003

4 Digitale Bildverarbeitung

Einleitung

„Bildverarbeitung erfolgt, seitdem Leben auf höherer Entwicklungsstufe existiert. Tiere und Menschen verarbeiten Bildeindrücke ohne äußere Hilfsmittel nach analogen Prinzipien. Dies geschieht on-line, parallel, multispektral über ein wissensbasiertes System." [Bähr, Vögtle 1991, S. 60]

"Die digitale Bildverarbeitung ist die Gesamtheit der Verfahren, durch die ein digitales Bild rechnerisch in ein verändertes digitales Bild überführt wird"
[Bill, Zehner 2001].

Die Veränderungen können sich auf die geometrischen und/oder die radiometrischen Aspekte eines Bildes beziehen. Außerdem können sie durch Klassifizierung hervorgerufen werden.

Die digitale Bildverarbeitung dient der Datenaufbereitung und der Datenauswertung. Die angewandten Verfahren dienen dazu, den Bildinhalt für den Betrachter deutlicher darzustellen oder bestimmte Aspekte des digitalen Bildes herauszuarbeiten.

Sie findet einerseits in der Fernerkundung Anwendung, andererseits mit ihren Algorithmen auch in der Datenanalyse.

Dieses Kapitel beinhaltet:

- Die radiometrische und geometrische Vorverarbeitung von Daten
- Bildaufbereitung zu Visualisierungszwecken
- Operationen mit mehreren Kanälen
- Filteroperationen im Frequenz- und im Ortsraum
- Binärbildverarbeitung
- Klassifikation

bmb+f Bundesministerium für Bildung und Forschung
Das Projekt Geoinformation – Neue Medien für die Einführung eines neuen Querschnittsfaches ist ein Projekt im Rahmen des BMBF-Programms Neue Medien in der Bildung.

zip Ein Projekt im Zukunfts-Investitions-programm der Bundesregierung
Das Projekt wird im Rahmen des Zukunftsinvestitions- programms der Bundesregierung (zip) gefördert.

Inhalt

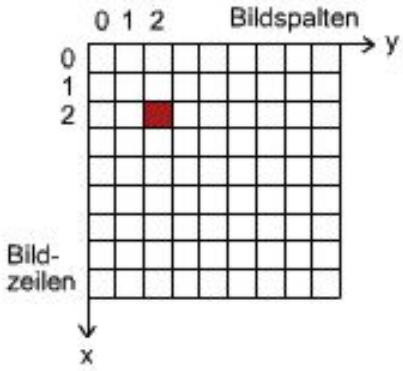
4. Digitale Bildverarbeitung

4.1 Grundlagen 2
4.1.1 Bildmatrix 3
4.1.2 Bildstatistik 5
4.1.2 Look-Up Table 7
4.2 Preprocessing 8
4.2.1 Radiometrische Korrektur 8
4.2.2 Geometrische Transformation 12
4.3 Bildverarbeitung zur Visualisierung 15
4.3.1 Kontraststreckung 16
4.3.2 Histogrammverebnung 17
4.3.3 Density Slicing 17
4.3.4 Pseudofarb-Transformation 18
4.3.5 Vergleich der Methoden 19
4.3.5 IHS-Transformation 20
4.3.6 Pan-sharpening 21
4.4 Mehrkanal-Operationen 22
4.4.1 Arithmetische Pixeloperationen 22
4.4.2 Ratiobildung 24
4.4.3 Hauptkomponentenanalyse 28
4.5 Filteroperationen 32
4.5.1 Ortsraum und Frequenzraum 32
4.5.2 Filterung im Frequenzraum 32
4.5.3 Filterung im Ortsraum 37
4.6 Klassifikation 45
4.6.1 Monospektrale Klassifikation 46
4.6.2 Multispektrale Klassifikation 47
4.6.3 Unüberwachte Klassifikation 49
4.6.4 Überwachte Klassifikation 50
4.6.5 Algorithmen der überwachten Klassifikation 53
4.6.6 Ablauf einer überwachten Klassifikation 55
4.7 Binärbildverarbeitung 60
4.7.1 Topologie / Nachbarschaft 61
4.7.2 Zusammenhangskomponenten 62
4.7.3 Distanzen und Distanztransformationen 64
4.7.4 Skelettierung 66
4.7.1 Topologische und geometrische Merkmale 68
4.7.2 Mathematische Morphologie 70
4.8 Literatur 73

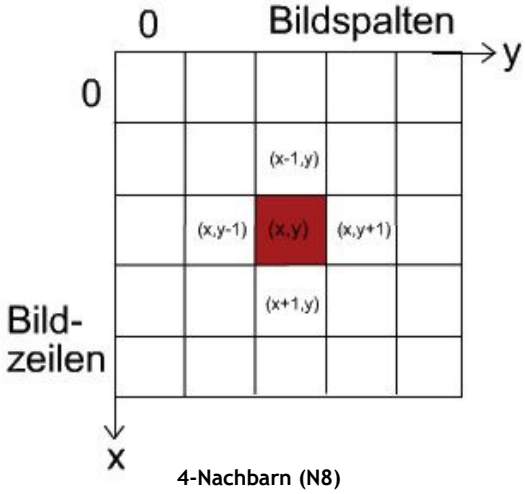
4.1 Grundlagen

4.1.1 Bildmatrix

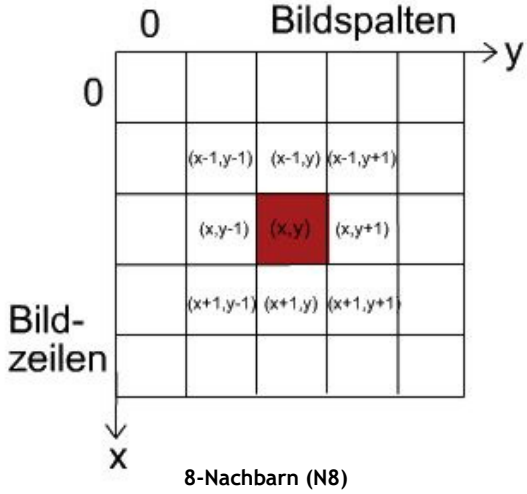
Eine Bildmatrix stellt eine regelmäßige Anordnung von Quadraten (auch Rechtecke sind möglich) dar. Sie gliedert sich in Bildzeilen (rows) und Bildspalten (columns), mit dem Nullpunkt in der linken oberen Ecke.



Jedem Pixel können bestimmte Nachbarschaftsbeziehungen zugewiesen werden:



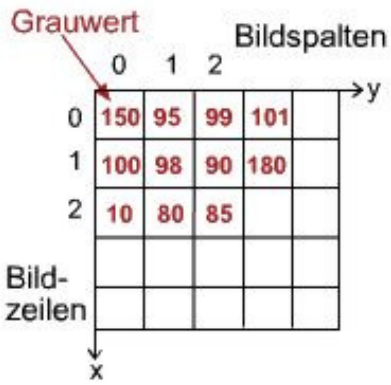
4-Nachbarn (N8)



8-Nachbarn (N8)

Berechnung des Abstands zwischen 2 Bildpunkten, (x_1, y_1) (x_2, y_2)= Euklidische Distanz: $d_e = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

Jedes Pixel besitzt einen Grauwert, z.B. bei einer radiometrischen Auflösung von 8-bit einen Wert von 0-255 . Ein Grauwert beschreibt die mittlere Reflexion für ein Gebiet (z.B. 30m) und ist immer eine positive ganze Zahl.



4.1.2 Bildstatistik

Hierzu gehören statistische Kennwerte, die die Eigenschaften von digitalisierten Bilddaten beschreiben. Sie erlauben Aussagen über die Grauwertverteilung im Bild. Wichtige Parameter sind:

- mittlerer Grauwert
- Standardabweichung
- Histogramm

mittlerer Grauwert:

$$\mu = \frac{1}{m \cdot n} \cdot \sum_{i=1}^{m} \sum_{j=1}^{n} P(i,j)$$

wobei *P(i,j)* ein Grauwertbild sei mit *m* Zeilen und *n* Spalten.

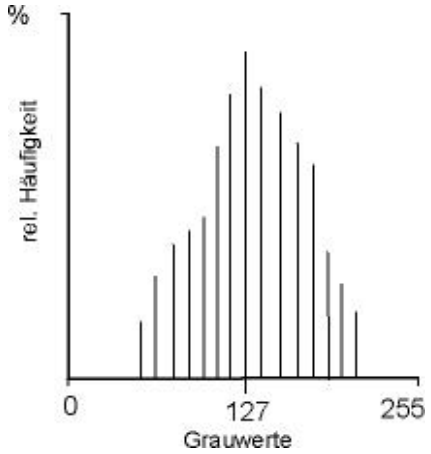
Standardabweichung: Ist ein statistisches Maß für die Verbreitung von Werten um ein Mittel.

$$\delta^2 = \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} (\mu - P(i,j))^2$$

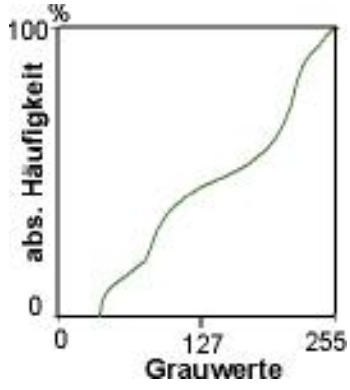
Histogramm:

Ein Histogramm ist eine tabellarische oder graphische Darstellung, die die Verteilung von Grauwerten zeigt (die absoluten oder relativen Häufigkeiten).

Auf der Abszisse werden die Grauwerte von 0-255 (bei 8-bit Auflösung), wobei 0=schwarz und 255=weiß ist, abgetragen, auf der Ordinate die prozentualen relativen Häufigkeiten der jeweiligen Grauwerte.

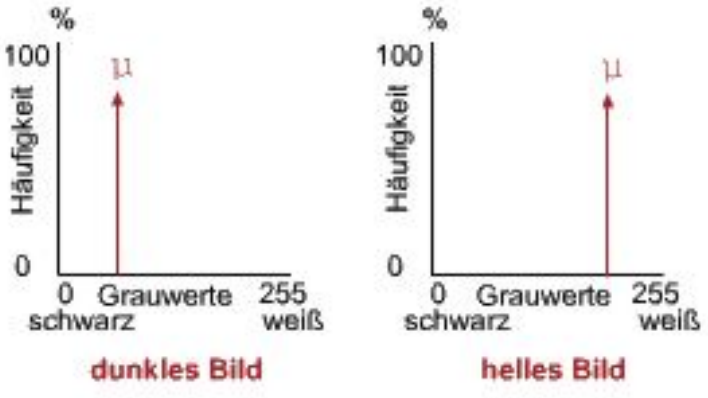


Bei einem kumulativem Histogramm (Abbildung rechts) wird die Häufigkeit für jeden Grauwert addiert und als Kurve eingezeichnet. Flache Abschnitte der Kurve entsprechen den Grauwerten, die im Bild nicht sehr häufig vorkommen, steile den häufig vorkommenden.



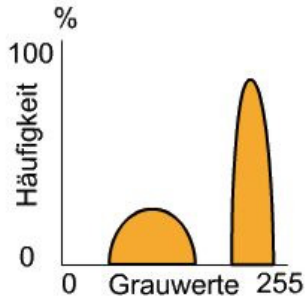
kumulatives Histogramm

Der mittlere Grauwert μ sagt etwas über die mittlere Gesamthelligkeit des Bildes aus. Hat die Mehrzahl der Grauwerte geringe Werte, handelt es sich um ein dunkles Bild, bei hohen Werten um ein helles.



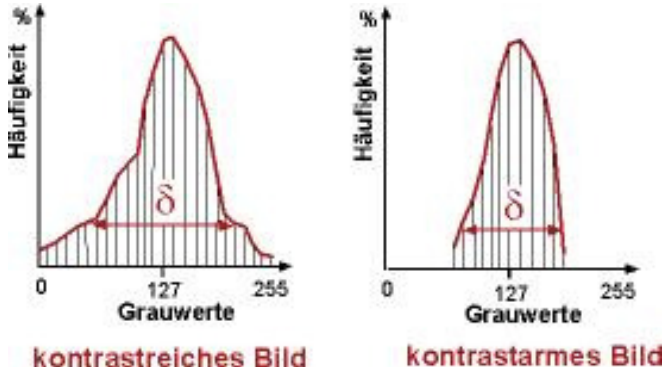
[nach Kappas 1994]

Ein Histogramm mit zwei lokalen Maxima, das also vorwiegend einen hellen und einen dunklen Bereich enthält, heißt **bimodal**.



bimodales Histogramm

Die Standardabweichung δ sagt etwas über den Kontrastumfang des Bildes aus:



[nach Kappas 1994]

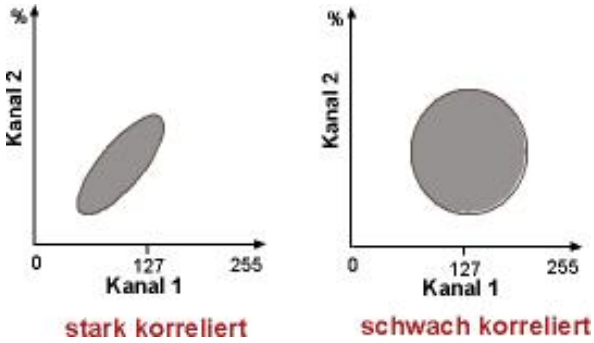
Das linke Histogramm ist das eines Bildes mit viel Kontrast; viele verschiedene Graustufen sind am Bildaufbau beteiligt; die Standardabweichung ist groß. Beim kontrastarmen Bild rechts gibt es nur eine geringe Streuung, es sind wenig Graustufen vertreten. Um die

Darstellung dieses Bildes zu verbessern, könnte man eine Kontraststreckung anwenden, ein Verfahren zur Bildverbesserung (siehe Kapitel 5.3.1)

Mehrdimensionale Histogramme:

Die bisher beschriebenen Histogramme bezogen sich auf einkanalige Bilder. Natürlich können sie auch im 2-dimensionalen oder n-dimensionalen Fall auftreten. Aus solchen Histogrammen lässt sich ablesen, wie stark Kanäle miteinander korreliert sind.

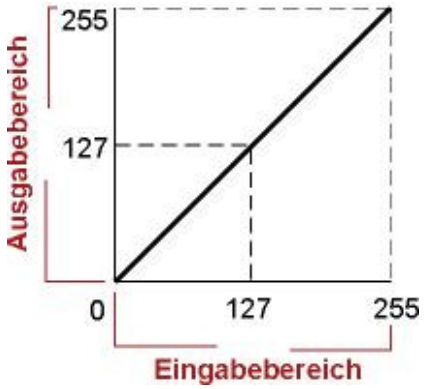
Auf der Ordinate werden die Grauwerte des einen Kanals abgetragen, auf der Abszisse die des anderen. Man erhält elliptisch geformte Grauwertverteilungen. Die Korrelation der beiden Kanäle kann dann an der Form der Ellipse abgelesen werden: Ist sie schmal und gestreckt, sind die beiden Kanäle stark korreliert; ist sie dagegen rund, sind sie nur schwach korreliert.



[nach Kappas 1994]

4.1.3 Look-Up Table

Die Darstellung von Farbbildern am Bildschirm erfordert, jedem Pixel eine Farbe aus einer Farbpalette zuzuweisen, was in einem Look-Up Table (LUT) gespeichert wird. Die Originaldaten werden dadurch nicht verändert, nur die Darstellung der Daten am Bildschirm wird modifiziert. Grauwerte werden dabei anderen Grauwerten zugewiesen, wobei meist 256 Farbintensitäten zur Verfügung stehen (entspricht 8 bit, 0-255).



Look-Up Table

LUTs werden zur Visualisierung (Pseudocolor, Klassifikation) und zur Unterdrückung oder Hervorhebung von Bildinformationen (z.B. Kontrastverstärkung) genutzt .

4.2 Preprocessing (Vorverarbeitung)

Rohe Fernerkundungsdaten beinhalten einige geometrische Fehler, fehlende Zeilen und Pixel, Rauschen und andere Probleme. Bevor diese rohen Daten analysiert und interpretiert werden können, müssen sie daher vorverarbeitet (engl: preprocessing) werden.

Methoden der Vorverarbeitung zielen darauf ab, eine oder mehrere der folgenden Effekte zu minimieren oder zu kompensieren:

Radiometrische Korrekturen:
- Atmosphärische Effekte
- Aufnahmezeitpunkt (Tageszeit, Jahreszeit)
- Topographie

Geometrische Korrekturen:
- Sensorbewegung
- Erdrotation
- Plattformbewegungen
- Topographie
- Erdkrümmung

4.2.1 Radiometrische Korrektur

Darunter versteht man Verfahren, bei denen fehlerhafte Bilddaten, die während der Datenaufnahme entstanden sind, korrigiert werden. Dabei werden die Grauwerte verändert, die geometrischen Eigenschaften jedoch bleiben erhalten. Die Korrekturen betreffen **atmosphärische Störungen**, **Sonnenstand** und **Sensorfehler**.

Viele dieser Korrekturen sind bereits von den Bodenempfangsstationen oder Einrichtungen, die digitale Luft- und Satellitenbilder vertreiben, durchgeführt worden. Die Definition, was als "Störung" betrachtet wird und daher korrigiert wird, ist abhängig von der Aufgabenstellung für die die Bilder gebraucht werden. Nicht alle der beschriebenen Operationen werden also notwendigerweise immer durchgeführt.

1. Einfluss der Atmosphäre:

Sonnenstrahlung interagiert mit der Atmosphäre durch Streuung und Absorption. Streuung führt zur Aufhellung des Bildes, Absorption (v.a. Wasserdampf oder andere Gase) bewirkt das Gegenteil. Dies führt zur Kontrastminderung im Bild. Die Wirkung ist abhängig vom Spektralbereich, kürzere Wellenlängen (z.B. sichtbares Licht) sind stärker davon betroffen als längere (z.B. Infrarot). Außerdem spielen die Zusammensetzung der Atmosphäre, der Sonnenstand und die Beobachtungsgeometrie eine Rolle.

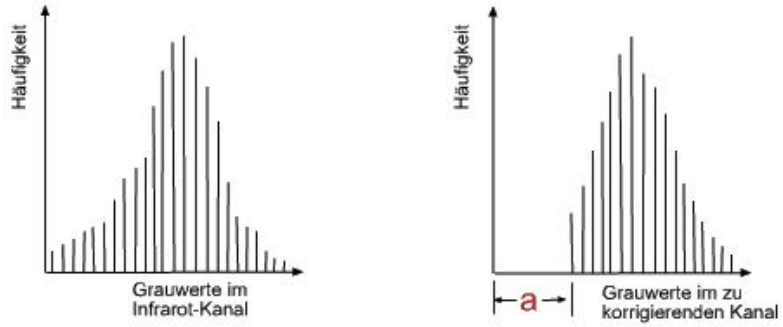
3 Möglichkeiten:

a. numerische Modelle

Die Algorithmen dieser Modelle gleichen die beobachtete Strahlung an Werte in klarer Atmosphäre an (z.B. LOWTRANZ, MODTRAN). Sie benutzen als Korrekturwerte Messungen von Absorption durch Aerosole und Gase (wie Wasserdampf oder Ozon).
Problem: die Anwendung ist sehr komplex und benötigt außerdem Zusatzinformationen meteorologischer Art, die oft nicht von dem Untersuchungsgebiet oder Zeitraum vorhanden sind.

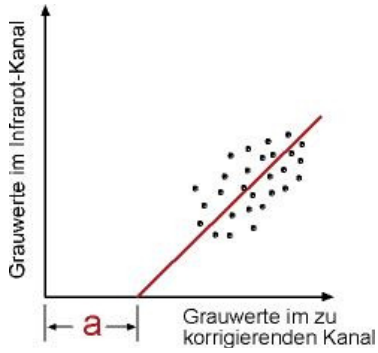
b. Histogrammvergleich (Histogram matching)

Da kürzere Wellenlängen von der Atmosphäre stärker beeinflusst werden als längere, geht man davon aus, dass der Infrarot-Kanal weitgehend frei davon ist. Grauwerte eines dunklen Gebietes im IR-Kanal (z.B. Wasserflächen, Schatten) werden herangenommen und in anderen Kanälen verglichen: die Verschiebung der niedrigsten Histogrammwerte wird als Folge von Atmosphäreneinfluss interpretiert und als Korrekturbetrag *a* genommen. Dieser Betrag wird von sämtlichen Pixeln des jeweiligen Bandes subtrahiert.
Problem: nur Näherung; kann nur für Streuungsfehler verwendet werden.



Histogrammvergleich (Histogramm matching) [nach Albertz 2001]

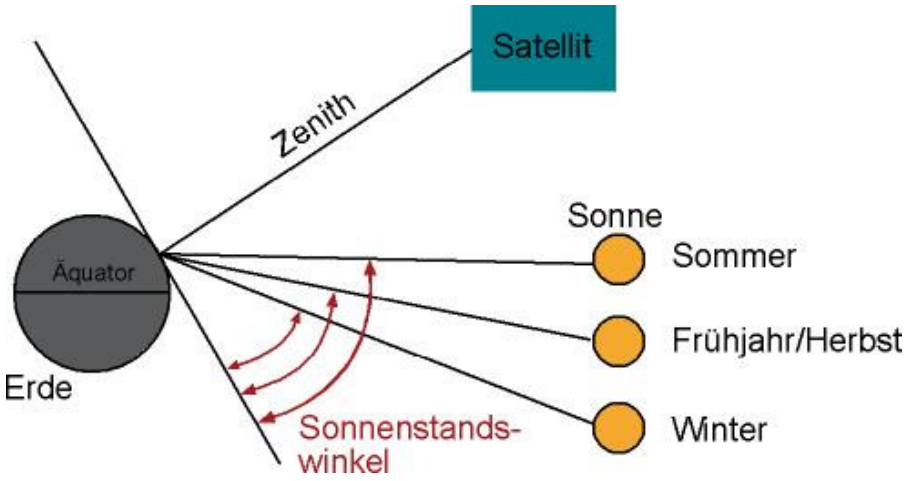
c. Regressionsmethode
Jeder zu korrigierende Kanal wird einzeln dem Infrarot- Kanal gegenübergestellt (s. Abb. unten) und eine Regressionsgerade durch den Punkthaufen gelegt. Die Gerade müsste ohne Atmosphäreneinfluss durch den Ursprung gehen. Der Wert bei dem sie die y-Achse schneidet entspricht dem Korrekturwert *a* für diesen Kanal, welcher von allen Grauwerten subtrahiert wird.



