

5 Digitale Bildanalyse

5.1 Grundlagen

5.1.1 Bildanalyseprozess

Seit dem Aufkommen digitaler Bildanalysetechniken haben sich die Möglichkeiten der photogrammetrischen Bildauswertung in großem Umfang verändert. Digitale Bilddaten erlauben wie in vielen anderen technischen Bereichen auch hier neue Vorgehensweisen in der Erfassung, Speicherung, Archivierung und Ausgabe der Bilder, vor allem aber in der (automatisierten) Verarbeitung der Bilder selbst.

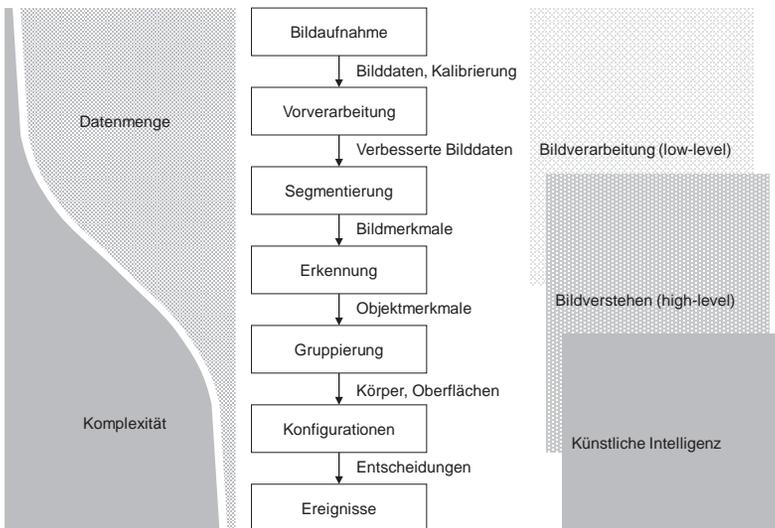


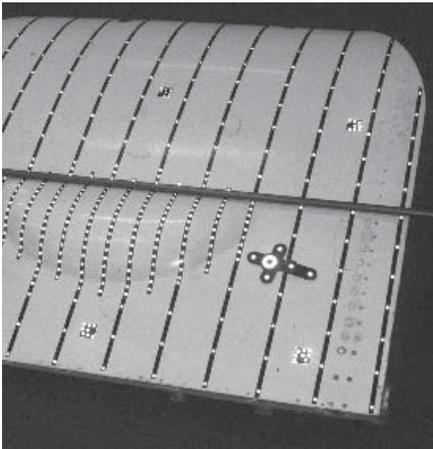
Abb. 5.1: Der Bildanalyseprozess

Abb. 5.1 zeigt den üblichen Prozess der technischen Bildanalyse, der von der Bildaufnahme bis zur intelligenten Auslösung von Ereignissen führt. Von oben nach unten nimmt die zu bewältigende Datenmenge ab, während die Komplexität der Verarbeitung zunimmt.

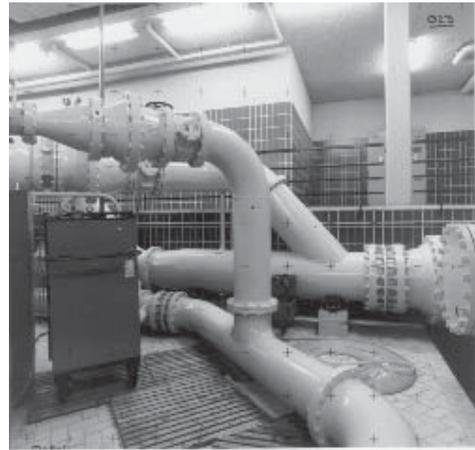
Photogrammetrische Methoden der Bildanalyse werden überwiegend in den Bereichen der Bildaufnahme (Sensortechnik, Kalibrierung), Vorverarbeitung und Segmentierung (Bildmessung, Merkmalerkennung, Merkmalszuordnung) entwickelt und eingesetzt. Bei diesen Methoden stehen die Automation relativ einfacher Messaufgaben und die Sicherstellung einer ausreichenden Robustheit und Genauigkeit im Vordergrund. Praxisreife Verfahren findet man daher überwiegend in Standardaufgaben der digitalen Punktmessung (signalisierte Objektpunkte) sowie bei Anwendungen der Stereophotogrammetrie (automatische Orientierung, Oberflächenmessungen, Fahrerassistenzsysteme). Methoden des Bildverstehens (Objekterkennung, Semantik) werden zunehmend erfolgreich im Zusammenhang mit lernenden Verfahren (deep learning, machine learning) eingesetzt.

