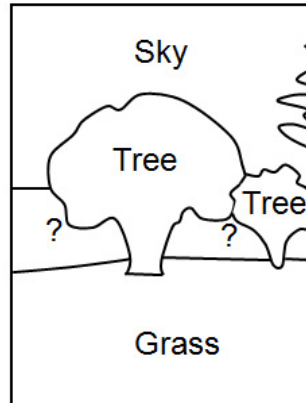


15. Image Segmentation

Segmentierung

Unter Segmentierung versteht man eine Operation, die für jeden einzelnen Bildpunkt entscheidet, ob er zu einem Objekt gehört oder nicht. Dabei entsteht ein Binärbild (bei nur einem Objekt im Bild), ein Bild mit mehreren eindeutige Nummern (Identifikatoren oder *Labels*). Den Pixeln des Hintergrundes wird eine Null zugewiesen. Die Abbildung links zeigt ein Beispiel, die Objekte Baum, Himmel und Wiese werden als zusammenhängende Pixelbereiche erkannt. Die Segmentierung steht an der Grenze zwischen der Bildverarbeitung und der Bildanalyse, da sie zusammenhängende Pixelbereiche (Objekte) extrahiert, die Erkennung dieser Objekte (z.B. ob es sich um Gras oder einen Baum handelt) zählt hingegen nicht zu den Aufgaben der Segmentierung. Nach der Segmentierung wissen wir also, welcher



Bildpunkt zu welcher segmentierten Region gehört. Damit ist ein Bild in Regionen eingeteilt, und wir kennen die Diskontinuitäten als die Ränder der Regionen. Segmentierungsverfahren lassen sich generell in 3 Klassen einteilen: **Punktorientierte Verfahren** verwenden nur die Grauwerte der einzelnen Pixel. **Kantenorientierte Verfahren** erkennen Kanten und versuchen ihnen zu folgen. **Regionenorientierte Verfahren** untersuchen Grauwerte in zusammenhängenden Regionen.

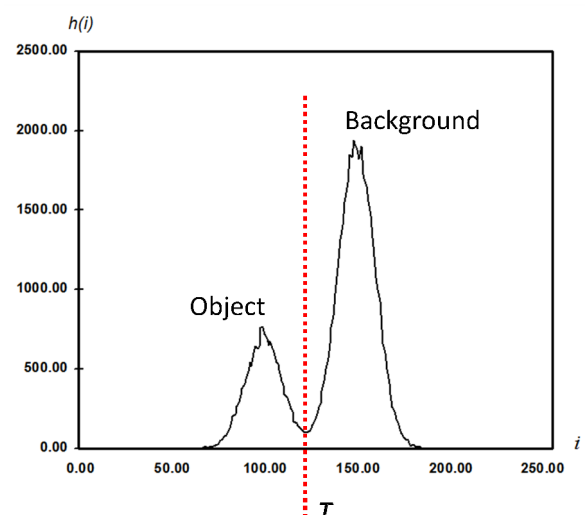
Bildpunkt zu welcher segmentierten Region gehört. Damit ist ein Bild in Regionen eingeteilt, und wir kennen die Diskontinuitäten als die Ränder der Regionen. Segmentierungsverfahren lassen sich generell in 3 Klassen einteilen: **Punktorientierte Verfahren** verwenden nur die Grauwerte der einzelnen Pixel. **Kantenorientierte Verfahren** erkennen Kanten und versuchen ihnen zu folgen. **Regionenorientierte Verfahren** untersuchen Grauwerte in zusammenhängenden Regionen.