Regressionsmethode [nach Albertz 2001]

2. Sonnenstandskorrektur

Wird angewendet um saisonale oder tageszeitliche Unterschiede im Sonnenstand auszugleichen. Das ist wichtig bei Untersuchungen von Satellitenbildern mit unterschiedlichen Aufnahmezeitpunkten oder beim "Mosaicking", dem Zusammenfügen mehrerer Bilder zu einem Bild.



[nach Lillesand/Kiefer 1994]

3. Korrekturen von Sensorfehlern

Umfasst Korrekturen, die darauf gerichtet sind, sichtbare Fehler und Rauschen im Bild zu entfernen. Defekte im Datenbestand können auftreten als periodisch oder zufällig fehlende Zeilen (line dropouts), Streifen (line striping) oder fehlende Pixel.

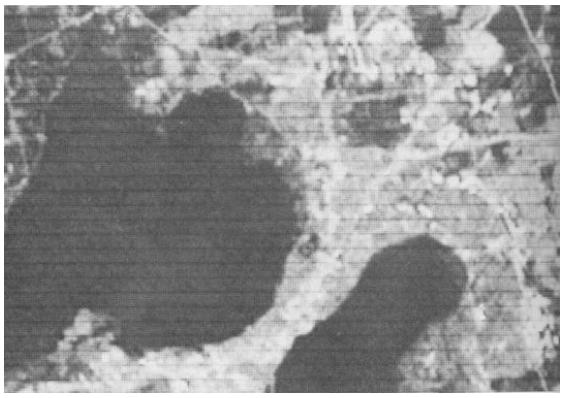
Zeilenausfall (line dropout)

Entsteht, wenn einer der Detektoren des Sensors ausfällt. Häufig v.a. beim Landsat MSS Sensor mit 6 Detektoren. Bei Ausfall eines Detektors entsteht jede 6. Zeile ein schwarzer Streifen (=Grauwert 0). Zur Restaurierung wird aus den umliegenden Pixeln ein künstlicher Grauwert berechnet.

Line dropout [Lillesand/Kiefer 1994]

Streifenfehler (Striping, Banding):

kommen dadurch zustande, dass einige Sensoren, obwohl sie vor dem Start sorgfältig kalibriert wurden, mit der Zeit degradieren. Das kann darin resultieren, dass Zeilen von solchen degradierten Sensoren unnatürlich höhere oder niedrigere Grauwerte haben als ihre Umgebung. Hier sind echte Daten vorhanden, sie müssen jedoch korrigiert werden. Dazu werden Algorithmen verwendet um inkorrekte Grauwerte anzupassen (genannt: Destriping); entweder werden solche Zeilen ersetzt mit Grauwerten, welche auf dem Durchschnitt angrenzender Pixel basieren, oder mit Werten, welche auf der Mittel- und Standardabweichung des Kanals basieren.



Streifenfehler [Lillesand/Kiefer 1994]

Rauschen (Noise):

Das sind Störpixel, die mit grob falschen Grauwerten belegt sind. Sie haben ihre Ursache z.B. in Fehlern bei der Datenübertragung zur Bodenstation. Auch hier wird zur Korrektur der mittlere Grauwert aus den Nachbarpixeln eingesetzt.

Rauschen [Lillesand/Kiefer 1994]

Achtung:
Alle diese radiometrischen Korrekturen stellen keine absolute Verbesserung der Grauwerte dar. In einigen Fällen werden auch neue Grauwerte "erfunden". Die Störungen werden nur insoweit beseitigt, dass die Daten einander angeglichen werden können und die Bildwirkung verbessert wird.

4.2.2 Geometrische Transformation

Luft- und Satellitenbilder werden in der Regel geometrisch korrigiert um Verzerrungen, die durch das Aufnahmesystem und das Geländerelief entstanden, zu eliminieren und die Daten in ein geodätisches Referenzsystem (Koordinatensystem) zu bringen. Hierfür werden Transformationsgleichungen herangezogen.
Die **Geokodierung** umfasst die Prozesse, welche eine geometrische Transformation der Originaldaten in ein gewähltes geodätisches Bezugssystem unter weitgehender Eliminierung der system- oder geländeabhängigen Verzerrungen erlauben.
Eine verwandte Technik ist die **Registrierung** (=image registration), die das Koordinatensystem des einen Bildes dem eines anderen Bildes anpasst. Sie wird z.B. für Beobachtungen der Veränderungen zwischen Satellitenbildern (= "change detection") verwendet.

2 Wege:

- Parametrische Verfahren:

  versuchen, die Aufnahmegeometrie im Raum mathematisch zu modellieren. Die Sensorgeometrie und die Raumlage (-bewegung) müssen bekannt sein (bei Luftbildern entspricht das der inneren und äußeren Orientierung). Zur Berechnung der äußeren Orientierung können entweder GPS/INS oder Paßpunkte verwendet werden.
  Paßpunkte werden benutzt um den Bezug zum Koordinatensystem herzustellen. Außerdem kann ein digitales Geländemodell mit einbezogen werden.

- Nicht-parametrische Verfahren:

  Die Aufnahmegeometrie wird nicht modelliert. Die unbekannten Koordinaten eines Pixels im Referenzsystem werden durch Interpolation zwischen benachbarten Paßpunkten berechnet. Zur Interpolation gibt es verschiedene mathematische Ansätze, z.B. Polynomverfahren, Warping.

Nutzung von Paßinformation:

<u>Paßpunkte:</u>

Das sind Punkte, die in den Bilddaten eindeutig identifizierbar sind und deren Koordinaten bekannt sind. Das können z.B. topographische Objekte, wie Kreuzungen, markante Geländepunkte, etc., sein. Die Koordinaten der Referenzpunkte können entweder bei einer entsprechend genauen Kartengrundlage direkt von der Karte abgelesen werden oder mittels GPS oder klassischer Vermessung ermittelt werden. Damit kann die geometrische Beziehung zwischen Bildpunkt und Geländepunkt hergestellt werden (z.B. x,y im Bild zu Rechts- und Hochwert in der Karte).
Die Koordinaten der Paßpunkte in den Bilddaten werden im allgemeinen durch interaktive Operationen am Monitor des Rechners bestimmt. Es werden aber auch automatische Verfahren angewandt.

Transformation der Koordinaten:

Mittels Transformationsgleichungen wird die Berechnung der x,y Koordinaten für jedes Pixel im neuen Bezugssystem von Bildpunkt zu Bildpunkt durchgeführt.

Die Transformationsgleichungen zur Interpolation zwischen zwei ebenen Punktfeldern lauten:

$$x' = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy$$

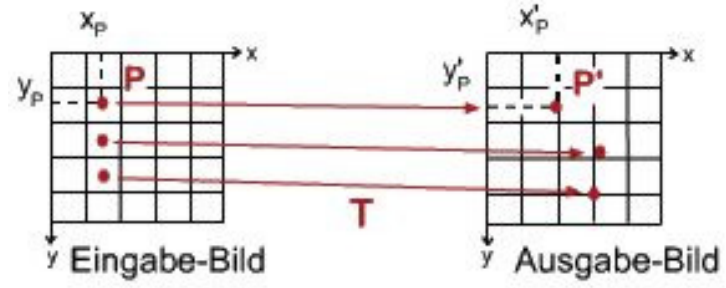
$$y' = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy$$

Daraus werden für jeden Passpunkt mit Bildkoordinaten x',y ' und Referenzkoordinaten x,y die Koeffizienten a und b berechnet.

Danach kann die eigentliche Entzerrung als direkte oder indirekte Transformation durchgeführt werden. Außerdem wird noch die radiometrische Information (=Grauwert) im Ausgabebild neu bestimmt.

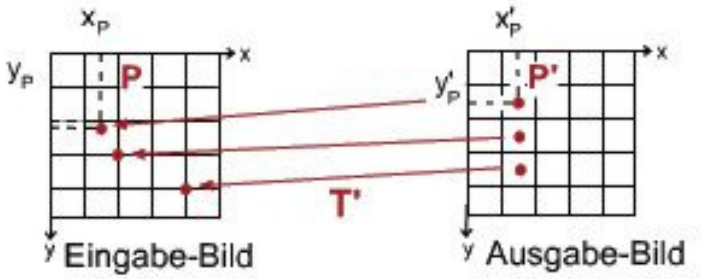
Es gibt 2 Wege der Transformation:

- Direkte Transformation:
berechnet für jedes Element des Eingabe-Bildes die Lage im Ausgabe-Bild und weist ihm den Grauwert aus dem Eingabe-Bild zu. Das entstehende unregelmäßige Bildraster muss nun noch durch geeignete Interpolation in eine regelmäßige Grauwertmatrix umgewandelt werden.



Schema der direkten Transformation

- Indirekte Transformation:
Hier werden für die Elemente des Ausgabe-Bildes die Grauwerte im Eingabe-Bild gesucht. Dadurch enthält man unmittelbar die entzerrte Bildmatrix, weshalb sich diese Arbeitsweise durchgesetzt hat.



Schema der indirekten Transformation

Resampling:

Durch die geometrische Transformation müssen die Daten des Eingabe-Bildes im Ausgabe-Bild neu geordnet werden. Dieser Vorgang wird als Resampling bezeichnet.

Jedem Pixel im Ausgabebild muss ein Grauwert zugewiesen werden. Da man in der Regel keine ganzzahligen Werte erhält, muss interpoliert werden um Grauwerte den Pixelmitten zuzuweisen. Dafür stehen mehrere Resampling-Methoden zur Verfügung, die im folgenden erläutert werden:
- Nächste Nachbarschaft (Nearest Neighbor)
- Bikubische Interpolation (Cubic Convolution)
- Bilineare Interpolation (Bilinear Interpolation)

Interpolationsverfahren:

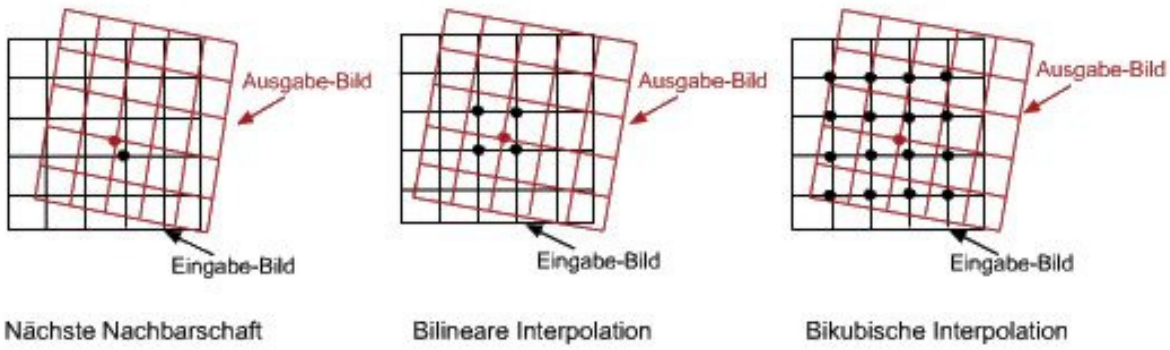
1. Verfahren der nächsten Nachbarschaft:
Es wird der Grauwert jenes Pixels im Eingabebild übernommen, welches den berechneten neuen Koordinaten x' und y ' am nächsten liegt. Die Originalgrauwerte bleiben daher unverändert, weshalb diese Methode bevorzugt angewandt wird, wenn eine nachfolgende Klassifikation der Daten erfolgt.
Das Verfahren benötigt am wenigsten CPU-Zeit, es kann aber zu einer Wiederholung von Grauwerten kommen und damit zu einem blockigen Aussehen des Bildes (vgl. Abbildung unten). Es führt zu einem hohen Genauigkeitsverlust [Bähr, Vögtle 1991].

Original **Nearest Neighbour** **Bikubische Interpolation**

Wiedergabe einer Bildszene bei verschiedenen Interpolationsmethoden:
Die Original-Bilddaten wurden durch eine geometrische Transformation um 30° gedreht [aus Albertz 2001].

2. Bilineare Interpolation:
Betrachtet die unmittelbare Nachbarschaft des Pixels (also die 4 Nachbarpixel) und interpoliert daraus den neuen Grauwert (gewichtete Mittel der Werte).
Das Verfahren ist etwas aufwendiger als das der nächsten Nachbarschaft und hat einen Glättungseffekt.

3. Bikubische Interpolation:
Hier werden die 16 Nachbarpixel des Bildpunktes interpoliert um den Grauwert für diese Pixel zu bestimmen. Diese Methode ist die komplizierteste und benötigt die längsten Rechenzeiten, besitzt dafür aber nicht das blockige Aussehen von Nearest Neighbour-Bildern oder eine extreme Glättung (bilineare Methode). Allerdings geht Information von hohen Frequenzen verloren, da der Interpolator im Grunde ein Tiefpassfilter ist.



Resampling-Methoden:
Die schwarzen Punkte stellen die Pixelmittelpunkte im Eingabe-Bild dar, die zur Interpolation des neuen Bildpunktes (=roter Punkt) im Ausgabe-Bild bei den verschiedenen Methoden verwendet werden.

4.3 Bildverarbeitung zur Visualisierung

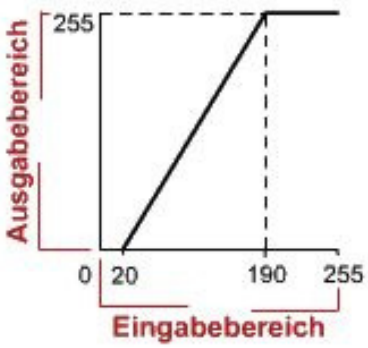
Der Sinn dieser Methoden liegt darin, die visuelle Erscheinung der Bilder zu verbessern, so dass die Interpretierbarkeit verbessert wird. Dabei werden die Grauwerte im Bild verändert. Das bedeutet, dass diese visuell verbesserten Bilder nicht für weitere Analyse (wie z.B. zur Klassifizierung) genommen werden sollten.

Es gibt mehrere Wege, die visuelle Interpretierbarkeit eines Bildes zu verbessern:

- 1. Veränderung des Kontrasts
  Methoden:
  Kontraststreckung, Histogrammverebnung, Äquidensitenherstellung
- 2. Konversion eines monochromatischen Bildes in ein Farbbild
  Methoden:
  Äquidensitenherstellung, Pseudofarben-Transformation
- 3. RGB zu IHS-Transformation
- 4. Pan-sharpening

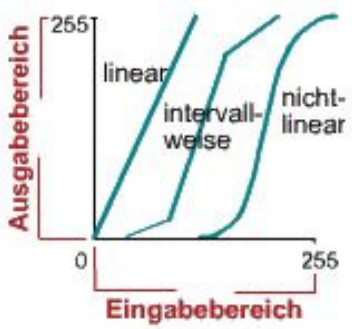
4.3.1 Kontraststreckung (contrast stretch)

Sie hat das Ziel, den Kontrast eines Bildes zu verbessern. Hierbei werden die Originalgrauwerte eines Kanals, DN_{min} bis DN_{max}, auf den vollen Ausgabebereich des Displays (d.h. 0-255 bei 8 bit) gespreizt, wobei DN_{min} der kleinste Grauwert im Bild sei und DN_{max} der größte. Alle Grauwerte dazwischen werden proportional dazu ebenfalls gestreckt. DN_{min} wird also auf 0 gesetzt, DN_{max} auf 255. Hierfür werden LUTs verwendet.

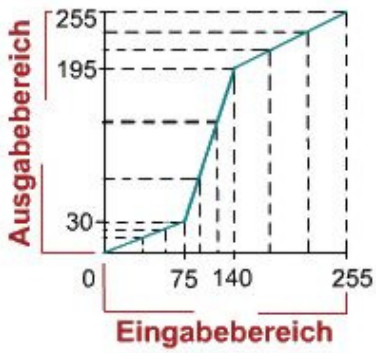


Look-Up Table mit linearer Kontrastveränderungs-Funktion. Die Grauwerte des Eingabebildes von 20 (=DN_{min}) bis 190 (=DN_{max}) werden auf den vollen Ausgabebereich 0 bis 255 gestreckt. Die Werte <20 und >190 werden auf 0 gesetzt.

Die Funktion der Kontraststreckung muss nicht immer linear verlaufen, wie in der vorherigen Abbildung gezeigt. Sie kann ebenso nicht-linear oder intervallweise (linear piecewise) verlaufen (siehe Abb rechts).



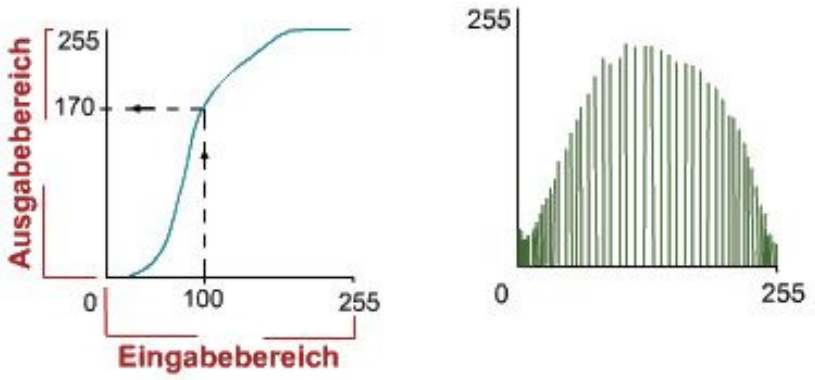
Die Abbildung rechts zeigt den LUT einer intervallweisen Kontraststreckung. Die Grauwerte des Eingabebereichs 75-140 werden in der Ausgabe auf 30-195 gestreckt, während die Bereiche darüber und darunter gestaucht werden. Das bewirkt eine Kontrastvergrößerung in dem gestreckten Bereich.



Die meisten Bildverarbeitungsprogramme erlauben es, eine eigene Übertragungsfunktion aufzustellen um mittels sogenannter "breakpoints" gestreckte und gestauchte Bereich selbst zu definieren.

4.3.2 Histogrammverebnung (histogram equalisation)

Das Histogramm eines Kanals wird so verändert, dass jede Histogrammklasse eine ungefähr gleiche Anzahl von Pixeln hat. Die Grauwerte des Ausgabebildes sollen also über die gesamte Grauwertskala gleich verteilt sein.
Ein Histogramm muss dabei zuerst in ein kumulatives Histogramm umgerechnet werden (siehe Kapitel 5.1.2). Die Teile des Bildes, die einen größeren Anteil an der Gesamtfläche haben, werden feiner dargestellt, die weniger gewichtigen Teile grober.
Die Pixel, die ursprünglich denselben Grauwert hatten, können allerdings nicht verschieden gedehnten Werten zugeteilt werden. Idealerweise entsteht ein einheitliches Histogramm und ein kontrastreiches Bild.

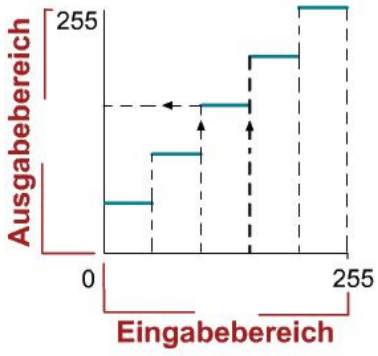


Histogrammverebnung: Graphische Darstellung einer Histogrammverebnung.

Die linke Abbildung zeigt die Funktion, die rechte das resultierende Histogramm: Die Zonen mit großer Häufigkeit werden gestreckt, die Zonen mit geringer Häufigkeit komprimiert (Kopf und Fuss des Histogramms).

4.3.3 Density Slicing (Äquidensitenherstellung)

Es wird eine treppenförmige Übertragungsfunktion angewandt, durch die Grauwertintervalle zu einem einzigen Grauwert zusammengefasst werden. Dabei werden Äquidensiten hergestellt, das sind Linien gleicher Grauwerte. Ein durchgehender Grauwertbereich (z.B. 0,1,2,3,4,5,6) heißt auf englisch "slice".



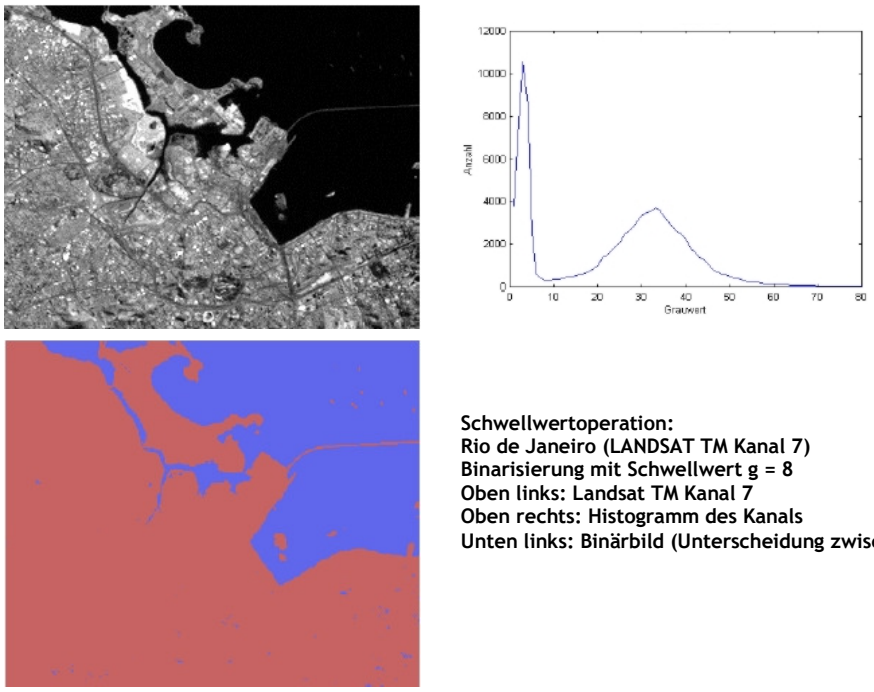
Density Slicing: Graphische Darstellung eines Look-Up Tables mit einer Übertragungsfunktion für die Äquidensiten.