In der Nahbereichsphotogrammetrie lassen sich zwei charakteristische Merkmale hervorheben, die Möglichkeiten und Grenzen der automatischen Bildauswertung besonders kennzeichnen:

- Konsequente Objektsignalisierung mit (teilweise) codierten, kontrastreichen Punktmarken führt bei geeigneter Belichtung zu quasi-binären Bildern, die sich praktisch vollautomatisch auswerten lassen (Abb. 5.2a).
- Beliebige Aufnahmeanordnungen führen zu größeren Maßstabsunterschieden im Bildverband, Verdeckungen, unvollständiger Objektabbildung usw. (Abb. 5.2b). Im Gegensatz zu einfachen Stereoanordnungen, wie sie z.B. in der Luftbildauswertung für Senkrechtaufnahmen vorliegen, sind Nahbereichsanwendungen also häufig durch komplexe Objektflächen und Bildverbände gekennzeichnet, die teilweise einen hohen interaktiven Anteil an der Auswertung erfordern.
- Bildanalyseverfahren und Orientierungsmethoden aus dem Bereich des Computer Vision eignen sich zunehmend zur vollautomatischen Orientierung von (hinreichend überlappenden) Bildsequenzen aus beliebigen (unkalibrierten) Kameras. Diese Verfahren (z.B. Structure-from-Motion) eröffnen interessante Potenziale, z.B. zur automatischen Näherungswertbestimmung oder zur vollautomatischen 3D-Rekonstruktion.



a) signalisiert



b) nicht signalisiert

Abb. 5.2: Beispielhafte Objektszenen der Nahbereichsphotogrammetrie

Der vorliegende Abschnitt befasst sich vor allem mit solchen Bildanalyseverfahren, die in der Praxis der Nahbereichsphotogrammetrie bereits erfolgreich eingesetzt werden. Der Schwerpunkt liegt auf Verfahren zur geometrischen Bildverarbeitung, Bildmessung und 3D-Rekonstruktion mit dem Ziel, eine hohe Qualität des Ergebnisses zu erzielen. Weiterführende Grundlagen und spezielle Algorithmen können der zahlreich vorhandenen Literatur zu Bildverarbeitung und Computer Vision entnommen werden, siehe Kap. 9.5.

5.1.2 Pixelkoordinatensystem

Zur Lösung von messtechnischen Bildverarbeitungsaufgaben ist die Definition des verwendeten *Pixelkoordinatensystems* von grundlegender Bedeutung (siehe Kap. 2.1.1). Für ein Bild

$$S = s(x, y) \quad (5.1)$$

wird in der Bildanalyse meistens ein bildschirmorientiertes linkshändiges xy -System von Zeilen und Spalten verwendet, in dem die x -Richtung in Zeilenrichtung und die y -Richtung in Spaltenrichtung definiert ist.

Der Ursprung dieses Systems liegt oben links, das erste Bildelement besitzt die Zeilen- und Spaltenkoordinaten $(0,0)$. Die Anzahl der Spalten beträgt m , die Anzahl der Zeilen n . Das letzte Bildelement besitzt die Koordinaten $(m-1, n-1)$. Breite und Höhe eines Pixels betragen jeweils 1 (Abb. 5.3).