Punktorientierte Segmentierungsverfahren

Die punktorientierte oder pixelbasierte Segmentierung stellt vom Ansatz her die einfachste Methode dar. Das Segmentierungskriterium bei der punktorientierten Segmentierung ist die Intensität eines isoliert betrachteten Pixels. Dazu verwendet man die Grauwertverteilung eines Bildes mit Hilfe des Histogramms und versucht, in diesem eine oder mehrere Schwellwerte zu finden. Ist das zu findende Objekt einfarbig und vom Hintergrund verschieden, so ergibt sich ein bimodales Histogramm mit räumlich getrennten Maxima. Idealerweise gibt es einen Bereich des Histogramms, der keine Punkte enthält. Durch Setzen einer Schwelle an dieser Stelle kann das Histogramm in mehrere Bereiche unterteilt werden. Es wäre auch denkbar, auf der Basis von Farbwerten zu segmentieren. Im dann betrachteten dreidimensionalen Raum muss das Verfahren aber so modifiziert werden, dass Punktwolken voneinander abgegrenzt werden können. Das Finden eines Schwellwerts im eindimensionalen Raum der Grauwerte ist erheblich einfacher, wobei jedoch anstelle der Grauwerte jedes andere Merkmal, dass eine Unterscheidung zulässt, verwendet werden kann.

Haben wir ein gutes Merkmal zur Trennung eines Objekts vom Hintergrund gefunden, zeigt das Histogramm der Grauwerte (oder allgemeiner der Merkmalswerte) eine bimodale Verteilung mit zwei getrennten Maxima wie in der Abbildung rechts. Wir können nicht erwarten, dass ein Bild eines einfarbigen Objektes vor einem einfarbigen Hintergrund nur 2 Werte aufweist, selbst wenn scharfe Objektkanten vorliegen, wird es wegen des optischen Systems und Bildsensors dazwischen liegende Grauwerte geben.

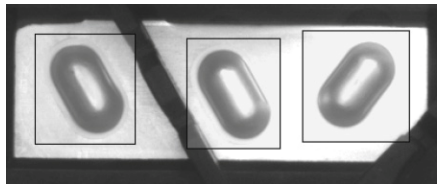
In der Abbildung ist der Schwellwert T so gewählt, dass er an der Talsohle der bimodalen Funktion zu liegen kommt, T trennt somit die beiden Grauwertregionen ideal. Die Größe der segmentierten Objekte ändert sich natürlich mit einer Variation des Schwellwertes deutlich. Ein korrekt gewählter Schwellwert ist also kritisch für eine fehlerfreie Bestimmung der geometrischen Merkmale eines Objekts. Dies ist ohne Wissen über den Typ der Kante zwischen Objekt und Hintergrund nicht möglich. Im Falle einer symmetrischen Kante entspricht der korrekte Schwellwert dem mittleren Grauwert aus Hintergrund- und Objektbildpunkten.



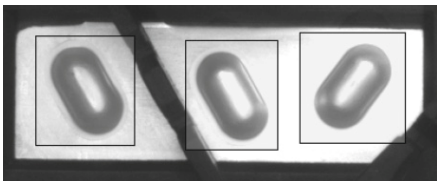
Schwellwerte

Wie oben ausgeführt, ist die Wahl des Schwellwertes inhärent mit der Qualität des Ergebnisses verknüpft. Die bisher geschilderte Schwellwertwahl nennt man **globaler Schwellwert**, der nur funktioniert,

wenn konstante Helligkeitsverteilungen im Bild existieren. Ist dies nicht der Fall, versagt die Methode wie in der Abbildung gezeigt, die Objekte fransen aus. Weder Bilder mit lokal unterschiedlicher Beleuchtung lassen sich so segmentieren, noch eignet sich ein globaler Schwellwert, wenn in einem Bereich des Bildes das gesuchte Objekt einen anderen Kontrast zum Hintergrund aufweist als in einem anderen.



Durch die Verwendung von **lokalen (ortsabhängigen) Schwellwerten** lässt sich dieser Nachteil vermeiden. Für jeden Bereich wird dabei ein eigener Schwellwert berechnet. Eine Methode wurde von *Chow* und *Kaneko* dafür entwickelt. Das Bild wird in eine Menge von überlappenden Teilbildern geteilt und in jedem Teilbild