Es werden 5 Grauwertstufen erzeugt. Der Grauwertbereich zwischen 0-255 wird zerlegt in 0-50 (=schwarz), 51-100 (=dunkelgrau), 101-150 (=mittelgrau), 151-200 (=hellgrau) und 201-255 (=weiß).

Densitiy Slicing ist auch eine Methode, durch die man ein panchromatisches Bild in eine Pseudofarbdarstellung überführen kann. Es wird dabei ein durchgehender Grauwertbereich auf einen Punkt im RGB-Farbwürfel abgebildet.

Im Ausgabe-Bild werden dadurch verschiedene aber intern homogene Gebiete hervorgehoben. Es kann so eine erhebliche Kontrastverbesserung erzielt werden. Allerdings kommt es zu Detailverlust, da ja Grauwerte zusammengefaßt werden.

Es können durch diese Methode auch Binärbilder erzeugt werden, indem man nur 2 Grauwertbereiche für die Zerlegung verwendet und ihnen in der Ausgabe 0 und 1 zuordnet. Näheres zur Binärbildverarbeitung siehe Kapitel 5.7.



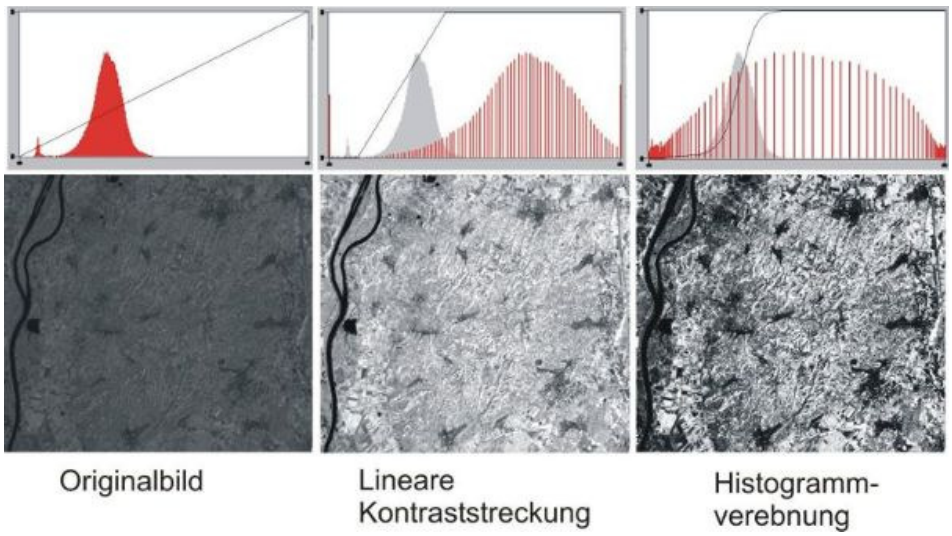
Schwellwertoperation:
Rio de Janeiro (LANDSAT TM Kanal 7)
Binarisierung mit Schwellwert g = 8
Oben links: Landsat TM Kanal 7
Oben rechts: Histogramm des Kanals
Unten links: Binärbild (Unterscheidung zwischen Land und Wasser)

4.3.4 Pseudofarben-Transformation

Im Unterschied zur Äquidensitenherstellung wird hier jedem einzelnen Grauwert ein Farbwert zugewiesen, wodurch ein Farbbild entsteht.

4.3.5 Vergleich der Methoden

Die Abbildung zeigt das Gebiet des Kaiserstuhls: Das Originalbild ist kontrastarm und schwer interpretierbar, die möglichen 256 Grauwerte werden nicht ausgenutzt. Die beiden anderen Bilder zeigen Bildverbesserungs-Methoden und die zugehörigen Histogramme mit der Übertragungsfunktion.

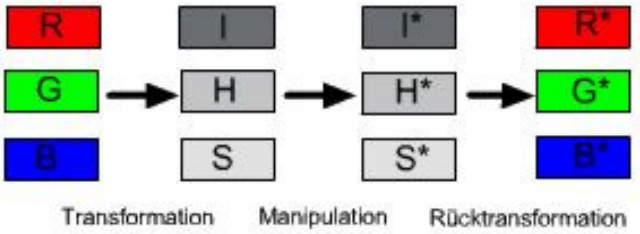


Die Abbildung zeigt das Gebiet des Kaiserstuhls: das Originalbild links, nach einer Kontraststreckung (Mitte) und nach einer Histogrammverebnung (rechts). Über den Bildern die zugehörigen Histogramme mit der Übertragungsfunktion. Das Histogramm des Originalbildes ist in Grau hinterlegt, das Histogramm des bearbeiteten Bildes ist rot dargestellt.

4.3.6 IHS-Transformation

Digitale Bilder werden üblicherweise mit den Primärfarben Rot, Grün und Blau dargestellt. Eine weitere Art der Farbdarstellung ist IHS (siehe Exkurs "Farbmodelle").

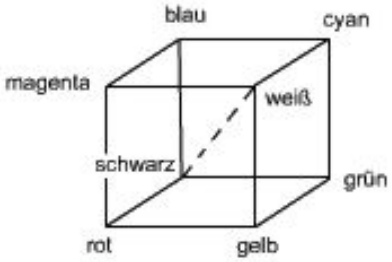
Transformationen vom RGB-Farbraum in den IHS-Farbraum werden zur Bildverbesserung benutzt. Die originalen RGB-Komponenten werden in ihre korrespondierenden IHS-Komponenten zerlegt. Diese werden dann mit den bereits besprochenen Methoden wie Kontraststreckung etc. manipuliert, um die gewünschten Charakteristika des Bildes herauszuarbeiten. Ein Vorteil dabei ist, dass die IHS-Komponenten unabhängig voneinander manipuliert werden können. So kann z.B. die "Intensität" durch Kontraststreckung verbessert werden, der Farbton und die Sättigung dagegen werden nicht verändert. Zum Schluss werden die modifizierten IHS-Komponenten wieder in den RGB-Farbraum zurücktransformiert.



Exkurs Farbmodelle:

1. RGB-Farbmodell

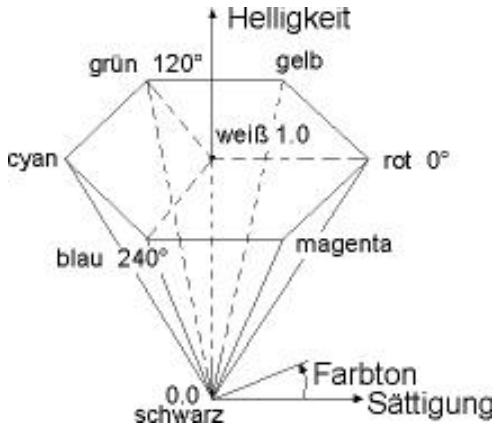
Eine Farbe wird beschrieben durch den Anteil an Farben der Additiven Farbmischung. So werden üblicherweise Farben in der digitalen Bildverarbeitung und auf Monitoren wiedergegeben.



RGB-Modell

2. IHS-Farmodell

Hier wird eine Farbe durch ihre Helligkeit (=Intensity), Farbton (=Hue) und Sättigung (=Saturation) beschrieben. Der Farbton gibt die Farbe wieder, die wir sehen (z.B. rot, grün etc.). Der Grad der Reinheit wird bestimmt durch die Sättigung und liegt zwischen 0 und 1. Eine reine Farbe ist 100% gesättigt (also 1). Die Intensität beschreibt die Helligkeit und reicht von 0.0 (schwarz) bis 1.0 (weiß).
Die Abbildung rechts zeigt eine geometrische Repräsentation des Modells: Der Farbton wird dargestellt durch die oberen Ecken des 6-eckigen Kegels, rot liegt bei 0°, grün bei 120° und blau bei 240°. Außen liegen die reinen Farben. Zusatz von weiß produziert weniger gesättigte, blassere Farben.



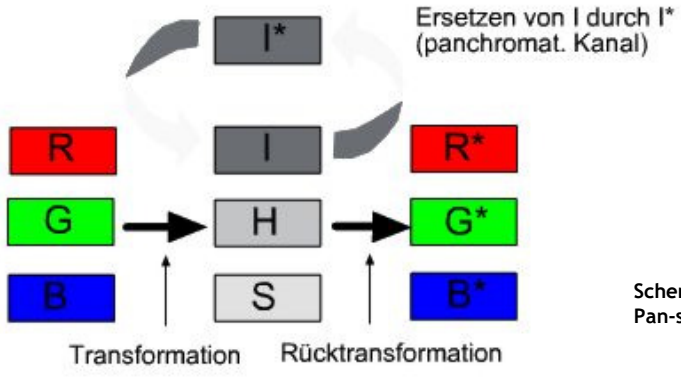
IHS-Modell

Dieses Farbmodell stellt eine bessere Farbrepräsentation dar von Farben, wie sie auch das menschliche Auge sieht, als das RGB-Farbmodell .

4.3.7 Pan-sharpening

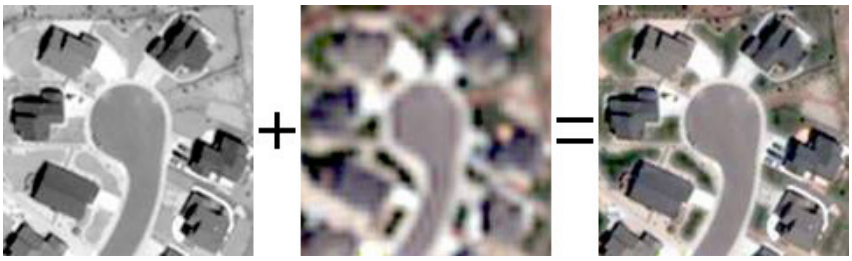
Hierbei wird ein panchromatischer Kanal mit hoher Auflösung mit multispektralen Kanälen verschnitten, welche eine geringere Auflösung haben. Dadurch entsteht ein Farbbild, das an Schärfe gewonnen hat. Die Farbinformation wird dadurch jedoch nicht verbessert.

Es gibt verschiedene Algorithmen für diese Operation. Eine davon verwendet die zuvor beschriebene IHS-Transformation: Der panchromatische Kanal wird in die Intensitäts-Komponente (I) umgewandelt, da I die Helligkeitsverteilung im Bild beschreibt und dem panchromatischen Kanal verwandt ist. Die Multispektralkanäle werden in Farbton (H) und Sättigung (S) transformiert. Die Transformation funktioniert nur dann, wenn der I-Kanal die Helligkeitsverteilung über das gesamte sichtbare Spektrum repräsentiert, wie das beim panchromatischen Kanal der Fall ist.



Schema des Ablaufs des Pan-sharpening

Beispiel: Ikonos besitzt 4 Multispektralkanäle mit 4 m Auflösung und einen panchromatischen Kanal mit 1 m Auflösung. Werden diese Datensätze kombiniert, ergibt das 4 Kanäle mit 1 m Auflösung (siehe Abbildung).



Quelle: Geomatica®:
http://www.pcigeomatics.com/product_ind/pansharpening.html

4.4 Mehrkanal-Operationen

Im vorherigen Kapitel wurden Bildverarbeitungsverfahren beschrieben, die jeweils nur einen Kanal des Bildes bearbeiten. Hier soll nun erläutert werden, wie man mehrere Bildkanäle miteinander verknüpft. Es können nicht nur multispektrale Datensätze verwendet werden, sondern auch multitemporale Datensätze (= Bilder, die ein Gebiet zu unterschiedlichen Zeiten zeigen; verwendet für Zeitreihenanalysen).
Die Verfahren dienen u.a. zur visuellen Weiterverarbeitung, Analyse von Zeitreihen, Bildung von Vegetationsindices und Hauptkomponententransformation.

Folgende Verfahren werden beschrieben:

- Arithmetische Pixeloperationen
- Ratiobildung
- Hauptkomponentenanalyse

4.4.1 Arithmetische Pixeloperationen

Darunter fallen Operationen der Addition, Subtraktion, Multiplikation und Division von 2 oder mehr Bildern des gleichen Gebietes. Diese Bilder müssen zueinander registriert sein, damit die Pixel lagegleich sind.

Die Operationen werden auf Pixelbasis durchgeführt und lassen einen neuen Bilddatensatz, ein Komposit, entstehen, das einen bestimmten Bildinhalt verstärkt wiedergibt.
Arithmetische Operationen sind deshalb möglich, weil Sensoren die variierenden Spektralcharakteristika unterschiedlicher Oberflächenmaterialien in verschiedenen Kanälen aufnehmen, welche deshalb unterschiedliche statistische Merkmale haben.

Addition:
Die Addition zweier Datensätze ist im Grunde eine Durchschnittsbildung. Wenn 2 Kanäle mit einem dynamischen Bereich von 0-255 addiert werden, hätte das resultierende Bild Werte von 0-510. Für eine 8-bit Ausgabe muss also auf 0-255 skaliert werden, indem man durch die Anzahl der verwendeten Kanäle teilt:

$$I_{sum}(x,y) = \frac{I_1(x,y) + I_2(x,y)}{2}$$

wobei $I_1(x,y)$= Bild 1 und $I_2(x,y)$= Bild 2. Das Ergebnis wird auf eine ganze Zahl gerundet.

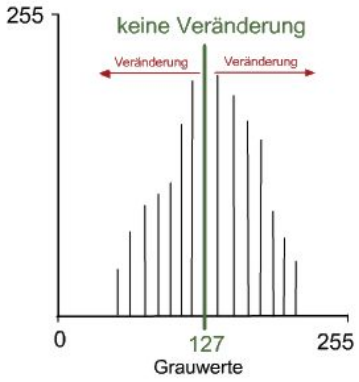
Durch Addition kann Rauschen im Bild (Noise) reduzieren werden. Außerdem kann diese Methode dazu benutzt werden, graphische Informationen in Form von Gittermustern, Text etc, in ein digitales Bild einzublenden [Haberäcker 1991] .

Subtraktion:
Die Subtraktion von Bildern wird für multitemporale Datensätze benutzt. Damit können die Unterschiede zwischen Kanälen sichtbar gemacht werden, weshalb diese Methode im Bereich der Änderungserkennung (engl. *change detection*) Verwendung findet.

Werden 2 Kanäle mit den Grauwerten 0-255 voneinander abgezogen, so muss eine Konstante, hier 255, dazu addiert werden, um die negativen Zahlen zu entfernen. Außerdem muss durch 2 geteilt werden, um den Display-Bereich von 0-255 zu erreichen:

$$I_{diff}(x,y) = \frac{(255 + I_1(x,y) - I_2(x,y))}{2}$$

Das entstandene Komposit verdeutlicht das Ausmaß und die Richtung der Veränderungen zwischen den beiden Zeitpunkten. Gebiet, in denen keine Veränderungen auftreten, haben den Grauwert 127 (=mittelgrau im Bild); die Gebiete, in denen Veränderungen stattfanden, haben Grauwerte höher oder niedriger als 127, je nach Richtung der Veränderung.



Multiplikation:
Diese Verfahren ist nützlich, wenn nur eine bestimmte Region im Bild betrachtet werden soll. Wenn z.B. Ein Bild aus 2 völlig unterschiedlich reflektierenden Regionen besteht, z.B. Land und Meer und nur das Meer betrachtet werden soll, lenken die Variationen in der Reflexion des Landes von den subtileren im Wasser ab. Also wird das Gebiet "Land" eliminiert. Das geschieht durch eine Maskierung (engl. *masking*).

Vorgehen:

- 1. Definieren der zu maskierenden Region:
  Das kann zum einen durch Einführung eines geeigneten Schwellenwertes geschehen oder durch die Verwendung eines Binärbildes. Bei beiden Methoden entsteht ein Datensatz mit Grauwerten von 0 und 1, wobei die Pixel unter der Maske 0 sind, die anderen 1.

- 2. Masking:
  Das Originalbild wird nun Pixel für Pixel mit der Maske multipliziert. Wird ein Grauwert mit 1 multipliziert, bleiben die Originalgrauwerte gleich. Wird ein Grauwert mit 0 multipliziert, wird er auf 0 gesetzt, also maskiert.

Division:
Dies ist die am meisten benutzte arithmetische Operation, weshalb sie auch im nächsten Kapitel 5.4.2 gesondert behandelt wird und hier nur kurz vorgestellt wird.

Diese Methode wird auch Ratiobildung (engl. *ratioing*) genannt. Sie wird dazu benutzt, Merkmale eines Bildes durch Unterschiede im Spektralverhalten herauszustellen. Indem man 2 verschiedene Kanäle zueinander ins Verhältnis setzt, verstärkt man im resultierenden Bild Variationen in den spektralen Reflexionskurven. Ein bekanntes Beispiel ist die Erstellung von Vegetationsindices.

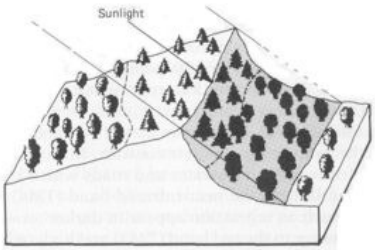
Die Ratiobildung findet auch Anwendung in der Beseitigung von Schatten und topographischen Effekten.

4.4.2 Ratiobildung

Ratiobildung ist ein weitverbreitetes Verfahren und wird gerne dazu benutzt, Indices herzustellen.

Verwendung:

1. Unerwünschte Einflüsse der durch das Geländerelief verursachten Beleuchtungsunterschiede können reduziert werden, da der Einfluss zwischen den einzelnen Kanälen stark korreliert ist. Obwohl also die Reflexionswerte der Objekte variieren, je nachdem ob sie auf der sonnenzu- oder -abgewandten Seite stehen, wird das Verhältnis dieser Grauwerte beinahe gleich sein

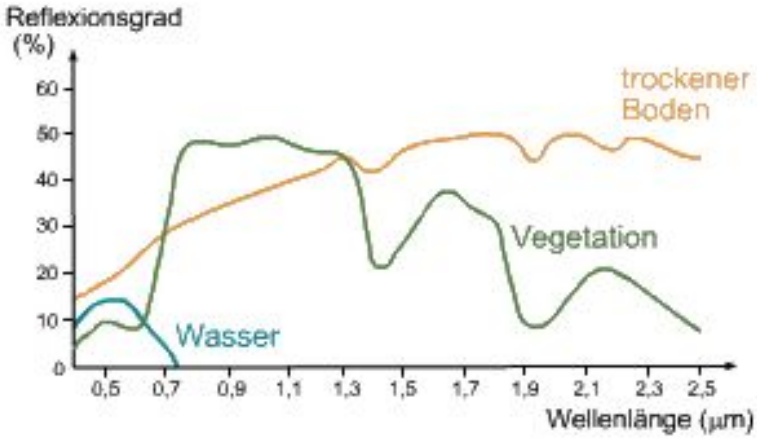


Land Cover/Illumination	Digital Number		
	Band A	Band B	Ratio (Band A/Band B)
Deciduous			
Sunlit	48	50	0.96
Shadow	18	19	0.95
Coniferous			
Sunlit	31	45	0.69
Shadow	11	16	0.69

Reduction of scene illumination effects through spectral ratioing. (Adapted from [165].)

Effekt von Ratiobildung: Verminderung der Effekte von Beleuchtungsunterschieden im Gelände [Lillesand, Kiefer 1994]

2. Spektraleigenschaften verschiedener Oberflächenmaterialien (z.B. Vegetation, Minerale, Gesteine, etc.) können im neuen Datensatz hervorgehoben werden. Das Verfahren ist besonders da erfolgreich, wo Objekte mit charakteristischen Reflexionseigenschaften betrachtet werden. Die Abbildung unten zeigt spektrale Reflexionskurven verschiedener Materialien:



Reflexionskurven verschiedener Materialien der Erdoberfläche

Charakteristisch ist der Anstieg der Vegetationskurve bei 0.68 µm, auch "red edge" genannt. Im nahen Infrarot hat Vegetation ein Reflexionsmaximum, im sichtbaren roten Licht ein Minimum. Setzt man also die Grauwerte im NIR und im roten Bereich des elektromagnetischen Spektrums zueinander ins Verhältnis, so wird der Wert für Vegetation höher sein als der für Wasser oder Boden. So lässt sich z.B. Vegetation von Nicht-Vegetation unterscheiden. Einen solchen Index nennt man **Vegetationsindex**.

Einfacher Vegetationsindex

Der **einfache Vegetationsindex** lautet

$$VI = \frac{NIR}{R},$$

bezogen auf ETM+ ist das. $\frac{Kanal4}{Kanal3}$

Der Index liegt bei Vegetation höher als bei Boden, da der Unterschied zwischen NIR und Rot bei Boden nicht so groß ist, daher also der Index kleiner ist.
Da die Werte des Index nicht bei 0-255 liegen, müssen sie zur Darstellung erst noch skaliert werden.

NDVI

Der bekannteste und am weitesten verbreitete Index ist der NDVI (=Normalized Difference Vegetation Index).

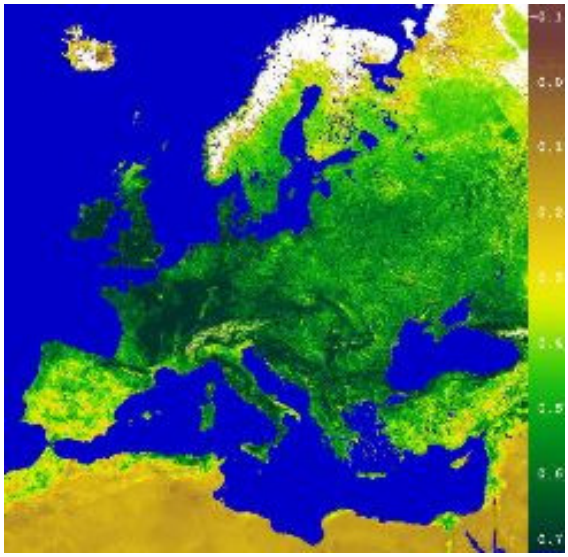
$$NDVI = \frac{NIR - R}{NIR + R}$$

für Landsat ETM+ Kanäle lautet er: $\frac{Kanal4 - Kanal3}{Kanal4 + Kanal3}$

Die Normierung führt dazu, dass der NDVI nur Werte zwischen -1 und +1 hat. Es werden dadurch Einflüsse der Geländeneigung und Beleuchtungsverhältnisse kompensiert. Auch hier muss der Datensatz zur Wiedergabe wieder auf 0-255 skaliert werden. Das wird erreicht indem man zuerst 1 addiert (Grauwertbereich somit 0-2) und das Resultat mit 127 multipliziert.