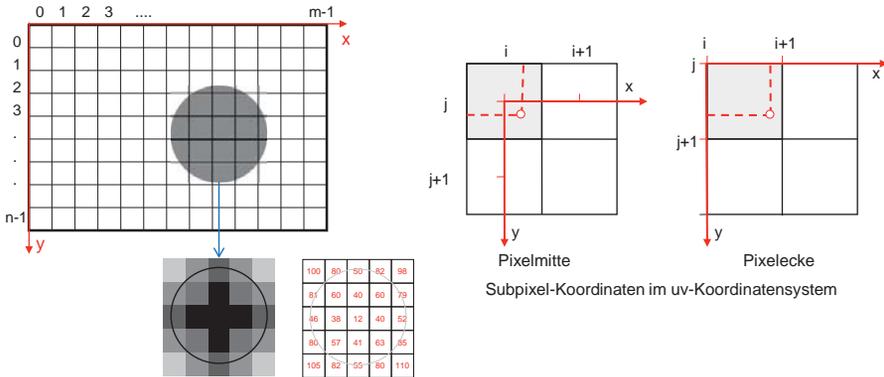


Abb. 5.3: Definition des Pixelkoordinatensystems

In diesem diskreten Raster hat jedes Bildelement ganzzahlige Koordinaten. Bei der optischen Abbildung eines Gegenstandes auf einen Sensor mit diesem Rastersystem erhält jedes Bildelement einen Intensitätswert (Grauwert¹) entsprechend der auf seine Fläche einwirkenden Helligkeit des Objektes. Die Grauwerte werden in der Regel mit 8 Bit quantisiert, sodass 256 Graustufen zwischen 0 (schwarz) und 255 (weiß) definiert werden können. Da der Mensch nur ca. 60 Grautöne unterscheiden kann, reicht diese Grauwerttiefe zur visuellen Darstellung von Bildern aus. In technischen Systemen kann der Informationsgehalt jedoch erheblich größer sein, sodass auch Quantisierungen mit 10 Bit (1024 Graustufen), 12 Bit oder 16 Bit pro Grauwert sinnvoll sein können. Echtfarbbilder werden im Allgemeinen mit 24 Bit (je 8 Bit pro RGB-Farbkanaal) abgespeichert.

Da ein abgebildetes Objekt (z.B. ein signalisierter Objektpunkt) durch den Abbildungsmaßstab und die optischen Übertragungseigenschaften (MTF, PSF) des Aufnahmesystems mehr als ein Pixel überdecken kann, können die Koordinaten dieses Objektes im Subpixelbereich

¹ Im Folgenden werden die Begriffe Grauwert und Farbwert gleichwertig verwendet.

verschoben sein. Die Angabe von Subpixelkoordinaten ist dann möglich, wenn die Lage eines Objektpunktes über mehrere Pixel interpoliert werden kann. Hierbei liegt die Annahme zugrunde, dass die Subpixelverschiebung einer Objektkante zu einer entsprechenden Veränderung der Grauwerte des abgebildeten Objektes führt. Bei der Angabe von Subpixelkoordinaten geht man in der Regel davon aus, dass die obere linke Ecke (seltener die Mitte des Pixels) der ganzzahligen xy -Koordinate (i,j) entspricht (Abb. 5.3).

Abgebildete Objekte müssen in der Regel mehrere Pixel überdecken, um im Bild erkannt oder verarbeitet werden zu können. Zu einem gemeinsamen Objekt gehörende, zusammenhängende Pixel zeichnen sich dadurch aus, dass ihre Grauwerte in einer begrenzten Nachbarschaft bestimmten einheitlichen Kriterien genügen müssen (*Konnektivität*). Im diskreten Bildraster besitzt jedes Pixel eine feste Zahl von Nachbarpixeln (Ausnahme: Bildrand). Zur Definition von Konnektivitäten werden benachbarte Pixel nach dem N4- oder dem N8-Schema gekennzeichnet (Abb. 5.4). In diesem Beispiel sind drei Objekte A, B und C abgebildet. Wird N4-Nachbarschaft zugrunde gelegt, zerfällt Objekt B in einzelne Pixel und die Objekte A und C sind getrennt. In der N8-Relation ist Objekt B ein zusammenhängender Bereich, über die Eckverbindung werden jedoch auch A und C zusammengelegt. Erweiterte Algorithmen zur Konnektivität müssen daher auch die Grauwertverteilung innerhalb von Teilregionen betrachten, z.B. durch geeignete Filter.