Verwendung:

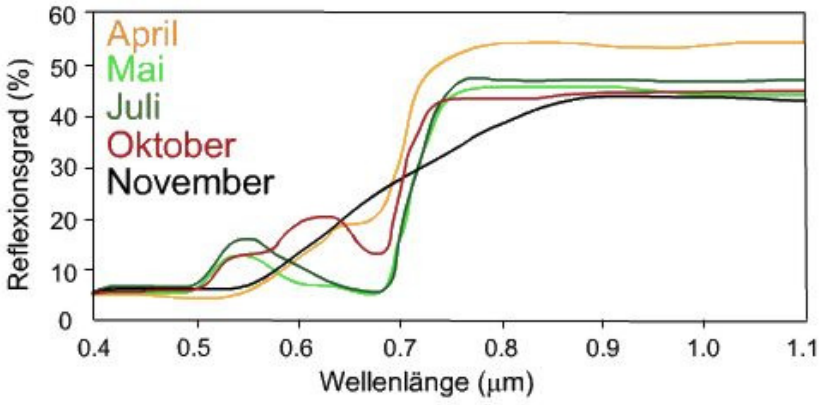
Der NDVI wird als Maß für die Vegetationsbedeckung einer Fläche verwendet. Er zeigt die Vitalität der Vegetation. In der Regel wird der Datensatz farbcodiert (density-sliced), wobei die Farbskala in der Regel von braun (keine Vegetation) über gelb bis dunkelgrün (gesunde Vegetation) reicht.



NDVI Europa, Mai 1997; Quelle: DLR

Jede Vegetationsart hat eine charakteristische Reflexionskurve. Der Reflexions- und Absorptionsgrad hängt von der Physiologie, den Pigmenten und der Orientierung der Pflanze zum Sensor ab. Diese Reflexionskurve schlägt sich im NDVI wieder, weshalb er auch zur **Klassifizierung von Vegetationsarten** herangenommen werden kann. Der NDVI für Vegetation liegt in der Regel im Bereich von -0.1 bis +0.7.

Die charakteristischen Reflexionseigenschaften der Pflanze ändern sich im jahreszeitlichen Verlauf, daher kann man saisonale Unterschiede im NDVI erkennen. Das ist wichtig für die Beobachtung der **Phänologie**. Die Abbildung zeigt den jahreszeitlichen Verlauf der Reflexion eines Eichenblattes [nach Gates 1970].



Jahreszeitlicher Verlauf der Reflexion eines Eichenblattes [nach Gates 1970].

Die Reflexionsunterschiede in NIR und Rot sind im Mai und Juli am höchsten. Im April, Oktober und November ist der Unterschied dagegen nicht so gut ausgeprägt, daher sind die NDVI-Werte geringer.

Bei **Krankheit und Stress** ändert sich der Wasser- und Pigmenthaushalt der Pflanze. Die Reflexion im IR-Bereich wird geringer, daher wird eine kranke Pflanze auch einen niedrigeren NDVI besitzen als eine gesunde.
Der NDVI dient auch als Basis zur Berechnung der Biomasse (=Volumen an pflanzlichem Material), sowie des Blattflächenindexes (engl. *leaf are index*).
Allerdings muss man beachten, dass bei Gebieten mit verstreuter Vegetation, wo Boden für den Sensor "sichtbar" ist, der Vegetationsindex niedriger ist, als er ohne den Einfluss des Bodens wäre.
Neben dem NDVI gibt es noch eine Vielzahl anderer Vegetationsindices, die verschiedene Kanäle unterschiedlich kombinieren und so versuchen, die Einflüsse des Bodens oder der Atmopsphäre auf den Index zu minimieren. Einige sollen hier exemplarisch erwähnt werden. Zur Vertiefung siehe: Bannari et al. 1995.

Transformierter NDVI:
0.5 wird zum NDVI dazu addiert um negative Werte zu vermeiden.

$$\sqrt{\frac{IR-R}{IR+R}+0{,}5}$$

PVI (Perpendicular Vegetation Index = lotrechter VI):
Versucht, Einflüsse des Bodens zu reduzieren.

$$PVI = \sqrt{(p_{soil}-p_{veg})_R^2+(p_{soil}-p_{veg})_{NIR}^2}$$

Grüner Vegetationsindex (GVI):
$$-0{,}29 \times (G) - 0{,}56 \times (R) + 0{,}60 \times (IR) + 0{,}49 \times (IR)$$

ARVI (Atmospherically Resistant Vegetation Index):
$$ARVI = \frac{(NIR - RB)}{(NIR + RB)}$$
RB = Differenz von rotem und blauem Kanal

Neben den Vegetationsindices gibt es natürlich noch Indices, die **Mineralien und Gesteine** hervorheben, da Mineralien ebenfalls charakteristische Reflexionskurven haben. Sie finden Anwendung in der Geologie. Einige seien hier erwähnt (nur gültig für TM-Kanäle):

Tonminerale: TM 5 / TM 7
Eisenminerale: TM 5 / TM 4
Eisenhydroxyde: TM 5 / TM 3
Eisenoxide: TM 3 / TM 1

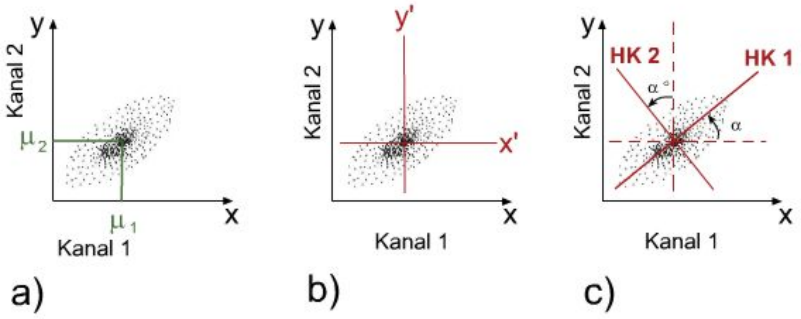
4.4.3 Hauptkomponentenanalyse

Die Bilddaten einer multispektralen Szene zeigen oft große Ähnlichkeit zwischen den Kanälen. Vor allem in benachbarten Spektralbereichen ist die Korrelation oft hoch. Das bedeutet, dass überschüssige, redundante Information in einer Szene vorhanden ist. Bei manchen Verarbeitungen ist es daher nützlich, die wesentlichen Informationen zu konzentrieren, damit die Spektralunterschiede zu verstärken und die Interpretierbarkeit zu verbessern. Dies führt außerdem zu einer Datenreduktion.

Eine Methode, dies zu erreichen, ist die Hauptkomponentenanalyse (engl. *Principal Component Analysis; PCA*), auch HKA genannt. Sie leitet aus den vorhandenen Bilddaten neue Datensätze ab, die die unkorrelierten Spektralunterschiede beinhalten.
Die mathematische Ableitung soll hier nicht ausführlich erklärt werden, sondern nur die wesentlichen Merkmale der HKA. Weiterführende Literatur: Kraus (1990), Haberäcker (1991).

Verwendung:
- Klassifikation (nur wenn Originalgrauwerte nicht erhalten werden müssen)
- Datenreduktion
- Erzeugung eines synthetischen Farbbildes

Im Folgenden soll die Hauptkomponententransformation annhand eines 2-kanaligen Bildes beschrieben werden. Die Transformation der Originalgrauwerte in neue Grauwerte erfolgt daher in einem 2-dimensionalen Merkmalsraum. Bei einer vollständigen Transformation besitzt der Merkmalsraum so viele Dimensionen wie Kanäle, z.B. bei Landsat TM mit 7 Kanälen 7 Dimensionen.



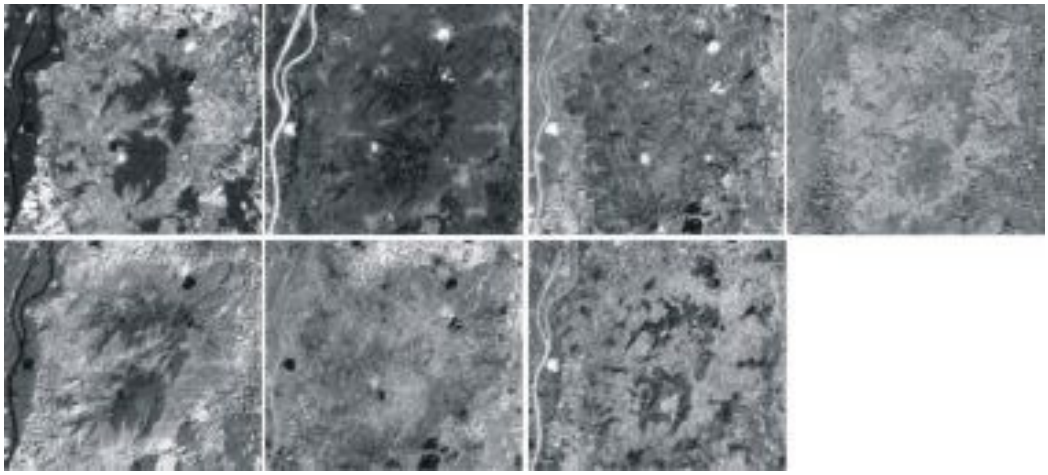
Die Grauwertverteilung **(Abb. a)** zeigt eine elliptische Form, die Daten sind also hoch korreliert. Die Eigenvektoren der Varianz-Covarianz-Matrix definieren dabei die Richtung der Halbachsen dieses Ellipsoids, die Eigenwerte ihre Länge. Aus der Covarianz-Matrix werden die Eigenvektoren und Eigenwerte berechnet.

Um die Hauptkomponenten (HK) zu ermitteln, wird ein neues Koordinatensystem x' und y' geschaffen, indem der Ursprung an die Mittelwerte μ_1 und μ_2 der beiden Kanäle versetzt wird. Die Originaldaten müssen also von x nach x' und y nach y' transformiert werden **(Abb. b)**.
Das neue Koordinatensystem wird dann um seinen Ursprung mit dem Winkel α rotiert. Damit ist seine x-Achse entlang der größten Varianz der Daten ausgerichtet **(Abb. c)**. Diese Achse bildet die erste Hauptkomponente und hat die Richtung der größten Streuung. Die senkrecht zu ihr stehende Achse ist die zweite Hauptkomponente.
Bei n-Dimensionen entstehen natürlich n-dimensionale Punktwolken mit entsprechender Anzahl von Raumachsen, die senkrecht zur 1. HK liegen.

Um die Originaldaten von den x'- und y'-Achsen auf die Hauptkomponenten-Achsen zu transformieren, müssen Koeffizienten benutzt werden, die aus der Covarianz-Matrix des Original-Datensatzes stammen. Als Ergebnis der Transformation ergeben sich die Matrix der Eigenwerte und der Eigenvektoren (geben Auskunft über die Korrelation zwischen den HK und den Originalkanälen), sowie die Varianz in jeder HK.

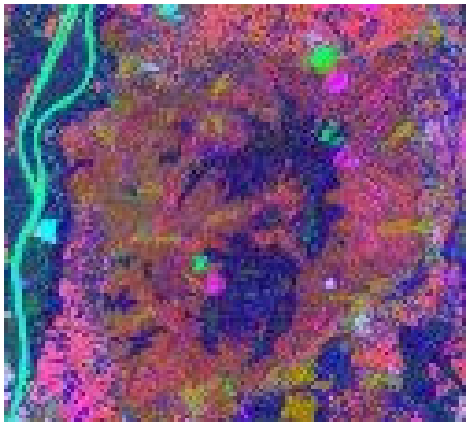
Die HKA versucht also Hauptkomponenten zu finden, mit deren Hilfe sich die gesamte Variation der ursprünglichen Spektralunterschiede reproduzieren läßt. Mit steigendem Ordnungsgrad der HK dominiert zunehmend das Rauschen im Bild (also immer geringere Unterschiede zwischen den Kanälen). Während die erste HK noch in der Regel 85% der Varianz beinhaltet, sind es bei der 2. HK nur noch 10 %. Erfahrungsgemäß stellen die ersten drei bis vier HK's eine gute Grundlage zur Interpretation dar.

Beispiel:
Die Abbildung unten zeigt die 7 Hauptkomponenten einer TM Szene (Gebiet: Kaiserstuhl, 2000). Das linke obere Bild ist die 1. HK mit dem größten Informationsgehalt. Der Informationsgehalt nimmt mit zunehmender Ordnungszahl der HK ab.



Hauptkomponenten einer TM-Szene (Kaiserstuhl, 2000):
Obere Reihe von links nach rechts: 1. bis 4. HK
Untere Reihe von links nach rechts: 5. bis 7. HK
Das linke obere Bild ist die 1. HK mit dem größten Informationsgehalt. Der Informationsgehalt nimmt mit zunehmender Ordnungszahl der HK ab. Allerdings besitzt jede Hauptkomponenten auch Merkmale, die andere HK nicht besitzen.

Der dazugehörige Report (siehe nächste Seite) zeigt die Covarianz-Matrix der Ausgangskanäle, die Eigenwerte und -vektoren, sowie die prozentuale Varianz der HK. Daraus geht hervor, dass die 1. HK 76 % der Spektralunterschiede beinhaltet, die 2. HK noch 16 % und die 3. HK nur noch 4 %. Die ersten drei HK zeigen also 96 % der Information des Originalbildes. Diese 3 Kanäle können zu einem RGB-Bild zusammengestellt werden (1. HK=Rot, 2. HK=Grün, 3. HK=Blau), das eine höhere Informationsdichte besitzt als jede andere aus den Originalkanälen erzeugte Farbkombination.



RGB- Farbbild der TM-Szene Kaiserstuhl
mit
1. HK = R,
2. HK = G
3. HK = B

Report

Kanal	Mittelwert	Standardabweichung
1	82.9998	13.5013
2	71.1263	16.5626
3	65.9486	26.0211
4	99.7704	21.8717
5	106.8032	29.7947
6	131.0970	6.0998
7	66.9065	30.5922

Covarianz-Matrix für Input-Kanäle

	1	2	3	4	5	6	7
1	182.284						
2	214.696	274.319					
3	326.500	419.511	677.098				
4	-81.424	-75.748	-156.038	478.372			
5	268.214	384.893	628.888	75.152	887.725		
6	34.727	50.371	87.457	-27.637	99.480	37.208	
7	316.448	429.139	714.350	-126.215	813.230	111.947	935.884

Eigenwerte

	Eigenwert	Standardabweichung	% Varianz
HK 1	2664.8442	51.6621	76.73 %
HK 2	565.2491	23.7750	16.28 %
HK 3	151.8291	12.3219	4.37 %
HK 4	51.0864	7.1477	1.47 %
HK 5	22.9648	4.7922	0.66 %
HK 6	12.2068	3.4938	0.35 %
Hk 7	4.7058	2.1693	0.14 %

Eigenvektoren

	HK 1	HK 2	HK 3	HK 4	HK 5	HK 6	HK 7
Kanal 1	0.22381	0.29811	0.48414	-0.06898	0.5342	0.07027	0.57584
Kanal 2	0.12562	0.08417	0.17654	-0.88709	-0.39777	0.02112	0.01934
Kanal 3	0.46504	0.44323	0.43252	0.32512	-0.31881	-0.10570	-0.42625
Kanal 4	-0.02691	0.02261	0.04596	-0.30045	0.65549	0.04238	-0.68914
Kanal 5	-0.13562	0.00951	0.14760	0.08533	-0.11795	0.96591	-0.07453
Kanal 6	-0.66251	-0.20288	0.67297	0.06035	-0.10321	-0.21748	-0.07375
Kanal 7	-0.50973	0.81599	-0.26569	-0.03756	-0.02341	-0.03694	0.02079

4.5 Filteroperationen

Filterung dient dazu, in einem Bild Erwünschtes von Unerwünschtem zu trennen. So kann es z.B. von Interesse sein, bestimmte Informationen, die im Bild nur schwach ausgeprägt sind, zu verstärken oder Bildstörungen und Rauschen herauszufiltern.

Es gibt 2 Räume, in denen man ein Bild beschreiben kann:
1. Ortsraum
2. Frequenzraum

Zuerst wird der Frequenzraum besprochen sowie die Fouriertransformation, die ein Bild vom gewohnten Ortsraum in den Frequenzraum überführt. Die Filtertypen werden dann im Ortraum näher erklärt.

4.5.1 Ortsraum und Frequenzraum

In der Digitalen Bildverarbeitung wird ein Bild als 2-dimensionales Signal dargestellt: Signal = G (x,y). G ist der Grauwert an den Ortskoordinaten x, y, so wird das Bild mit Zeilen und Spalten aufgebaut. Der Wertebereich für G liegt bei einer 8-bit Darstellung bei 0-255. Das ist der Ortsraum.

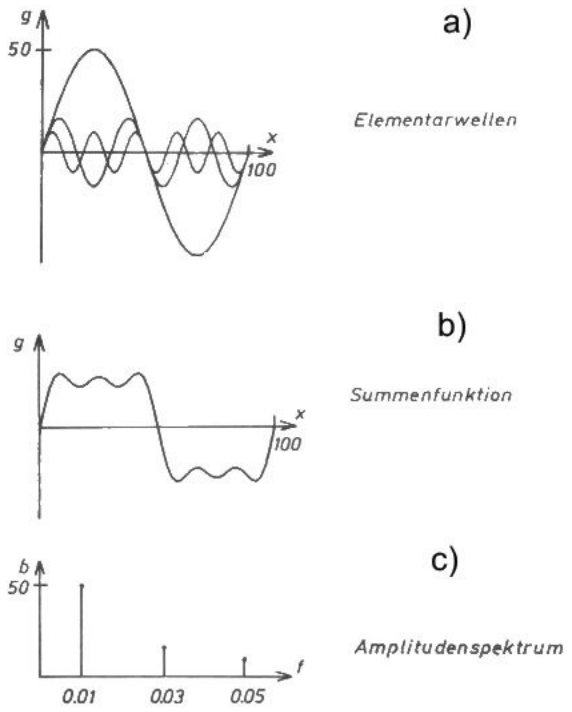
Im Frequenzraum wird das Signal als Zusammensetzung aus einzelnen Grundfrequenzen (Elementarwellen) beschrieben. Die Koeffizienten zu den einzelnen Frequenzen sind die Amplituden, mit denen die Frequenzen im Signal vorkommen. Die Frequenzen sind nichts anderes als die Reziprok-Werte der Wellenlängen.

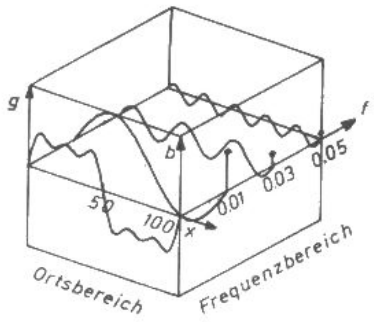
Eine Faltung im Ortsbereich entspricht einer Übertragungsfunktion im Frequenzbereich. Jedes Bild kann in jedem beliebigen Raum gefiltert werden und wieder rücktransformiert werden.

4.5.2 Filterung im Frequenzraum

Ein Bild wird vom Ortsraum in den Frequenzraum überführt, indem man eine mathematische Operation anwendet, bekannt als Fourier-Transfomation. Dabei wird eine kontinuierliche Funktion durch die Grauwerte gelegt. Diese Kurve wiederum kann durch eine unendliche Summe von Sinus- und Cosinusfunktionen dargestellt werden mit verschiedenen Amplituden, Frequenzen und Phasen.

Zur Verdeutlichung betrachten wir es im eindimensionalen Raum (siehe Abb nächste Seite): Funktion g(x) ist eine eindimensionale Grauwertfunktion. Die Abbildung zeigt eine komplexe Wellenform (**Abb b**), die z.B. eine Zeile eines Bildes beschreibt. **Abb. a** zeigt, wie diese Funktion in mehrere Sinus- und Cosinus-Wellen zerlegt werden kann mit verschiedenen Frequenzen. Sie setzt sich aus den 3 Elementarwellen mit den Frequenzen 0.01, 0.03 und 0.05 zusammen. Als Amplitudenspektrum (**Abb c**) bezeichnet man die Amplituden der Elementarwellen, die über den Frequenzen 0.01, 0.03 und 0.05 aufgetragen werden.

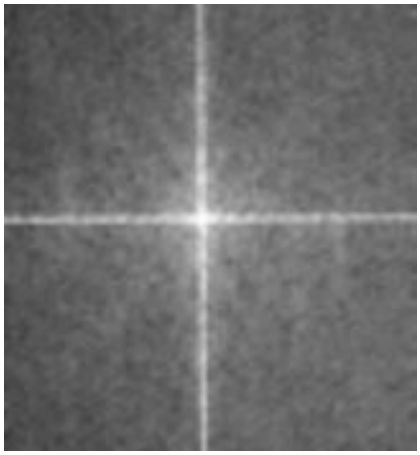




Die Abbildungen zeigen den Zusammenhang zwischen Orts- und Frequenzbereich
[Kraus 1990]

Ein 2-dimensionales digitales Bild kann man in m Grauwertzeilen oder n Grauwertspalten zerlegen und für jede dieser Zeilen bzw. Spalten eine eindimensionale Fourieranalyse durchführen. Eleganter ist die Betrachtung als 2-dimensionale Grauwertverteilung und damit 2-dimensionaler Fourier-Transformation. Der Unterschied ist, dass die Sinus- und Cosinus-Wellen 2-dimensionale Wellenformen sind und jede dieser Wellen nicht nur eine Amplitude sondern auch eine Richtung (=Phase) hat.

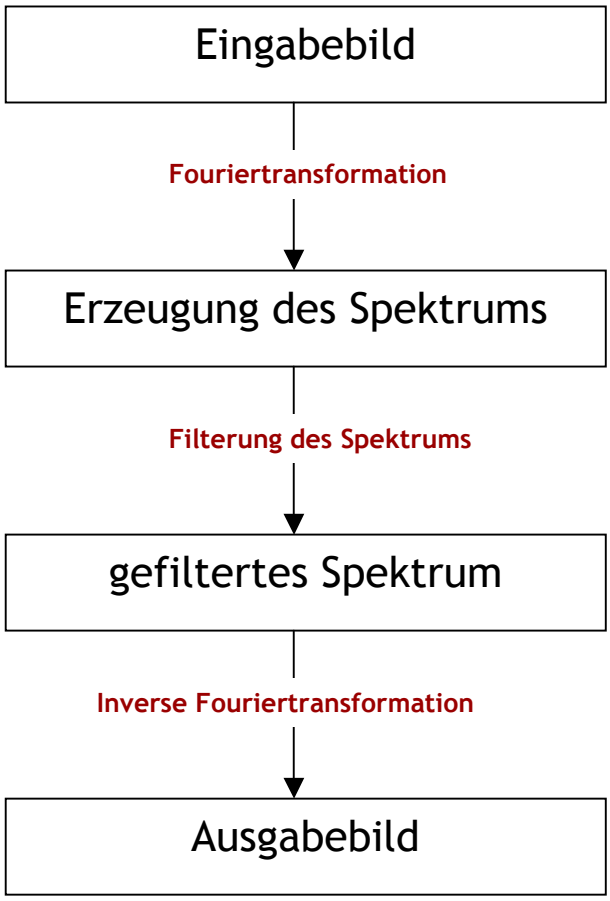
Nachdem ein Bild in seine Ortsfrequenzen umgewandelt wurde, ist es möglich, diese Werte in einem 2-dimensionalen Plot darzustellen, genannt Fourier-Spektrum oder Amplituden-Spektrum (siehe Abbildung). Die niedrigen Frequenzen (=große Wellenlängen) sind im Zentrum dargestellt und hell wiedergegeben. Sie überwiegen im allgemeinen bei einem Bild. Die höheren Frequenzen sind außen dargestellt. Der Ursprung ist in der Mitte des Plots. Features, die horizontal im Originalbild laufen resultieren in vertikalen Komponenten im Fourier-Spektrum; Features die vertikal angeordnet sind im Original-Bild resultieren in horizontalen Komponenten.



Filtertypen

Die Fourieranalyse ist hier nicht mathematische wiedergegeben worden, sondern es soll auf weiterführende Literatur von Bähr, Vögtle (1991), Kraus (1990), Haberäcker (1991) und Mather (1987) verwiesen werden. Im Exkurs „Fourier-Transformation" sind die Formeln für die Fourier-Transformation und die Inverse FT dargestellt.

Mit Hilfe der Fourieranalyse kann also ein Bild im Ortsraum in den Frequenzraum umgewandelt werden. Das Bild kann im Frequenzraum gefiltert werden indem das Spektrum mit eine Übertragungsfunktion multipliziert wird, die bestimmt, welche Frequenzkomponenten beeinflusst werden. Das gefilterte Spektrum wird dann wieder durch eine Inverse Fouriertransformation in ein Bild im Ortsraum umgewandelt (dazu benötigt man den Amplitudenanteil und den Phasenanteil, welche die relative Lage der Objekte zur Bildmitte beschreibt). Die Amplituden erlauben es, Aussagen über das Frequenzverhalten des Bildes zu machen.



Zur Filterung der Bilder im Orts- und Frequenzbereich gibt es unterschiedliche Filtertypen, die hier nur kurz angerissen werden sollen. Sie werden im nächsten Kapitel, 5.5.3 Filterung im Ortsraum, weiter ausgeführt.

Exkurs „Fourier-Transformation"

Eine allgemeine Funktion f(x) lässt sich mit der **Fouriertransformation** in die Funktion ihrer Frequenzen u überführen:

$$F\{f(x)\} = F(u) = \int_{-\infty}^{+\infty} f(x) \cdot e^{(-j2\pi ux)} dx$$

wobei F(u) das Spektrum von f(x) ist.

Erweiterung für den 2-dimensionalen Fall, wie er in der Bildverarbeitung in der Regel vorkommt:

$$F\{f(x,y)\} = F(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \cdot e^{(-j2\pi(ux+vy))} dx\, dy$$

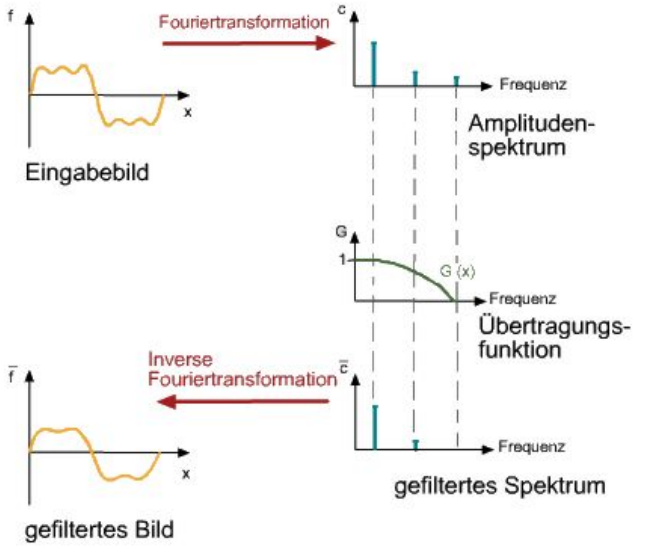
wobei F(u,v) die Fouriertransformierte (Bild-) Funktion f(x,y) ist.

Inverse Fouriertransformation:

$$f(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u,v) \cdot e^{(j2\pi(ux+vy))} du\, dv$$

Quelle: [Bähr, Vögtle 1991]
weiterführende Literatur: [Bähr, Vögtle 1991], Kraus 1990], [Richards 1994]

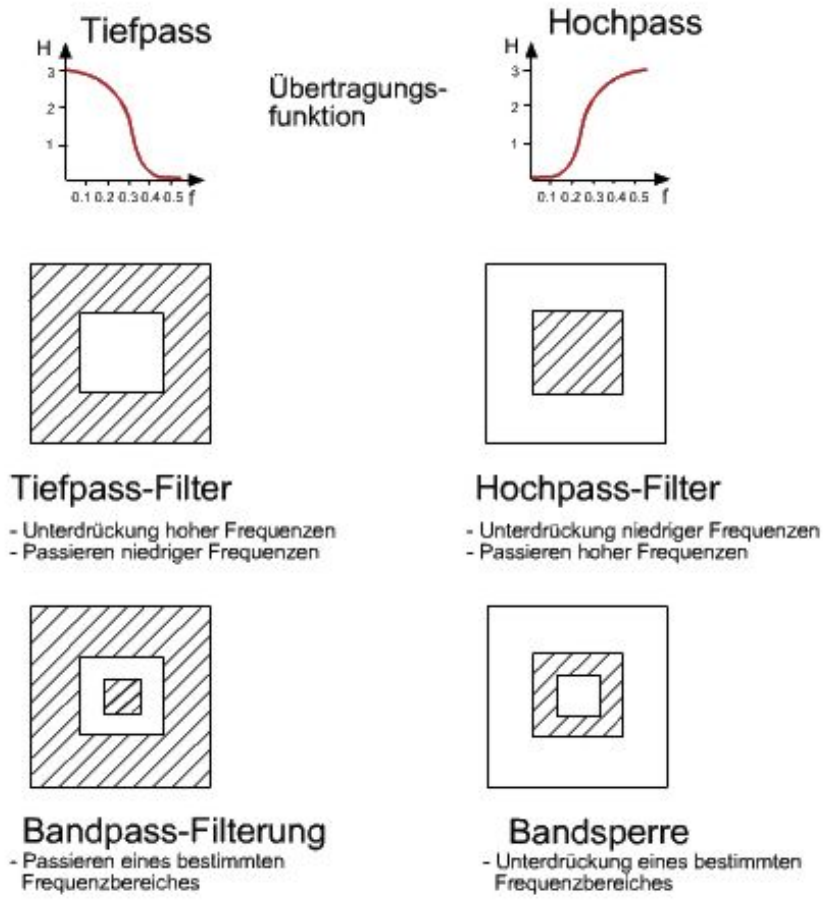
Die Abbildung unten zeigt links das Summensignal des Eingabebildes. Die Fouriertransformation liefert das Amplitudenspektrum. Durch Multiplikation mit einer Übertragungsfunktion H(f) entsteht das gefilterte Spektrum, das wiederum rücktransformiert werden kann (gefiltertes Bild). Man sieht, dass durch die Filterfunktion hohe Frequenzen ganz entfernt werden, mittlere gedämpft werden und niedrige Frequenzen gar nicht betroffen sind. In diesem Fall ist die Übertragungsfunktion ein **Tiefpassfilter**: er unterdrückt die hohen Frequenzen und lässt die tiefen Frequenzen passieren.



nach [Kraus 1990]

Ein **Hochpassfilter** bewirkt das Gegenteil: er unterdrückt die tiefen Frequenzen und lässt die hohen passieren.

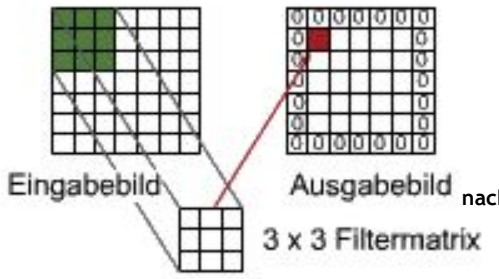
Die untenstehende Abbildung zeigt die Wirkung von Hoch- und Tiefpassfiltern im Frequenzraum. Die schraffierten Bereich werden eliminiert, die weißen durchgelassen.



Filtertypen (nach [Kraus 1990], [Kappas 1994])

4.5.3 Filterung im Ortsraum

Filterung im Ortsbereich funktioniert so, dass ein sogenanntes "gleitendes Fenster" (moving window) über die Rastermatrix läuft. Dieses Fenster ist die Filtermatrix (engl. Kernel), die eine bestimmte Ausdehnung in Zeilen- und Spaltenrichtung hat. Diese Ausdehnung umfasst in der Regel eine ungerade Zahl von Pixeln, z.B. 3×3, 5×5, 7×7. Die Matrix wird in der linken oberen Ecke des Bildes platziert (siehe Abb.). Der Grauwert für das im Zentrum der Matrix liegende Pixel wird ausgerechnet und ersetzt. Der berechnete Wert ist also eine gewichtete Summe der Pixel die unter der Filtermatrix liegen. Als Gewichte dienen die Maskenwerte.



nach [Mather 1996]

Dann wird die Matrix ein Pixel weiter nach rechts gesetzt, der Grauwert für das neue mittige Pixel ausgerechnet usw. Das wird so weitergeführt bis die Filtermatrix im rechten unteren Bildrand angekommen ist.
Das Resultat ist ein gefiltertes Bild, das allerdings ungefilterte Pixel am Rand besitzt, die dann entweder auf 0 gesetzt werden oder die Grauwerte der Nachbarpixel erhalten, um das Ausgabebild in derselben Größe zu halten wie das Eingabebild.

Die Operation dieser sog. Faltung kann dann folgendermaßen geschrieben werden:

$$f'(x,y) = f(x,y) \times g(x,y)$$

wobei f(x,y) das Ausgangsbild und g(x,y) die Filtermatrix bezeichnet.

Es gibt, wie schon dargestellt wurde, verschiedene Filtertypen. Im Ortsraum werden **Hoch- und Tiefpassfilter** verwendet. Filter wirken auf die räumliche Frequenz (Ortsfrequenz, engl. *spatial frequency*) eines Bildes, d.h. die räumliche Anordnung von Pixeln, also die Grauwertveränderungen pro Raumeinheit in einem bestimmten Teil des Bildes:

- geringe Veränderungen:
  = einheitliche Grauwerte, sanfte Übergänge
  entspricht Wellenmuster mit niedriger Frequenz, d.h. langen Wellenlängen im Frequenzraum
- hohe Veränderungen:
  = extreme Kontraste zwischen den Pixeln
  entspricht Wellenmuster mit hoher Frequenz, d.h. kurzen Wellenlängen

Die Hochpassfilter lassen also nur hochfrequente Bildanteile durch und führen so zu einer Datenverschärfung. Entgegengesetzt dazu führen Tiefpassfilter zu einer Datenglättung, da sie tieffrequente Bildanteile passieren lassen.

Tiefpassfilter

Sie setzen die räumliche Variabilität herab. Sie führen zu:

- Glättung oder Entfernung von Details im Bild;
  kann auch zu Unschärfe führen (engl. *blurring*)
- Zerstörung/Abschwächung von Bildkanten
- Filterung von hochfrequentem Rauschen

1. Mittelwert-Filter

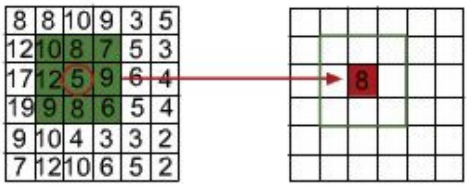
Dabei werden die Grauwerte im Fenster mit den Werten der Filtermatrix multipliziert und dann durch die Anzahl der Elemente in der Matrix geteilt.
Beispiel: bei einer 3×3 Matrix würde durch 9 geteilt werden 5×5 durch 25 usw.
Eine 5×5 -Matrix·

$$g(x,y) = \frac{1}{25} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Für ein 3×3 -Fenster sieht die Filtermatrix folgendermaßen aus:

$$g(x,y) = \frac{1}{9} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Beispiel mit dieser Matrix:



$$\frac{(1\cdot 10)+(1\cdot 8)+(1\cdot 7)+(1\cdot 12)+(1\cdot 5)+(1\cdot 9)+(1\cdot 9)+(1\cdot 8)+(1\cdot 6)}{9} = \frac{74}{9} = 8{,}2222,\ abgerun\,det = 8$$

Bei dieser Matrix sind alle Bildpunkte gleich gewichtet. Durch Änderung der Werte der Filtermatrix kann eine gezielte Gewichtung herbeigeführt werden, z.B.:

$$g(x,y) = \frac{1}{10} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Die bisher beschriebenen Mittelwert-Filter sind lineare Filteroperatoren. Sie können auch im Frequenzraum durchgeführt werden. Nichtlineare Operatoren können durch eine einfache Faltung nicht realisiert werden, weshalb sie auch im Frequenzraum nicht ausgeführt werden können.

Ein solcher Operator ist der Mittelwert-Filter mit Schwelle. Hier wird die Mittelung der Grauwerte nur dann wirksam, wenn die Differenz zwischen dem Mittelwert und dem betreffenden Originalgrauwert einen bestimmten Schwellenwert T unterschreitet, sonst bleibt der Originalgrauwert erhalten. Kanten im Bild werden dadurch, abhängig von T, unterschiedlich beeinflusst.

$$f'(x,y) = \begin{cases} f(x,y) \times g(x,y) & falls \ \left| f(x,y) - (f(x,y) \times g(x,y)) \right| < T \\ f(x,y) & sonst \end{cases}$$

2. Median-Filter

Der Median-Filter ist ein weiterer nichtlinearer Filter. Die Grauwerte innerhalb des Fensters *m* werden der Größe nach sortiert. Dem Zentralpixel wird dann der Medianwert, der eine Population in 2 gleich große Hälften teilt, zugewiesen.

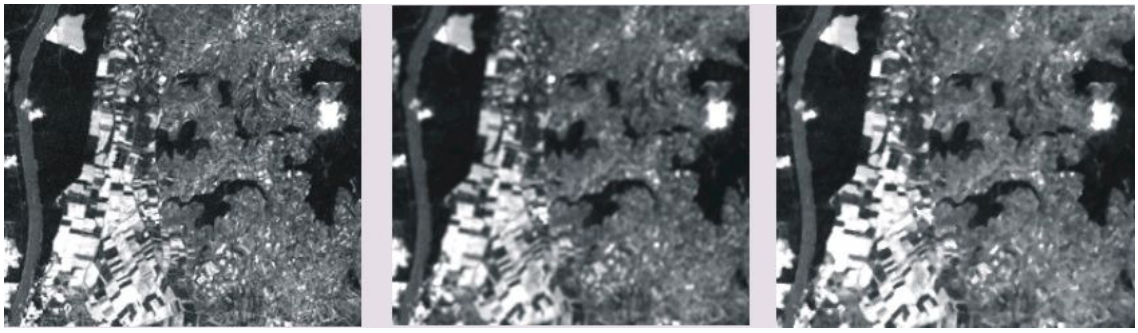
Beispiel:

5	3	1
2	8	10
9	5	21

m = {5, 3, 1, 2, 8, 10, 9, 5, 21}
m' = {1, 2, 3, 5, **5**, 8, 9, 10, 21}

Das Zentralpixel bekommt also den Wert **5**.
Der Mittelwert wäre 7,111 gewesen, abgerundet 7. Dieser Wert kommt bei den Originalgrauwerten nicht vor und ist von dem extremen Grauwert 21 beeinflusst. Der Medianfilter entfernt also extreme Grauwerte, die oft Rauschanteile sind. Außerdem werden beim Median-Filter die Originalgrauwerte nicht verändert. Er kann daher oft ein besserer Kompromiss zwischen Glättung und Kantenerhaltung sein als der Mittelwert-Filter.

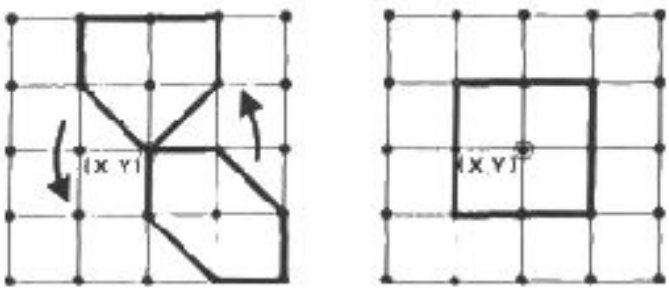
Vergleich Mittelwert- und Median-Filter:



Originalbild **Mittelwert-Filter** **Median-Filter**

3. Adaptive Filter:

Adaptive Filter sind nicht-lineare Filter, die eine kantenerhaltende Glättung durchführen. Sie berücksichtigen die räumliche Verteilung der Grauwerte innerhalb des Fensters, im Gegensatz zu Median- als auch Mittelwertfilter, die festgelegte Gewichtungsfaktoren benutzen.



Die verschiedenen Untermengen von zusammenhängenden Pixeln in einem 5×5 Fenster [Bähr, Vögtle 1991]

Vorgehen:

1. Rotation einer definierten Maske in einem mindestens 5 ⊠5 Pixeln großen Fenster
2. Berechnung der Varianz der Grauwerte der zusammenhängenden Untermenge in jeder Position
3. Ermittlung der kleinsten Varianz aus den Mittelwerten der Untermenge
4. Zuweisung des neuen Wertes dem zentralen Pixel (x,y)
5. Mehrfache Wiederholung des Prozesses bis die Grauwertänderung einen bestimmten Schwellenwert unterschreitet (iterative Anwendung)

[Bähr, Vögtle 1991]

Stellen hoher Grauwertvarianz, wie sie in der Umgebung scharfer Kanten auftreten, werden somit nicht in den Glättungsprozess mit einbezogen.

Hochpassfilter

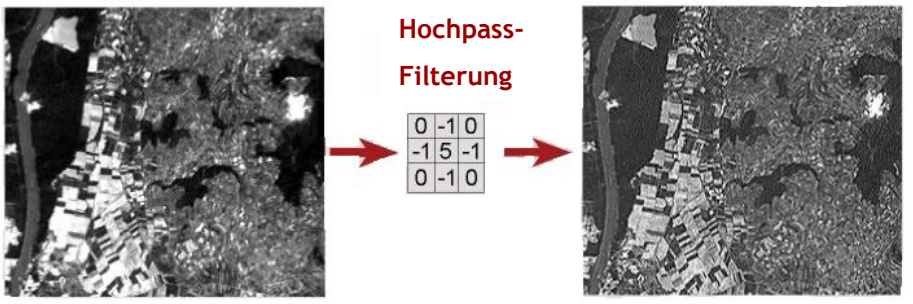
Einen gegenteiligen Effekt haben Hochpassfilter. Sie verstärken die Grauwertunterschiede im Bild und betonen Grauwertkanten und Linien. Man erhält dadurch ein kontrastreicheres, "schärferes" Bild, allerdings wird das Bildrauschen ebenfalls verstärkt.
Sie lassen auch negative Werte zu, so ergeben sich sog. Differenzoperatoren.

Beispiele für Hochpassfilter

$$g(x,y) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{pmatrix} \qquad g(x,y) = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

Je stärker die Gewichtung des zentralen Pixels in der Filtermatrix, desto größer ist der Effekt der Filterung.

Die Abbildung unten zeigt den Effekt der Hochpass-Filterung:

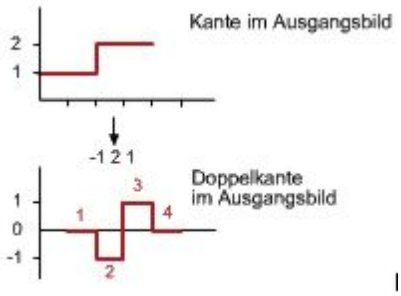


Kantenextraktion (*edge detection*):

Bei dieser Art von Filtern werden die Kanten regelrecht herausmodelliert. Sie sind sogenannte "Nullsummen-Filter", d.h. dass die Summe der Elemente innerhalb der Maske gleich 0 ist. Dadurch erscheinen alle homogenen Flächen schwarz und die starken Grauwertsprünge (also Kanten) werden hell dargestellt. Bei Grauwertübergängen entsteht eine Maßzahl für die Stärke des Übergangs.
Es gibt diverse Filter mit verschiedenen Matrizen, z.B. Laplace-Filter, Sobel-Filter, Prewitt-Filter, Roberts-Filter etc.