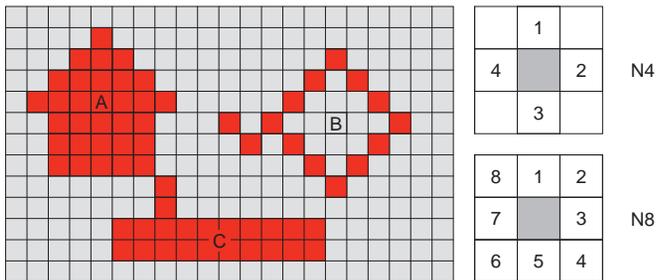


Abb. 5.4: Objekte und Nachbarschaften

5.1.3 Bilddatenverwaltung

5.1.3.1 Bildpyramiden

Unter einer *Bild-* oder *Auflösungspyramide* versteht man eine Bildserie, in der das jeweils folgende Bild gegenüber dem Vorgänger verkleinert ist, in der Regel um den Faktor 2 (Abb. 5.5). Vor der Verkleinerung sollte das Bild geglättet werden, zum Beispiel mit Gauß-Filtern (vgl. Kap. 5.2.3.2). Kleine Bildstrukturen verschwinden dabei nach und nach mit abnehmender Bildauflösung, der Informationsgehalt sinkt mit zunehmender Verkleinerung (Abb. 5.6). Die Datenmenge der gesamten Bildpyramide wächst nur um ca. 30% gegenüber dem Originalbild. Bildpyramiden dienen vor allem zur hierarchischen Mustererkennung, bei der zunächst in der geringsten Auflösung nach groben Bildmerkmalen gesucht wird. Mit zunehmender Auflösung (Bildgröße) wird die Suche verfeinert, wobei nur noch die in der vorhergehenden Stufe ermittelten Bildbereiche näher untersucht werden.



Original



Gauß-Tiefpassfilter



Faktor 2



4



8



16

Abb. 5.5: Bildpyramide mit 5 Stufen

Faktor 4



Faktor 16

Abb. 5.6: Informationsgehalt der Auflösungsstufen Faktor 4 und 16

5.1.3.2 Datenformate

Die Speicherung digitaler Bilddaten kann auf vielfältige Weise organisiert werden. In den Bereichen der digitalen Bildverarbeitung und der Rastergrafik wurden dazu zahlreiche Datenformate entwickelt, die in der Lage sind, neben den eigentlichen Bilddaten weitere Informationen wie Bildbeschreibungen, Farbtabelle, Übersichtsbilder usw. abzuspeichern. Für photogrammetrische Belange ist dabei die eindeutige Reproduzierbarkeit des Original-

bildes von wesentlicher Bedeutung. Datenverluste können durch den Einsatz von Datenkompressionsverfahren, aber auch durch Verwendung unzureichender Grauwerttiefen (Bits pro Pixel) oder Zuweisung falscher Farbtabelle entstehen.

Aus der großen Vielfalt von Bildformaten seien hier nur einige Standardformate genannt:

- **Speicherung von Rohdaten:**

Hierbei werden nur die Farbwerte des Bildes unkomprimiert in einer Binärdatei gespeichert. Bei einem Byte pro Pixel entspricht die entstehende Dateigröße exakt der Gesamtzahl von Pixeln im Bild. Zum Einlesen der Rohdaten müssen Informationen über die Größe des Originalbildes separat gespeichert oder bereitgestellt werden. Für Mehrkanalbilder kann die Reihenfolge der Spektralbereiche pixelweise (pixel-interleaved), zeilenweise (line-interleaved) oder kanalweise (band-interleaved) erfolgen. Einige Kamerahersteller bieten eigene Formate zur Speicherung der Rohbilder an.

- **TIFF: Tagged Image File Format:**

Das TIFF-Datenformat hat sich durch seine universellen Einsatzmöglichkeiten weit verbreitet. Es basiert auf einem zeigerorientierten Verzeichnis (directory) von wichtigen Bildinformationen (Bildgröße, Farbtiefe, Paletten, Auflösung etc.). TIFF erlaubt verschiedene Methoden zur verlustfreien Bildkompression (LZW, Huffman).

- **BMP: Windows Bitmaps:**

Das BMP-Format ist der Standard für Rasterbilder in der Microsoft-Windows-Umgebung. Es erlaubt die verlustfreie Abspeicherung beliebiger Grau- und Farbbilder (bis 24 Bit) mit unterschiedlicher Anzahl von Grau- oder Farbstufen.

- **GIF: Graphic Interchange Format:**

Bilder im GIF-Format sind standardmäßig verlustfrei komprimiert. Grau- und Farbbilder können nur bis zu 8 Bit pro Pixel enthalten.

- **PNG: Portable Network Graphics:**

Das PNG-Format gilt als Nachfolger des GIF-Formates und bietet diesem gegenüber umfangreichere Transparenzeinstellungen (Alpha-Kanal), RGB-Bildspeicherung und eine höhere Kompressionsrate.

- **JPEG: Joint Photographic Expert Group:**

Das JPEG-Format bietet große Kompressionen bis zum Faktor 100, wobei bewusst ein Datenverlust in Kauf genommen wird (siehe Kap. 5.1.3.3). Mit JPEG2000 steht ein erweiterter Standard zur Verfügung.

Abb. 5.7 zeigt die Kompressionsraten verschiedener Datenformate für zwei Bilder mit sehr unterschiedlicher Struktur. Das Bild „Mosaik“ (Abb. 5.9, Ausschnitt aus Abb. 5.24, blaue Balken) kann mit seinen inhomogenen Strukturen nur auf ca. 78% der Originalgröße verlustfrei komprimiert werden (PNG), während das homogenere Bild „Marken“ (Abb. 5.10, rote Balken) bis 40% komprimiert werden kann. Die verlustbehaftete JPEG-Kompression erlaubt dagegen noch deutlich höhere Reduktionen, wobei die Bildqualität ab einem Faktor von 50 deutlich abnimmt.

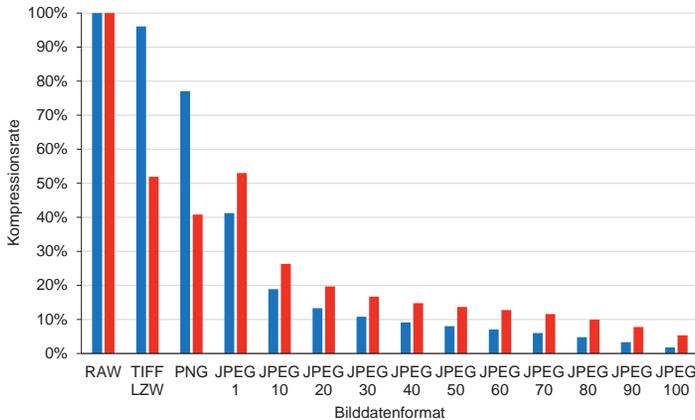


Abb. 5.7: Kompressionsraten verschiedener Bildformate
(blau: Bild „Mosaik“, Abb. 5.9; rot: Bild „Marken“, Abb. 5.10)

5.1.3.3 Bildkompression

Die Kompression von Bilddaten hat bei den in der digitalen Photogrammetrie heute schon üblichen Datenmengen eine große praktische Bedeutung. So benötigt ein monochromatischer Bildverband von 50 Bildern à 3000 x 2000 Pixel eine Rohdatenmenge von ca. 300MB, während ein UAV-Bildflug mit 400 Farbbildern à 4000 x 3000 Pixel mehr als 1.4TB Speicherplatz für die Rohdaten benötigt.

Laufängencodierte Kompressionsverfahren zählen die Anzahl identischer Grauwerte in einer Zeile oder Fläche und codieren den entsprechenden Bildbereich durch Angabe des Grauwertes und des Wiederholungsfaktors. Sie eignen sich nur bei Bildszenen mit größeren homogenen Bereichen und führen bei natürlichen Bildszenen häufig sogar zu einer Vergrößerung der Datenmenge.

Frequenzbasierte Kompressionsverfahren unterziehen das Bild einer Spektralanalyse (Fourier-, Cosinus- oder Wavelet-Transformation, siehe Kap. 5.2.3.1) und speichern die Koeffizienten der zugehörigen Linearkombinationen. Durch Elimination nicht signifikanter Koeffizienten kann eine verlustbehaftete Datenreduktion vorgenommen werden.

Das JPEG-Bildformat basiert ebenfalls auf einer verlustbehafteten Datenkomprimierung. Ziel ist die Erhaltung des wesentlichen Bildinhaltes, sodass auch bei großen Kompressionsraten nur ein geringer visueller Qualitätsverlust bemerkbar ist. JPEG hat sich aufgrund der hohen Kompressionsleistung in fast allen Bereichen durchgesetzt. Das Verfahren ist standardisiert.