1. Laplace-Filter:

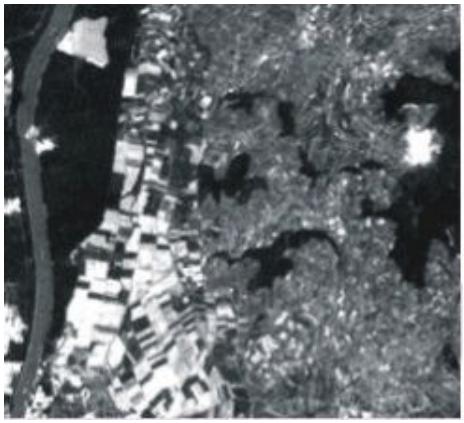
Er ist ein richtungsunabhängiger Kantenfilter, arbeitet also in horizontaler, vertikaler und diagonaler Richtung. Er wird zur Extraktion von Grauwertkanten eingesetzt: läuft er über eine Kante, verstärkt er diese durch Bildung einer Doppelkante [Kappas1994]: Der Operator nimmt einen negativen Wert an am Fuß der Kante, einen positiven am Gipfel.



[Kappas 1994]

Der horizontale Filter mit dem vertikalen Filter ergibt den Laplace-Filter:

$$(-1 \quad 2 \quad -1)\, mit \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} ergibt \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$



Originalbild

Laplace-Filter

Weitere Laplace-Filter: $g(x,y) = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$ $g(x,y) = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$

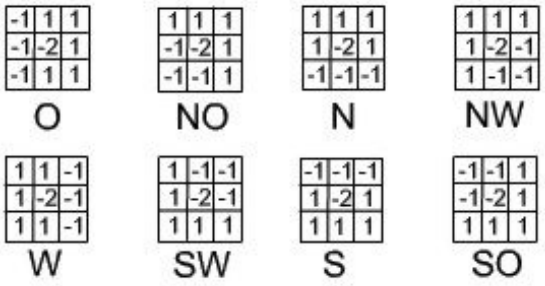
Ein kantenverstärktes Bild würde durch eine Subtraktion des Laplace-Filters vom Originalbild entstehen:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

Originalbild Laplace-Filter kantenverstärktes Bild

2. Richungsabhängige Filter:

Sie extrahieren Kanten, die in bestimmten Richtungen verlaufen. Sie werden Kompassgradienten genannt. Für alle 8 Richtungen gibt es Masken [Haberäcker 1991].



Beispiel für richtungsabhängige Extraktion von Kanten:

Beispiel für richtungsabhängige Extraktion von Kanten
links: Originalbild; links unten: NW-SO verlaufende Kanten; rechts unten: NO-SW verlaufende Kanten

3. Sobel-Filter:

Der Sobel-Filter ist ebenfalls ein richtungsabhängiger Filter. Er arbeitet mit 2 Masken in horizontaler und vertikaler Richtung:

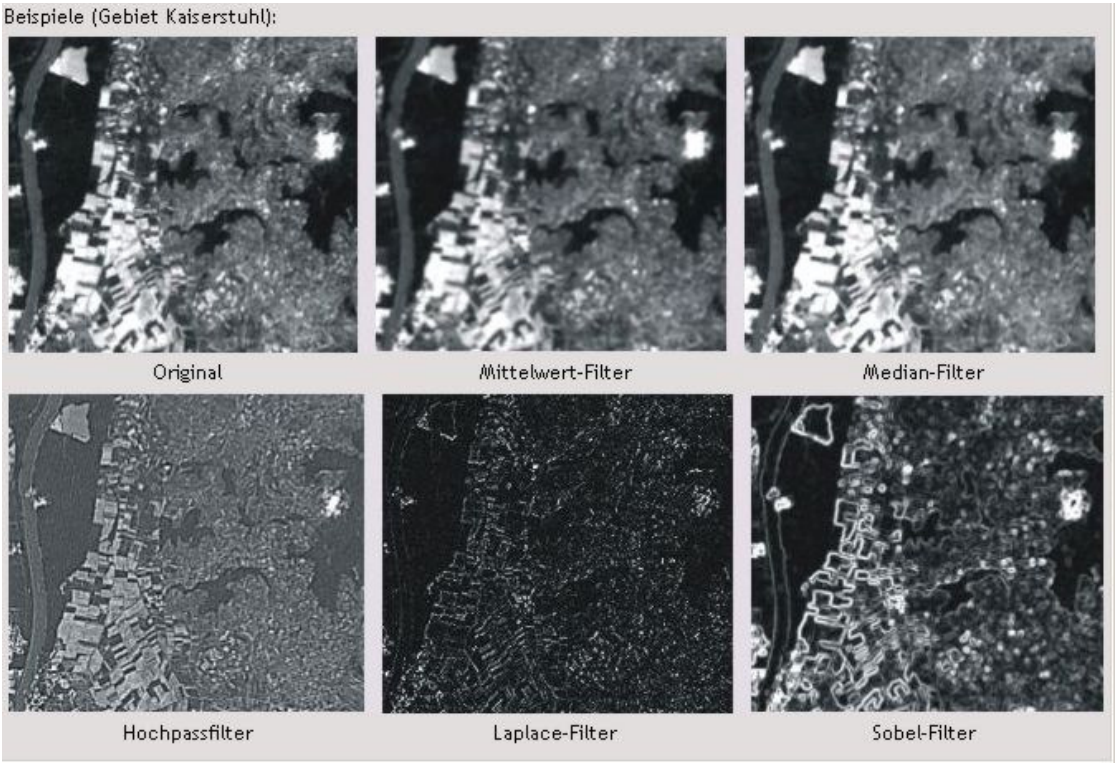
$$G_x = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad \text{und} \quad G_y = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$$

Sie mitteln einerseits in Zeilen- bzw. Spaltenrichtung (bei starker Betonung der zentralen Werte) und glätten andererseits in Spalten- bzw. Zeilenrichtung, da jeweils eine Zeile (bzw. Spalte) übersprungen wird.

Die gefalteten Bilder werden dann folgendermaßen zusammengefasst unter Berechnung des Gesamtgradienten: $G = \sqrt{G_x^2 + G_y^2}$

Sobel-Filterung

Zusammenfassung der Filtertypen



4.6 Klassifikation

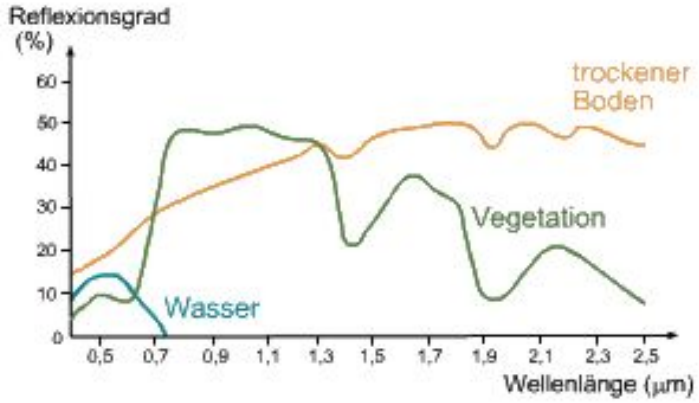
Definition:

Klassifikation bedeutet allgemein gesagt das Zuordnen von Pixeln zu abgegrenzten Klassen mit unterschiedlichen Charakteristika. [Bähr 1999]

Dieses Kapitel befasst sich mit der Klassifikation nach spektralen Merkmalen, daher sind die spektralen Signaturen der Objekte die wichtigsten Parameter. Die spektralen Signaturen werden in Grauwerten wiedergegeben, welche den jeweiligen Pixeln im Bild zugeordnet sind.

$$g_{i,k} = x_{i,k} = (x_1, x_2, x_3 \ldots x_n)_{i,k}$$

wobei n = die Anzahl der Spektralbänder. Der Grauwert $g_{i,k}$ wird von n spektralen Komponenten aufgebaut mit den jeweiligen Grauwerten ($x_1 \ldots x_n$).
Effekte benachbarter Pixel werden bei der spektralen Klassifikation nicht mit einbezogen im Gegensatz zur objektorientierten Klassifikation.
Den Pixel werden bei der spektralen Klassifikation anhand ihrer Reflexionseigenschaften (spektralen Signaturen) thematische Eigenschaften zugeordnet und so in Objektklassen eingeteilt. Objektlassen wie z.B. Wasser, Boden, Vegetation weisen sehr unterschiedliche Reflexionseigenschaften auf (siehe Abbildung).



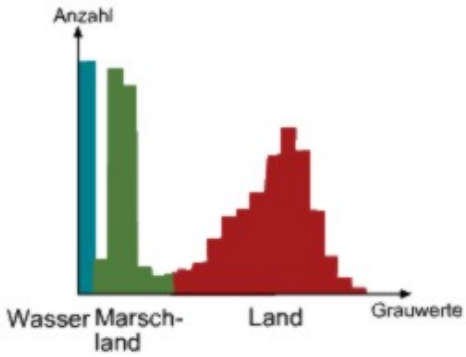
Reflexionsgrade verschiedener Oberflächenmaterialien

Die Funktionsweise der Klassifikation kann am besten durch Histogramme erklärt werden. Varianz, Mittel und Standardabweichung sind dabei wichtige Parameter zur Einteilung der Pixel in Klassen. Siehe Kapitel 5.1.2 zur Erläuterung von Histogrammen.

4.6.1. Monospektrale Klassifikation

Wir haben in Kapitel 5.3.3 bereits ein Klassifizierungsverfahren kennengelernt: die Äquidensitenherstellung. Dabei werden Grauwerte eines Kanals durch Schwellenwerte voneinander getrennt. Das untenstehende Bild zeigt ein einkanaliges Satellitenbild des Jadebusens (Landsat MSS, Kanal 7 - Infrarot). 3 verschiedene Bereiche können im Bild visuell unterschieden werden. Das Histogramm (Abb. rechts) zeigt dies ebenfalls: die dunklen Bereiche entsprechen Wasser, die hellen Land und die mittleren Marschland. Anhand des Histogramms kann auch erkannt werden, dass die Klasse "Land" eine größere Varianz aufweist als die Klasse "Wasser".

Landsat MSS Kanal 7; Jadebusen bei Wilhelmshafen
aus: [Bähr, Vögtle 1999]

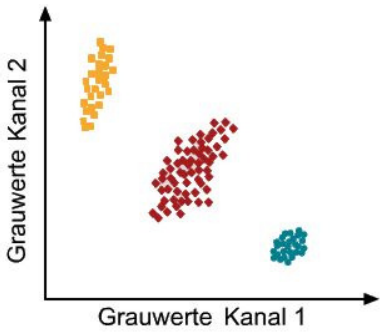


Histogramm der Abbildung oben (generalisiert)
[nach Bähr, Vögtle 1999]

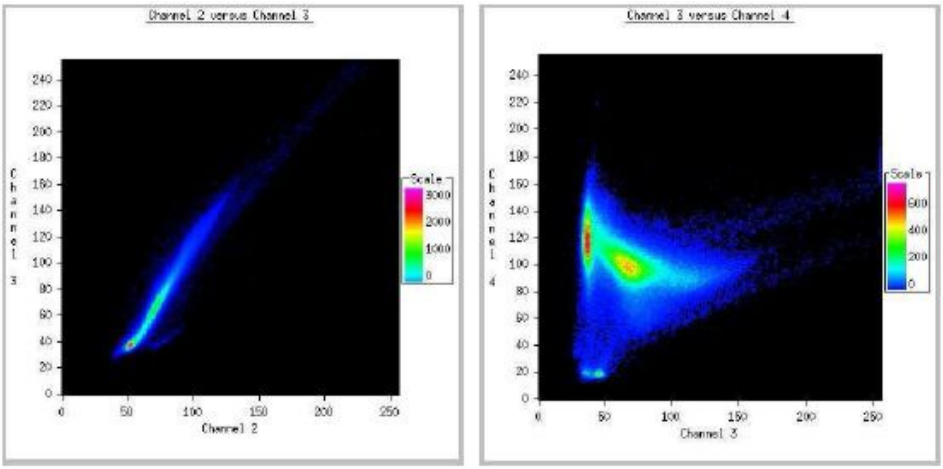
Die Genauigkeit der Klassifikation kann dadurch erhöht werden, dass mehr als ein Kanal benutzt wird, da infolge der unregelmäßigen Streuung in einer Objektklasse eine Unterscheidung in Klassen nur sehr schwer ausgeführt werden kann.

4.6.2. Multispektralklassifikation

Zur besseren Anschaulichkeit sei zuerst die Klassifikation mit 2 Kanälen erklärt.
Wenn die Grauwerte von Kanal 1 gegen die Grauwerte von Kanal 2 aufgetragen werden entstehen Punktwolken (=*Cluster*) (siehe Abb. unten). Größe und Form dieser Cluster sind eine Funktion der Variabilität der Objekttypen, des Sensorrauschens, topographischer Effekte etc. Die Grauwerte einer bestimmten Objektklasse (z.B. Vegetation) unterscheiden sich von Kanal zu Kanal, obwohl es Ähnlichkeiten in bestimmten Kanälen geben kann (d.h. hohe Korrelation zwischen Kanälen).



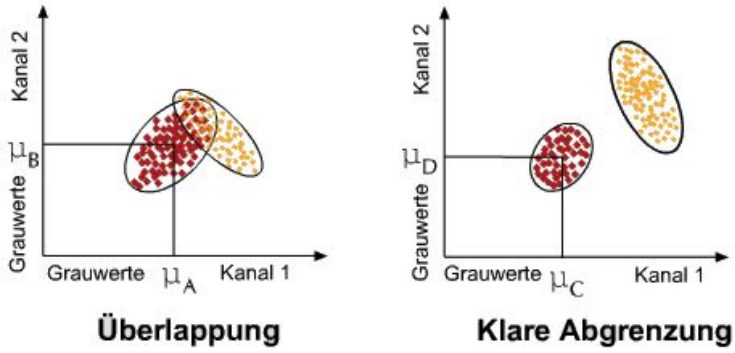
2-dimensionales Histogramm



Clusterdarstellung (=Scattergramm) eines realen Bildes

Je näher zwei Objekt-Punkte im Merkmalsraum zueinander liegen, um so ähnlicher sind sie einander und mit um so größerer Wahrscheinlichkeit werden sie demselben Typ angehören.

Die Punktwolken sind dann gut voneinander abgegrenzt, wenn die Schwerpunkte der Punktwolken weit auseinander liegen und die Klasse homogen ist. Falls die Grauwerte der einzelnen Klassen stark streuen, können sich die Punktwolken überlappen und somit die Klassen nicht mehr klar getrennt werden. Je größer die Schnittmenge, desto schwerer ist die Klassenzuordnung:



Lage zweier Punktwolken im 2-dimensionalen Merkmalsraum

Nimmt man weitere Kanäle hinzu, bekommt man einen mehrdimensionalen Merkmalsraum (engl. *feature space*), durch den die Bildelemente besser in Klassen einzuteilen sind. Bei 5 Kanälen/Spektralbereichen erhält man z.B. einen 5-dimensionalen Merkmalsraum.

Man kann verschiedene Klassifikationsverfahren unterscheiden, die sich in zwei große Gruppen einteilen lassen:

- unüberwachte Klassifikation (unsupervised classification)
- überwachte Klassifikation (supervised classification)

Bei der unüberwachten Klassifikation werden die multispektralen Daten auf Grundlage der Ähnlichkeit in Gruppen bzw. Klassen gruppiert (statistische Verfahren) und dann diese Klassen identifiziert indem man ihnen - sofern möglich - thematische Kategorien zuordnet. Bei der überwachten Klassifikation geht man umgekehrt vor: man nimmt Stichproben ("Trainingsgebiete") von zu klassifizierenden Oberflächenklassen und untersucht das gesamte Gebiet auf Grundlagen der Charakteristika der Trainingsgebiete.
Im folgenden werden diese beiden Verfahren näher erläutert.

4.6.3 Unüberwachte Klassifikation

Basierend auf statistischen Informationen der Daten werden spektrale Klassen zusammengestellt. Danach werden sie von dem Analyst Informationsklassen zugeordnet. Programme, die **Clustering Algorithmen** genannt werden, werden benutzt um die Pixel im multispektralen Merkmalsraum nach statistischen Gesichtspunkten zu gruppieren. Um die "Ähnlichkeit" von Pixeln festzustellen wird auf Distanzmaße zurückgegriffen, z.B. die Euklidische Distanz. Ziel ist, eine optimale Unterteilung des Merkmalsraumes zu erreichen. Normalerweise kann der Analyst spezifizieren wie viele Cluster er haben will, außerdem kann er noch Parameter zur Trennung von Klassen einstellen.

Bei diesem Verfahren bringt der Analyst nicht sein a-priori Wissen ein wie bei der Überwachten Klassifikation. Die Daten werden ("automatisch") in Klassen aufgespalten ohne dass man die "richtigen" Informationsklassen kennt (z.B. Landbedeckungsklassen wie Vegetation, Wasser, Boden). Als Resultat des Klassifikationsprozesses bekommt man also eine Verteilung von Grauwerten und der Analyst muss diese Cluster den Informationsklassen zuordnen. Dabei nimmt er die original Satelliten- oder Luftbildszene zu Hilfe um einem Cluster die wahrscheinlichste Bodenbedeckungsklasse zuzuordnen.

Im allgemeinen führt die unüberwachte Klassifikation nicht zu zufriedenstellenden Ergebnissen. Sie kann aber sehr gut genutzt werden um mögliche Klassen herauszufinden und die statistischen Merkmale der Daten kennenzulernen. Daher wird sie oft vor der überwachten Klassifikation durchgeführt.

Ein bekannter Algorithmus ist der ISODATA-Algorithmus.

Isodata Algorithmus:

(ISODATA - **I**terative **S**elf-**O**rganizing Data Analysis Technique)

- Initialisierung von K-Clusterzentren $c_{i,\,t}$ mit i=1,...,K (Anzahl der cluster), t=1,...,t_{Max} (terationsschritt)
- Zuordnung der Stichproben x
  $$x \in C_{j,t} falls \left\| x - C_{j,t} \right\| < \left\| x - C_{i,t} \right\|, i \neq j$$
- Verwerfung von kleinen Clustern
- Berechnung von neuen Clusterzentren $c_{i,t+1}$, z.B. bei minimaler Quadratsumme der Distanzen:
  $$C_{j,t+1} = \frac{1}{N_j} \sum_{x \in C_{j,t}} x$$
- falls Cluster verworfen, weiter mit Zuordnung der Stichproben x
- Zusammenfassung von Clustern

4.6.3 Überwachte Klassifikation

Überwachte Klassifikation basiert ebenfalls auf statistischen Parametern. Der Bearbeiter bringt sein a-priori Wissen am Anfang des Prozesses ein, nicht wie bei der unüberwachten Klassifikation am Ende.

Statistische Parameter:

$x = (x_1, x_2 \ldots x_n)^T$	beschreibt Grauwertvektor für 1 Pixel in n Kanälen
$\Omega \equiv (\omega_1, \omega_2 \ldots \omega_m)$	Vektor von *m* Klassen
$p(x \mid \omega_j)$	Verteilung von x in Klasse ω_j
$p(\omega_j)$	a-priori Wahrscheinlichkeit für das Auftreten der Klasse ω_j
$p(\omega_j \mid x)$	a-posteriori Wahrscheinlichkeit von x in Klasse ω_j = die Wahrscheinlichkeit dass ein Grauwert x der Klasse ω_j angehört

Generell ist die a-priori Wahrscheinlichkeit unbekannt. Sie wird allerdings durch "Erfahrung" hergeleitet. Es ist zum Beispiel sehr unwahrscheinlich (aber nicht unmöglich) dass eine Stadt sich mitten im Wasser befindet. Die Wahrscheinlichkeit dass eine bestimmte Klasse in einem bestimmten Umfeld existiert ist eine wichtige Information.
In der Klassifikation wird die a-posteriori Wahrscheinlichkeit gesucht, also die Wahrscheinlichkeit mit der der Grauwertvektor x ein Element einer bestimmten Klasse ω_j ist. Also werden Verfahren angewandt, mit dem x einer definierten Klasse ω_j zugeordnet werden kann.

Bayes-Regel:

Dies ist ein grundlegendes Gesetz in der Statistik. Es erlaubt die Bestimmung der Wahrscheinlichkeit, dass ein Pixel x zu einer Klasse ω_j gehört.

Die a-posterior Warhscheinlichkeit ist gegeben durch:

$$p(\omega_j \mid x) = \frac{p(x \mid \omega_j) p(\omega_j)}{p(x)} \quad (1)$$

mit

$$p(x) = \sum_{i=1}^{m} p(x \mid \omega_j) p(\omega_j) \quad (2)$$

(2) in (1) = **Bayes-Regel**

$$p(\omega_j \mid x) = \frac{p(x \mid \omega_j) p(\omega_j)}{\sum_{i=1}^{m} p(x \mid \omega_j) p(\omega_j)}$$

Die Wahrscheinlichkeit ist hoch wenn die Position eines Pixels x dicht an dem erwarteten zentralen Wert der Verteilungsfunktion liegt.

Ein Pixel x wird also der Klasse zugeordnet, welche die größte Wahrscheinlichkeit zeigt. Dieser einfache Grundsatz wird ausgedrückt durch das Klassifizierungsgesetz:

Klassifizierungsgesetz:

Der Mermalsvektor x wird der Klasse wj zugeordnet, wenn:

$$p(x \mid \omega_i)\, p(\omega_i) > p(x \mid \omega_j)\, p(\omega_j)$$

$$g_i(x) > g_j(x)$$ **Bewertungsfunktion**

$$g_i(x) = g_j(x)$$ **Trennfunktion**

Man kann Trennfunktionen bestimmen wenn beide Elemente der obersten Gleichung gleich gesetzt werden.

Ein wichtiger Parameter im Klassifizierungsgesetz ist die Verteilungsfunktion $p(x \mid \omega_j)$:

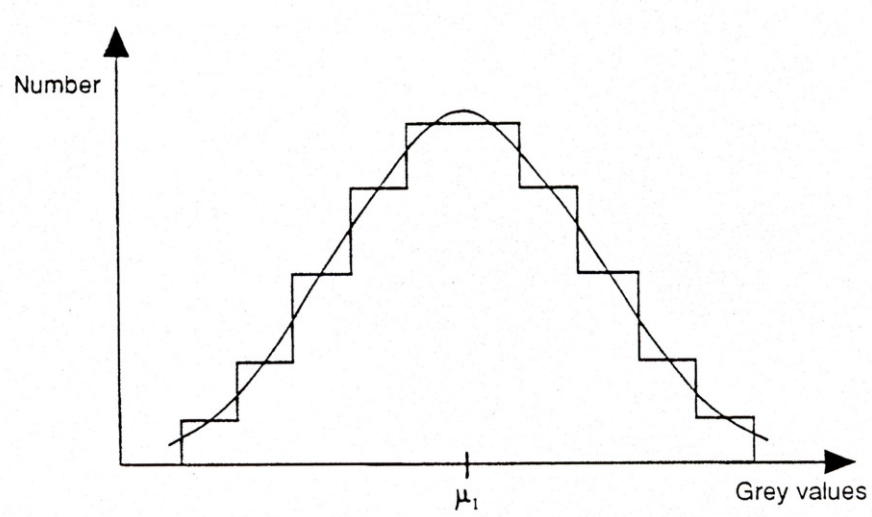
$$p(x \mid \omega_j) = \frac{1}{\sqrt{2\pi^n \mid C_j}} e^{-\frac{1}{2}(x-\mu_i)^T C_j^{-1} (x-\mu_i)}$$

Mittelwert vektor: $\mu_i = \sum_{i=1}^{N} x_i$

Kovarianz matrix: $C_j = (\sigma_{kl}) = \frac{\sum_{i=1}^{N} (x_i - \mu_j)(x_i - \mu_k)}{N-1}$

i: Index des Pixels der Stichprobe
j, k: Index der Klassen

Dies ist ein Algorithmus für eine multidimensionale Normalverteilung. In der Statistik wird sie oft verwendet, da ihre mathematische Manipulation nicht zu komplex ist und sie außerdem die Natur gut beschreibt. Die Normalverteilung für eine Dimension ist die Gauß'sche Normalverteilung:



Normalverteilung (Gauß'sche Funktion) [Bähr, Vögtle 1999]

Da wie schon erwähnt, jede Klasse eine eigene spezifische Verteilung hat, hat die Gauß'sche Funktion jeder dieser Klassen verschiedene Mittelwerte μ und Varianzen σ^2. Bei 2 Dimensionen haben die Verteilungen die Form von Ellipsen.
- Mittelwert = Mittelpunkt des Custers
- Varianz = Form und Orientierung der Ellipsen

Bei multispektraler Verteilung bekommen wir statt der einfachen Varianz σ^2 die Kovarianzmatrix C

$$C_i = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \cdots & & & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}_i$$

Für eine multispektrale Klassifikation brauchen wir die realen Werte der Klassen, d.h. die unbekannten Parameter der Normalverteilung (Mittelwert μ und die Kovarianzmatrix). Diese Unbekannten werden für Probenflächen bestimmt, sogenannte "Trainingsgebiete".

Das Resultat einer Klassifikation hängt in großem Maße davon ab, wie sehr diese mathematischen Verteilungsfunktionen der Klassen der wirklichen Verteilung in der Natur entsprechen. Die Grundfrage ist dabei natürlich: "Wie korrekt lässt sich die Natur in Klassen unterteilen, die der Normalverteilung folgen?"

4.6.5 Algorithmen für die überwachte Klassifikation

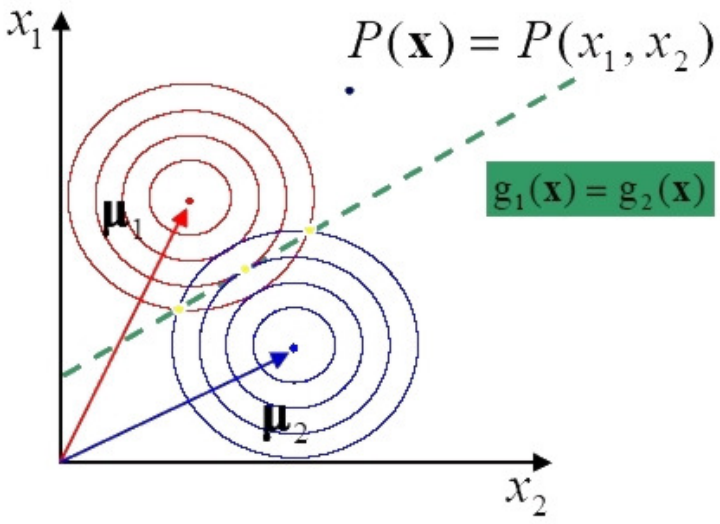
Im Folgenden werden Algorithmen für die überwachte Klassifikation erklärt. Es gibt für die Algorithmen keine Einschränkungen in Bezug auf Klassenanzahl oder Anzahl der benutzten Kanäle, sie unterscheiden sich jedoch in ihrer Kovarianzmatrix.

- Minimum Distance – Klassifikator
- Mahalanobis Klassifikator
- Maximum-Likelihood Klassifikator
  außerdem noch:
  Parallelpiped Klassifikator (Quaderklassifikator)

1. Minimum Distance-Klassifikator

$$C_i = E$$

Die Verteilungen der Klassen i und j sind konzentrische Kreise, deren Mittelpunkt der Mittelwert μ ist. Für jedes Pixel wird der euklidische Abstand zu den Mittelpunkten aller Klassen ermittelt. Das Pixel $p(x_i/x_2)$ wird jener Klasse zugeteilt, zu deren Mittelpunkt der Abstand am geringsten ist (auch Nearest-Neighbour-Klassifikator genannt).



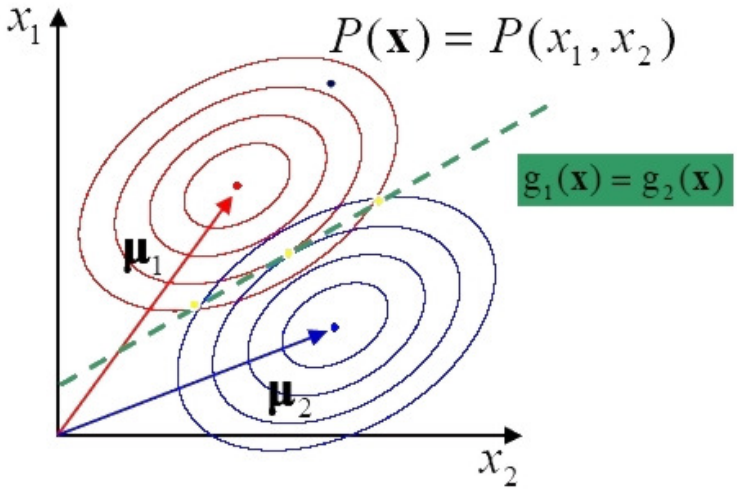
Prinzip des Minimum Distance Klassifikators

Innerhalb eines 2D-Merkmalsraumes sind die Trennfunktionen Geraden. Diese Methode ergibt keine guten Resultate, da die Klassen nicht ihrer wahren Natur entsprechend modelliert werden: die Cluster stellen nämlich keine Kreise dar wie im Algorithmus, sondern Ellipsen.

2. Mahalanobis Klassifikator

$C_i = C$

Hier erfolgt die Zuordnung eines Pixels über den kürzesten statistischen Abstand (die Mahalonobis-Distanz). Beim Mahalonobis-Klassifikator wird nur eine Kovarianzmatrix benutzt statt verschiedene für jede der i-Klassen.
Im 2-dimensionalen Merkmalsraum sind die Cluster Ellipsen, keine Kreise wie bei der Minimum Distance-Methode. Die Trennfunktion ist wieder eine Gerade, jedoch ist ihr Verlauf durch die Kovarianzmatrix bestimmt und nicht mehr die Gerade zwischen den Klassenmittelpunkten und der Position des Pixels.



Prinzip des Mahalanobis Klassifikators

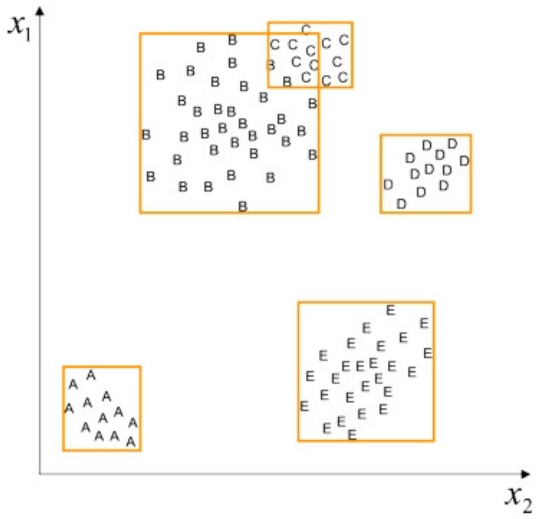
3. Maximum-Likelihood-Klassifikator

$C_i \neq C_j$

Dieser Klassifikator benutzt verschiedene Kovarianzmatrizen für jede Klasse. Er liefert bessere Resultate als die vorhergehenden, hat allerdings eine höhere Rechenzeit, da jede Kovarianzmatrix der i-Klassen berechnet werden muss.
Da die Kovarianzmatrix für jede Klasse verschieden ist, ist die Trennfunktion keine gerade Linie, sondern, bei 2 Kanälen eine Kurve, bei mehr Kanälen sind sie Flächen höherer Ordnung.

Quadermethode (Parallelepiped-Verfahren):

Es werden für die Trainingsgebiete jeder Objektklasse die obere und untere Spektralgrenze definiert (im 2D-Raum sind das Rechtecke, im 3D-Raum Quader). Alle Pixel werden nun entsprechend dieser Eckdaten auf ihre Zugehörigkeit geprüft. Pixel, welche in keine Klasse fallen, bleiben unklassifiziert.



Prinzip des Quaderklassifikators

4.6.6 Ablauf einer überwachten Klassifikation

Wichtige Arbeitsschritte:

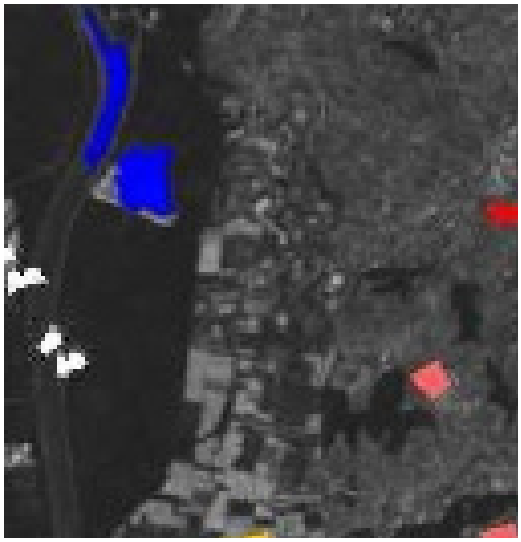
- Festlegung der Landnutzungsklassen (spektrale Klassen)
- Festlegung von Trainingsgebieten für jede Klasse
- Schätzung der Merkmale (Signaturen) für jede Klasse
- Durchführung der eigentlichen Klassifikation
- Bewertung und Überprüfung der Ergebnisse

1. Festlegung der Landnutzungsklassen:
Der Bearbeiter muss die zu klassifizierenden Landnutzungsklassen festlegen. Häufig erfolgt eine hierarchische Klassenfestlegung.
Beispiel: Klasse: Wald
Unterklasse: Laubwald, Nadelwald

2. Festlegung von Trainingsgebieten
Sie werden unter Nutzung von Karten, GIS, Luftbildern, etc. bestimmt.

Anforderungen an Trainingsgebiete:

- sie sollen repräsentativ sein
- homogen aber nicht uniform
- so groß wie nötig aber so klein wie möglich



Gebiet Kaiserstuhl. Festlegung der Trainingsgebiete

3. Schätzung der Merkmale für jede Klasse

Für jede Klasse werden Schätzungen der Merkmale erstellt. Die Varianz und der Mittelwert werden für jeden zu nutzenden Kanal berechnet.

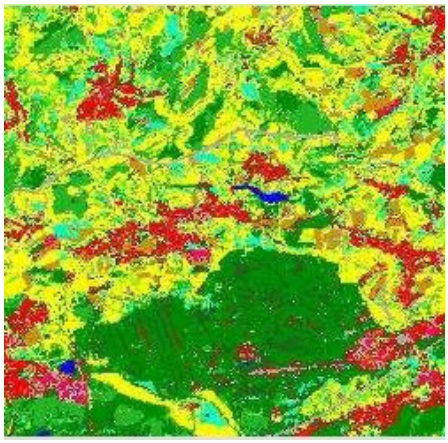
	Value	Name	Color	Threshold	Bias
18	18	wein1		7.00	1.00
19	19	wein2		7.00	1.00
20	20	wein3		7.00	1.00
21	21	obst1		7.00	1.00

General | Matrices

Channel	Mean	Std.Dev
1	84.82	3.07044
2	74.64	3.3212
3	74.44	5.98384
4	91.2	2.87054
5	121.14	8.79709
7	80.28	8.13883
8	64.88	1.72789
9	27.06	1.19013
10	25.66	2.0439
11	74.52	3.36593
12	86.96	4.76727
14	24.36	2.3984

Samples: 50

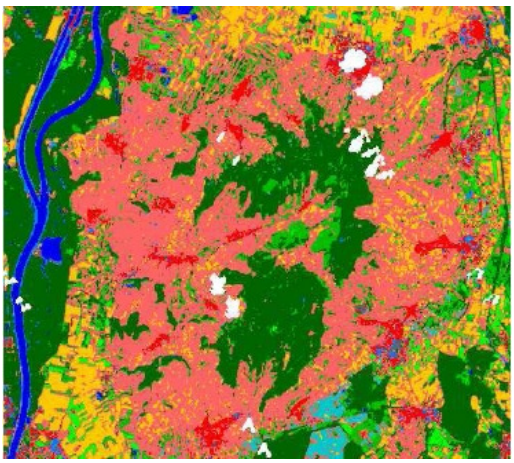
4. Durchführung der eigentlichen Klassifikation
z.B. mit Maximum-Likelihood-Klassifikator



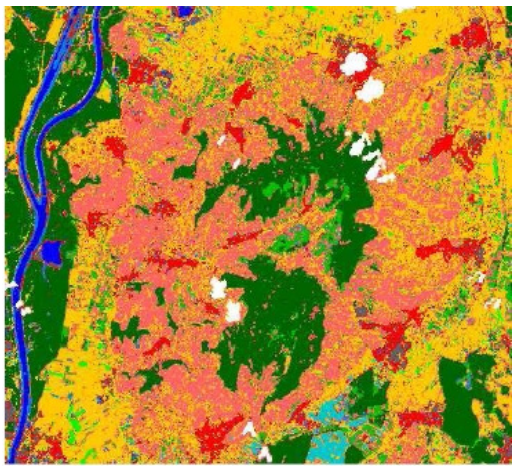
Beispiel einer Klassifikation

Legende	Groesse (ha)	Anteil
nicht klassifiziert	4.154	4,2 %
Strassen, Parkplatz, Stein	2.020	2,1 %
lockere Besiedlung	5.154	5,3 %
dichte Besiedlung	1.079	1,1 %
Industrie, Gewerbe, Hallen	848	0,9 %
Nadelwald	18.384	18,9 %
Laubwald	20.035	20,6 %
extensiv genutztes Gruenland	3.941	4,1 %
Wasserflaechen	80	0,1 %
Ackerland (offener Boden)	7.073	7,3 %
Ackerland (vegetationsbedeckt)	30.337	31,2 %
Ackerland (folienbedeckt)	118	0,1 %
Wiese, Weide	850	0,9 %
Windwurfflaechen	3.089	3,2 %
Eisenbahnanlagen	38	0,04 %
Summe	97.200	100,0 %

Vergleich der Ergebnisse einer Klassifikation mit dem Minimum-Distance Klassifikator und dem Maximum-Likelihood Klassifikator:

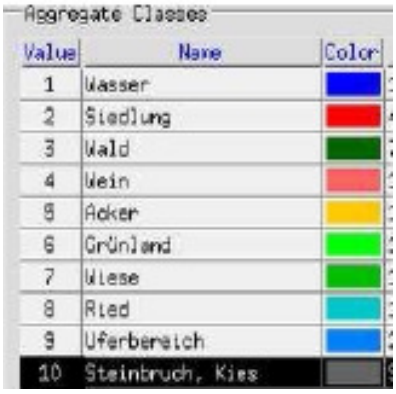


Minimum Distance



Maximum-Likelihood

Vergleich von Klassifikationen mit Minimum Distance und Maximum-Likelihood Klassifikator
Gebiet Kaiserstuhl, Landsat TM 2000
Rechts die Legende. Die weißen Gebiete sind Wolken.



5. Bewertung und Überprüfung der Ergebnisse
Überprüft wird:
- Wahrscheinlichkeit der Klassenzugehörigkeit
- Differenz der Wahrscheinlichkeiten zur nächsten Klasse

Dafür wird eine Konfusionsmatrix der Trainingsgebiete erstellt, die die Eignung der Trainingsgebiete aufzeigt.
Unter Nutzung von Karten, GIS, Feldvergleichen werden Kontrollgebiete festgelegt. Es wird eine Konfusionsmatrix der Kontrollgebiete erstellt und anhand dieser Referenz die Klassifikation bewertet.

Beispiel für Konfusionsmatrix:

Klassifikation	Referenz: Siedlung	Wein	Acker	Grünland	Wald	Total	C.A. (%)
Siedlung	400	5	2	0	0	407	98.3
Wein	3	200	1	0	2	206	97.1
Acker	2	5	180	0	0	187	96.3
Grünland	0	1	1	210	5	217	96.8
Wald	0	1	0	4	300	305	98.4
Total	405	212	184	214	307	1322	
P.A.(%)	98.8	94.3	97.8	98.1	97.7		97.6

producer's accuracy: M_{ii} / M_{+i} overall accuracy: $\Sigma M_{ii} / M_{++}$

consumer's accuracy: M_{ii} / M_{i+}

Konfusionsmatrix zur Überprüfung der Klassifikationsergebnisse

Zusammenfassung:

- Für multispektrale Klassifikation ist das a-priori Wissen des Anwenders nötig
- Unüberwachte Klassifikation basiert auf der Evaluation von Histogrammen und wird meist als pre-processing Methode angewandt.
- Ein Problem ist die statistische Verteilung in Klassen und die reale Welt
- Diese Methode liefert gute Resultate und kann sehr effektiv sein wenn sie mit Bedacht und Sorgfalt angewandt wird.

4.7 Binärbildverarbeitung

Binärbild

Ein Binärbild hat die Werte 0 und 1.

Wertebereich {1, 0}

{wahr, falsch}

{Vordergrund, Hintergrund}}

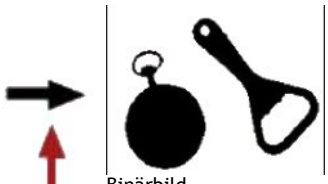
{Objekt, Nicht-Objekt}

Pixel mit dem Wert 1:
- relevante Information: gesetzte Pixel
- Gesamtmenge dieser Pixel: Bildvordergrund

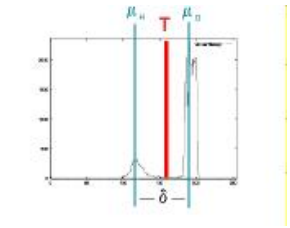
$$b(r,c) = \begin{cases} 1 & \text{falls Objekt} \\ 0 & \text{sonst} \end{cases}$$

Die folgende Abbildung zeigt, wie man von einem Grauwertbild zu einem Binärbild gelangt.

Grauwertbild



Binärbild

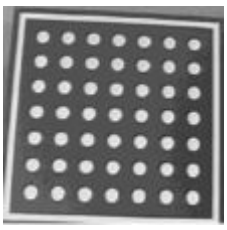


Histogramm

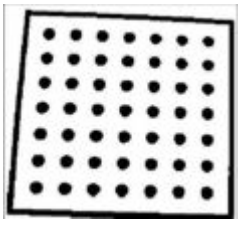
$$\delta = \mu_H - \mu_O$$

Die Binärbildverarbeitung wird häufig im Bereich industrieller Bildverarbeitung eingesetzt, wo es darauf ankommt, schnell Objekte im Bild zu selektieren, zu bestimmen und zu erkennen.

Beispiel: Bestimmung des Punktabstandes



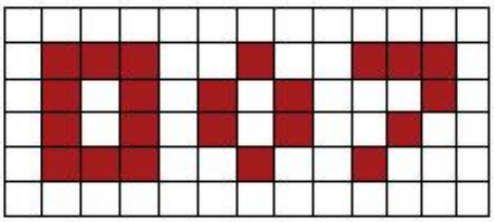
Grauwertbild



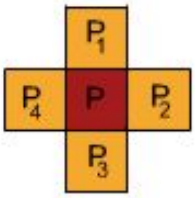
Binärbild

4.7.1 Topologie / Nachbarschaft

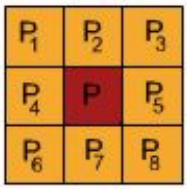
Beispiel:
Was hängt zusammen? Was als Linie bezeichnet wird, hängt von der Nachbarschaft ab.



In einem Rasterbild lassen sich mehrere Arten von Nachbarschaft definieren. Die wichtigsten sind:



4er - Nachbarschaft (N4):
alle Pixel, die horizontal oder vertikal an das gegebene Pixel P angrenzen



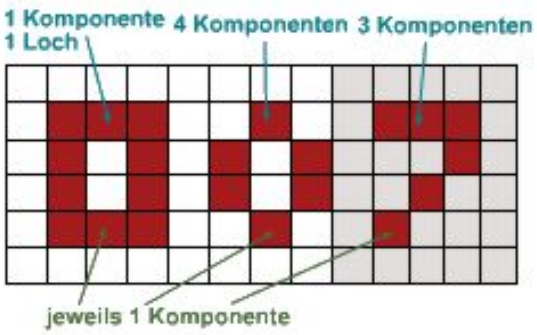
8er - Nachbarschaft (N8):
alle Pixel der 4er-Nachbarschaft plus alle Pixel, die diagonal an das gegebene Pixel P angrenzen.