Die Kompressionsmethode basiert auf dem dreistufigen *Baseline-Sequential*-Verfahren (Abb. 5.8). In disjunkten 8x8-Pixelblöcken wird eine Diskrete Kosinus-Transformation (DCT) berechnet, deren Koeffizienten durch eine (wählbare) Quantisierungstabelle gewichtet und laufängencodiert gespeichert werden. Die Dekomprimierung erfolgt durch den inversen Vorgang. Für Farbbilder erfolgt zusätzlich eine IHS-Farbtransformation.

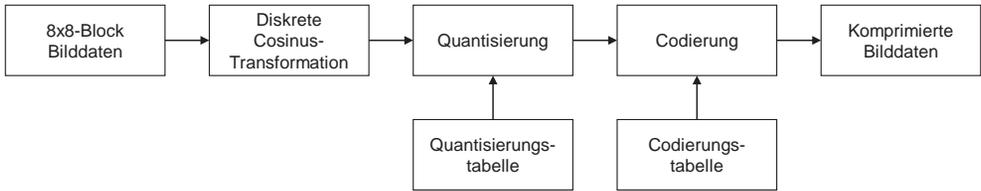


Abb. 5.8: JPEG-Bildkompression

Der eigentliche Datenverlust wird durch die Wahl der Intervalle in der Quantisierungstabelle gesteuert. Sie ist in der Regel so aufgebaut, dass bei visueller Bildbetrachtung kein wesentlicher Informationsverlust auffällt (Abb. 5.9 und Abb. 5.10).

Mit JPEG2000 stehen einige Erweiterungen des JPEG-Formates zur Verfügung. Die Kosinus-Transformation wurde durch eine waveletbasierte Kompression ersetzt. Weiterhin können verschiedene Bildausschnitte unterschiedlich stark komprimiert werden. JPEG2000 unterstützt transparente Bilder.

Die Auswirkungen auf photogrammetrische Messungen hängen stark vom Bildinhalt ab. Generell verursacht die JPEG-Kompression Lageverschiebungen im Bild, die Beträge von 0.1 bis mehr als 1 Pixel annehmen können und damit das mögliche Genauigkeitspotenzial automatischer Punktmessungen von 0.02–0.05 Pixel deutlich überschreiten (Kap. 5.4.2.6). Zusätzlich kann es durch die 8x8-Blockbildung zu unerwünschten Effekten kommen.

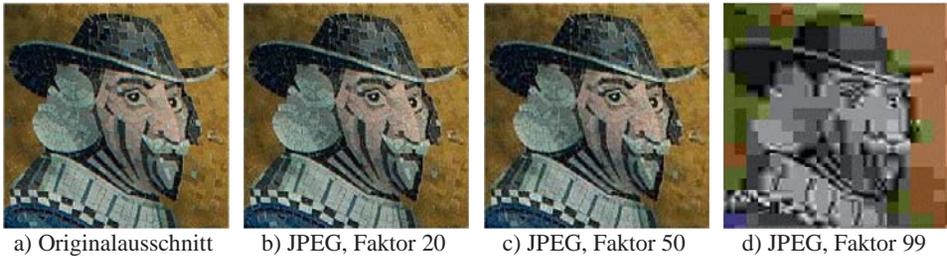


Abb. 5.9: Auswirkungen verlustbehafteter Bildkompression im Bild „Mosaik“ (vgl. Abb. 5.7)

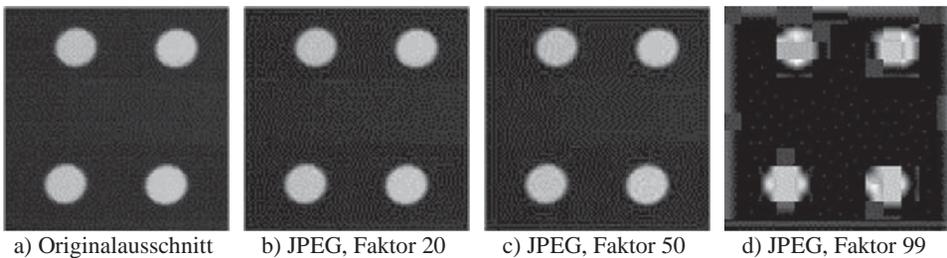


Abb. 5.10: Auswirkungen verlustbehafteter Bildkompression im Bild „Marken“ (vgl. Abb. 5.7)