Achtung: Je nach Art der Nachbarschaft werden Pixel als zusammenhängend bezeichnet.

Jordan'scher Kurvensatz: Jede einfach zusammenhängende Linie teilt die Ebene in zwei zusammenhängende Gebiete

Damit der Jordan'sche Kurvensatz erfüllt wird, wendet man eine Kombination der Nachbarschaft an.

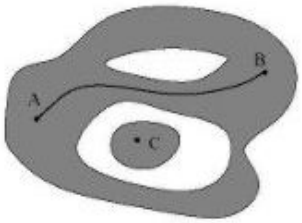
- **N4 für Vordergund, N8 für Hintergrund** oder
- **N8 für Vordergrund, N4 für Hintergrund**



4.7.2 Zusammenhangskomponenten

(engl. *connected components*)

Bei der Ermittlung von Zusammenhangskomponenten werden zusammenhängende Bereiche bestimmt. Bereiche sind dann zusammenhängend wenn ein Pfad (=eine Pixelsequenz) von einer Position zur anderen Position gelegt werden kann.



Von Punkt A nach Punkt B kann mindestens ein Pfad gelegt werden, ohne das Objekt zu verlassen.
wobei $A,B \in O; A \neq B$

Verfahren zur Bestimmung:

- 1. Verfahren: Grapheinfärbung
- 2. Verfahren: nach Rosenfeld/Pfalz 1966

1. Grapheinfärbung

- Setze Anzahl der Komponenten = 0
- Wiederhole
  1. Nächste Komponente K=K+1
  2. Anzahl Pixel in S={ }
  3. Suche nicht gekennzeichnete Vordergrundpixel
  (b(r,c)= 1; k(r,c)=0);
  setze S={(r,c)}, bezeichne k(r,c)=K
  4. Wiederhole
     1. Suche alle Nachbarn von S mit b(r,c)=1: N (S)
     2. Bezeichne alle diese mit k(r,c)=K
     3. S=S+N(S)
     Bis keine Nachbarn mehr gefunden werden

  Bis alle Pixel mit b(r,,c)=1 bezeichnet sind

Bemerkung:
→ Diese Methode nutzt die Nachbarschaftsbeziehungen nicht aus, die im Raster enthalten sind (im Gegensatz zur Rosenfeld/Pfalz).
→ Der 2. Schritt ist aufwendig, wenn die Anzahl der Komponenten hoch ist
→ Nachbarschaft N(S) kann bei komplexen Komponenten sehr groß werden

2. Methode von Rosenfeld/Pfalz 1966

- Zahl der Komponenten K=0; Äquivalenztabelle E={ }

  Für alle Zeilen r=0, ... N_r-1

  Für alle Spalten c=0, ..., N_c-1

  falls b(r,c)=1

  A=N(r,c)

  falls Anzahl Pixel A=0

  neue Komponente mit Nummer K=K+1

  bezeichne k(r,c)=1

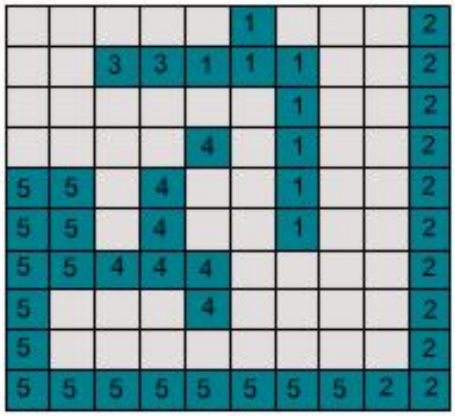
  falls Anzahl Pixel A>1

  k(r,c)=min (k(r,c) in A) (kleinste vergebene Nummer

  erweitere Äquivalenztabelle E=E+{K,k(r,c)}
- Vereinfache Äquivalenztabelle (=Grapheinfärbung)
- ersetze Bezeichner k(r,c) durch Bezeichner der Komponente von E

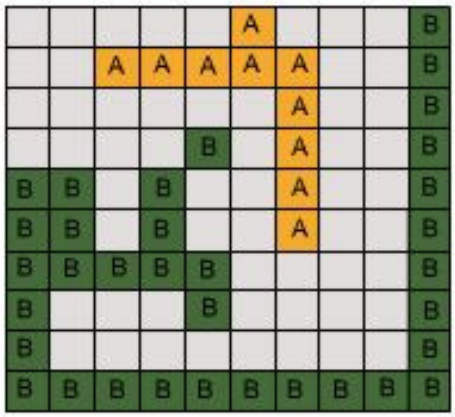
Bemerkungen:
→ ein Durchgang durch das Bild
→ Nutzung der regelmäßigen Nachbarschaftsstruktur
→ Führung einer Äquivalenztabelle
→ Grapeinfärbung des Äquivalenzgraphen

Beispiel:



Äquivalenztabelle
(1,3)
()4,5)
(2,5)

Äquivalenzgraph
(1,3)
(2,4,5)



Äquivalenztabelle
(1,3)
(4,5)
(2,5)

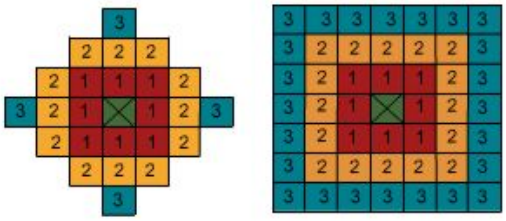
Äquivalenzgraph
(1,3): A
(2,4,5): B

4.7.3 Distanzen und Distanztransformationen

Distanzen zwischen den Pixeln können mit Hilfe der Euklidischen Distanz, Distanz mit 4er-Nachbarschaft (N4) oder 8er-Nachbarschaft(N8) bestimmt werden.

Der Abstand ist also abhängig von der Nachbarschaft.

Beispiel: digitale Kreise



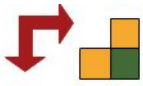
Beispiele
- minimaler Abstand eines Objektpixels vom Rand (=innere Distanz)
- minimaler Abstand eines Hintergrundpixels zum nächsten Objekt (=äußere Distanz)

Verfahren: **Distanztransformation**

Die Distanztransformation wird in 2 Durchgängen durchgeführt:

von links oben nach rechts unten

$$d(r,c) = \min(d(r,c), d(r,c-1)+1, d(r-1,c)+1)$$



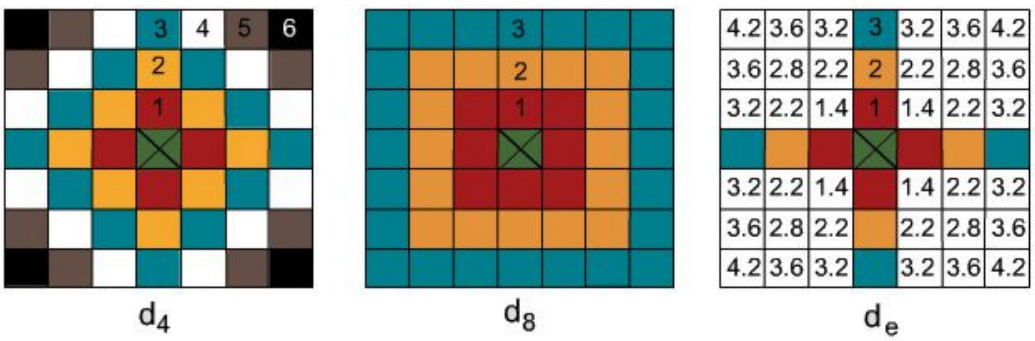
und

von rechts unten nach links oben

$$d(r,c) = \min(d(r,c), d(r,c+1)+1, d(r+1,c)+1)$$

Innere Distanz: MAX_DIST für das Objekt, 0 für Hintergrund
Äußere Distanz: MAX_DIST für Hintergrund, 0 für das Objekt.

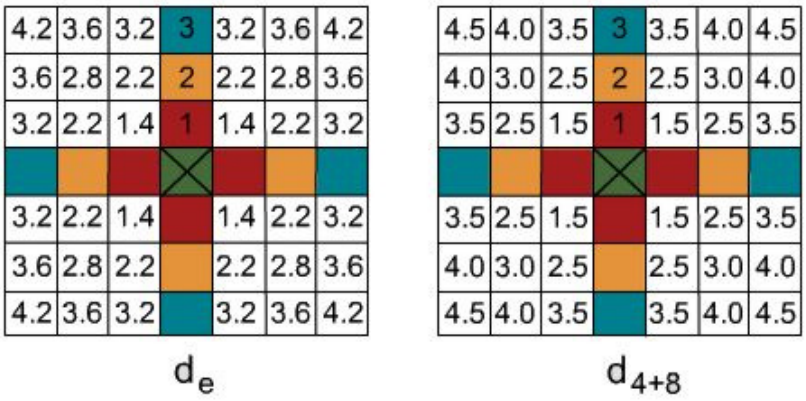
Je nachdem, mit welcher Nachbarschaft Distanzen berechnet werden, sind die Distanzen im Vergleich zur Euklidischen Distanz zu groß bzw. zu klein.



Vergleich der Distanzen: **1. N4-Nachbarschaft - d_4** **2. N8-Nachbarschaft - d_8** **3. Euklidische Distanz - d_e**

N4-Nachbarschaft: d4 - zu große Abstände
N8-Nachbarschaft: d8 - zu kleine Abstände

Um diesen Effekt zu umgehen, kann entweder gemittelt werden:
$d_{4+8} = 1/2\ (d_4 + d_8)$



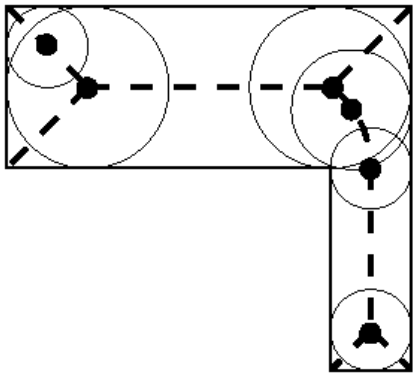
Vergleich zwischen Euklidischer Distanz d_e und den Mittelwertdistanzen von N4 und N8 ($=d_{4+8}$)

oder N8 durch einen Faktor modifiziert werden.

4.7.4 Skelettierung

Definition Skelett:

Menge aller Mittelpunkte von Kreisscheiben, die vollständig im Objekt liegen und dessen Rand mindestens zweimal berühren



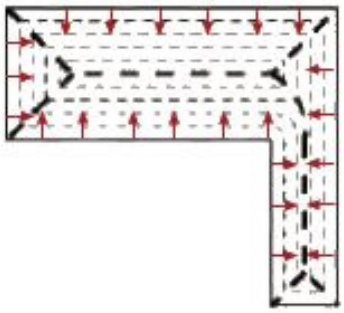
Skelett

Wird häufig in der Schrifterkennung verwendet. Es reduziert auf "topologische" Linien, die ein Pixel breit sind.

Das Objekt und das Skelett besitzen identische topologische Eigenschaften.

Berechnung der Mittelachse:
Steppenbrand-Algorithmus

Am Rand des Objekts werden von Norden, Süden, Westen und Osten Pixel nach bestimmten Kriterien abgeschält ("Feuer gelegt"). Dort wo die "Feuerfronten" aufeinandertreffen, befindet sich die Mittelachse (siehe Abbildung).



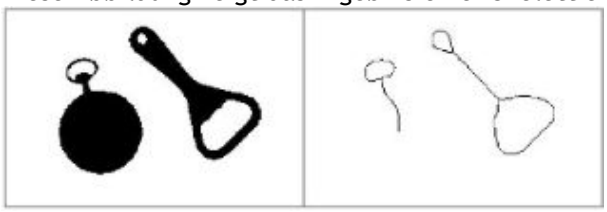
Schema des Steppenbrand-Algorithmus

Dabei müssen Endpixel erhalten bleiben und die Nachbarschaften beibehalten werden. Aufgrund des letzteren ist die Wahl der Anzahl der Nachbarn wichtig ("Wann ist ein Pixel Endpixel?") (siehe Abbildung unten).

Verfahren:

```
DO
{
   DO von N,S,W,O
   {
      Eliminiere Randpixel, wenn sie
      - nur eine Nachbarschaftskomponente haben
und
   - kein Endpixel
   sind
}
}  WHILE Änderungen
```

Diese Abbildung zeigt das Ergebnis einer Skelettierung:



Binärbild **Skelett**

Wann ein Pixel ein Endpixel ist, hängt von der Wahl der Nachbarschaft ab.
Daraus resultieren also für ein und dasselbe Bild unterschiedliche Ergebnisse.

4.7.5 Topologische und geometrische Merkmale

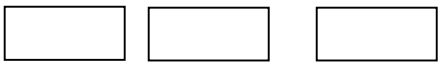
1. topologische Merkmale:

Topologische Merkmale sind:

- Anzahl Komponenten
- Anzahl Löcher
- Eulerzahl

Die **Anzahl der Komponenten** pro Objekt ist abhängig davon, was wir als Objekt auffassen.

Beispiel



a. wenn Objekt = Gebäude, dann hat jedes Objekt eine Komponente
b. wenn Objekt = Gebäudereihe, dann hat das Objekt mehrere Komponenten

3 Gebäide

Zu den topologischen Merkmalen gehört auch die Anzahl der **Löcher** pro Objekt

1 Komponente, 3 Löcher

Die **Eulerzahl** (Euler'sche Charakteristik) E ist definiert als die Differenz der Anzahl der Zusammenhangskomponenten eines Bildes K und der Anzahl der Löcher L.
Ein Loch bezeichnet die Zusammenhangskomponenten des Bildhintergrundes, die nicht mit dem Rand des Bildes verbunden sind.

$E = K - L$

Beispiel: K=1; L=3 E = 1-3 = -2

Sind in einem Bild nur Objekte ohne Löcher vorhanden, gibt die Eulerzahl die Zahl der Objekte an. Wenn die Eulerzahl für ein Objekt mit K=1 berechnet wird, so gibt sie die Zahl der Löcher in diesem Objekt wieder.
Für die Eulerzahl ist es notwendig, entweder 8er Nachbarschaft für den Bildvordergrund und 4er-Nachbarschaft für den Hintergrund oder umgekehrt zu verwenden.

2. Geometrische Merkmale:

Das sind z.B. Fläche, Umfang, Formfaktor k, Orientierung, Schwerpunkt, Verhältnis Länge zu Breite eines umschreibenden Rechtecks.

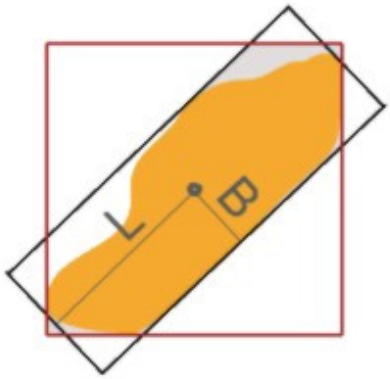
Der Faktor k gibt die Rundheit und Kompaktheit eines Objektes an.

$$k = \frac{U^2}{4\pi F}$$

wobei U=Umfang und F=Fläche

Da ein Kreis bei gegebener Fläche den geringsten Umfang aufweist, gilt: $k \geq 1$ (Kreis=1)

Umschreibendes Rechteck und gedrehtes Rechteck:
Verhältnis Länge L zu Breite B



Die Berechnung von einigen der angegebenen Größen erfolgt über die Bestimmung von Momenten und zentralen Momenten.

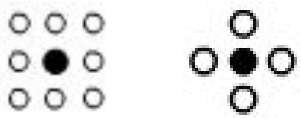
4.7.6 Mathematische Morphologie

Anwendungen:

- Bildverbesserung
- Vorverarbeitung von Segmentierung
- Kantendetektion
- Elimination von Störungen
- Texturanalyse
- Glättung von Kantenverläufen
- ...

Mathematische Morphologie wird hier unter dem Kaptitel Binärbildverarbeitung erklärt. Die mathematischen Operatoren können auch für eine Anwendung auf Grauwertbilder modifiziert werden.

Das strukturierende Element B definiert die Nachbarschaft für den Operator.

Beispiele 

Die Form des Elementes kann ein Rechteck, Quadrat, Kreis etc. sein.

Dilatation:

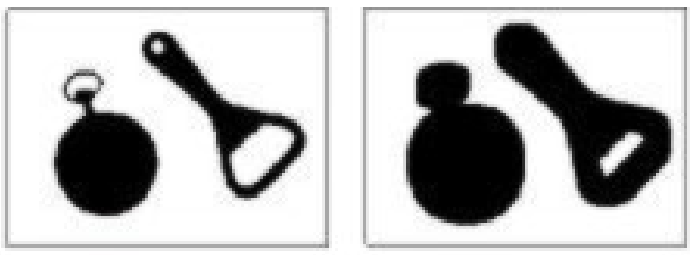
$$A_{dil} = A \oplus B$$

wobei A = Binärbild
B = strukturierendes Element

Die Elemente zweier Mengen werden addiert (*Minkowski Addition*): Wenn ein Bildpunkt in der durch das strukturierende Element B definierten Nachbarschaft der Wert des Bezugspunktes auf 1 gesetzt [Haberäcker 1991].

Die Dilatation bewirkt eine Vergrößerung aller Flächen.

Beispiel:



Originalbild **Dilatation**

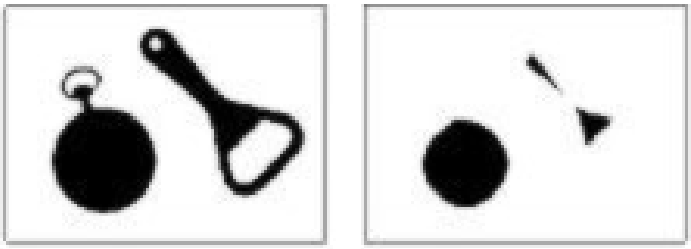
Erosion:

$$A_{ero} = A \ominus B$$

wobei A = Binärbild
B = strukturierendes Element

Es bleiben die Objektpixel des Ausgangsbildes A erhalten, deren Strukturpartner ebenfalls dem Objekt angehören, also den Wert 1 haben (logische UND-Verknüpfung). Ist mindestens ein Nachbar eines Bildelementes kein Objektpunkt, wird das Bildelement auf 0 gesetzt (=Hintergrund) (*Minkowski Subtraktion*) [Haberäcker 1991].
Das bewirkt ein Verkleinern aller Objektflächen.

Beispiel:



Originalbild **Erosion**

Dilatation und Erosion werden oft in Kombination eingesetzt. Erosion ist die entgegengesetzte Operation zu Dilatation.
Daraus ergeben sich folgende Operatoren:

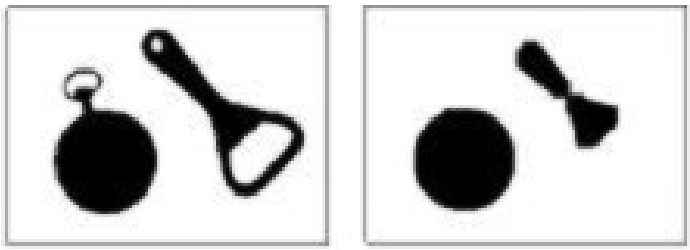
1. Opening ist die Kombination von Erosion mit anschließender Dilatation.

$$(A \ominus B) \oplus B = A \circ B$$

Eigenschaften des Opening:

- Elimination von Spitzen
- Elimination schmaler Verbindungen zw. Objektteilen
- Vergrößerung von Löchern

Beispiel:



Originalbild **Opening**

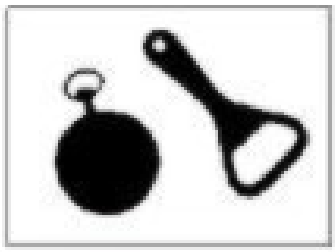
2. Closing ist die Kombination von Dilatation gefolgt von Erosion.

$$(A \oplus B) \ominus B = A \bullet B$$

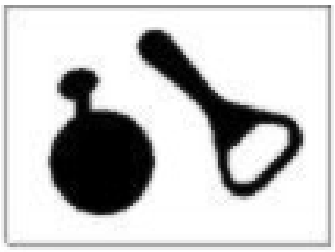
Eigenschaften des Closing:

- Verstärkung dünner Verbindungsstücke
- Elimination kleiner Löcher durch Auffüllen

Beispiel:



Originalbild



Closing

4.8 Literatur

Albertz, J. 2001: Einführung in die Fernerkundung. Wissenschaftliche Buchgesellschaft Darmstadt.

Bähr, H.-P. & Vögtle, T. 1991: Digitale Bildverarbeitung. 2. Auflage. Wichmann Verlag, Heidelberg.

Bähr, H.-P. & Vögtle, T. (Hrsg.) 1999: GIS for Environmental Monitoring. Schweizerbart, Stuttgart.

Bannari et al. 1995: A review of vegetation indices. In: Remote Sensing Reviews 13, pp. 95-120.

Bill, R.; Zehner, M. L. 2001: Lexikon der Geoinformatik. Wichmann Verlag, Heidelberg.

Gates D.M. 1970: Physical and physiological properties of plants. In: Remote Sensing with special reference to agriculture and forestry. National Academy of Science, Washington D.C.

Haberäcker, P. 1991: Digitale Bildverarbeitung: Grundlagen und Anwendungen. 4. Auflage. Hanser Verlag, München.

Kappas, M. 1994: Fernerkundung nah gebracht. Dümmler Verlag, Bonn.

Kraus, K. 1990: Fernerkundung, Bd.2 Auswertung photographischer und digitaler Bilder. Dümmler Verlag, Bonn.

Lillesand, T.M & Kiefer, R.W. 1994: Remote Sensing and Image Interpretation. 3rd ed., John Wiley & Sons, Chichester.

Mather, P.M. 1996: Computer processing of remotely-sensed images. John Wiley & Sons, Chichester.

Richards, J.A. 1994: Remote Sensing Digital Image Analysis: An Introduction. 2. Auflage. Springer Verlag, Heidelberg.