wird dann ein optimaler Schwellwert durch Analyse des lokalen Histogramms bestimmt. Für jedes einzelne Pixel wird der ideale Schwellwert durch Interpolation aller Ergebnisse der Teilbilder bestimmt. Problematiken entstehen auch hier wieder, falls ein Teilbild nicht bimodal ist. Eine wichtige Größe ist hierbei die Standardabweichung σ der Grauwerte im Teilbild. Ist sie besonders klein, kann man darauf schließen, dass sich das Teilbild in einer annähernd homogenen Region befindet, sprich komplett im Hintergrund oder im Objekt liegt. Weist sie hingegen einen hohen Wert auf, ist davon auszugehen, dass sich Objekt und Hintergrund im Bildausschnitt befinden. Die Grundidee des lokal adaptiven Verfahrens ist es, die Teilbilder solange zu vergrößern, bis sich die Standardabweichung des aktuell betrachteten Teilbildes der Standardabweichung des gesamten Bildes σ_{ges} genähert hat. Somit erreicht man auch durch die lokalen Verfahren nur dann eine optimale Trennung, wenn gewährleistet werden kann, dass jede Region Objekt und Hintergrund enthält. Lokale Schwellwertverfahren eignen sich am besten für schlecht ausgeleuchtete Aufnahmen, in denen viele kleine Objekte verteilt sind, wie z.B. Mikroskopaufnahmen.

Kantenorientierte Segmentierungsverfahren

Punktorientierte Segmentierung führt selbst bei perfekter Beleuchtung zu einem Fehler der Größe segmentierter Objekte, wenn die Objekte Variationen in ihren Grauwerten zeigen. Die Ursache für diesen Fehler liegt darin, dass sich die Grauwerte an den Objektkanten nur allmählich vom Hintergrund zum Objektwert ändern. Die korrekte Bestimmung des Schwellwertes zur Trennung der Objekte vom Hintergrund ist somit schwierig. Mit einer kantenbasierten Segmentierung lässt sich dieser Fehler vermeiden, da die Position einer Kante durch den Maximalwert der Ableitung erster Ordnung oder einen Nulldurchgang der Ableitung zweiter Ordnung charakterisiert ist. Eine kantenbasierte Segmentierung liefert auch selbst dann eine korrekte Kantenlage, wenn die Bildhelligkeit linear variiert, da sie nicht von der Intensität der Kante abhängt. Zur Kantendetektion wird ein Bild Zeile für Zeile nach Maxima des Betrags des Gradienten abgetastet. Daraufhin versucht ein Konturverfolgungsalgorithmus, dem Maximum des Gradienten um das Objekt herum zu folgen, bis der Ausgangspunkt wieder erreicht ist. Somit berücksichtigt die kantenorientierte Segmentierung, dass ein Objekt eine zusammenhängende Region ist. Einfache Algorithmen zur Konturverfolgung verbinden Kantensegmente, die nicht weiter als eine Distanz ε voneinander entfernt sind. Komplexere Algorithmen verbinden Kantensegmente, indem zusätzlich die Richtung der Kante in die Berechnung miteinbezogen wird. Die Kante wird also möglichst ihrem Verlauf entsprechend fortgesetzt. Andere Algorithmen verwenden die Hough-Transformation (siehe Kap. 14). Im Bild kann so der Anteil der darin enthaltenen Geraden bestimmt werden. In der Praxis sind kantenorientierte Verfahren nur halbautomatisch realisierbar.



Regionenorientierte Segmentierungsverfahren

Regionenorientierte Verfahren bringen einen neuen Aspekt in den Segmentierungsprozess. Bei den punktorientierten Verfahren wird ein Bildpunkt aufgrund seines Grauwertes dem Objekt zugeordnet, unabhängig davon, was mit den

Nachbarnpixeln geschieht. Damit können, losgelöst vom eigentlichen Objekt, einzelne isolierte Punkte oder kleine Bereiche entstehen. Das entscheidende Merkmal eines Objekts ist aber, dass es zusammenhängend ist, und daher sollte die Nachbarschaft des Pixels bei dessen Zuordnung zu einem Objekt miteinbezogen werden.

Relaxation Labelling

Bei diesem Verfahren erhält jedes Pixel ein Label mit einer vorgegebenen Wahrscheinlichkeit: $p(i)$ ist die Wahrscheinlichkeit, dass Pixel i zum Hintergrund gehört, $(1-p(i))$ die Wahrscheinlichkeit, dass es zum Vordergrund gehört. Benachbarte kompatible Pixel unterstützen sich mit Hilfe von fix festgelegten Konsistenzwerten c_p und c_n . Sind Labels gleich haben sie eine positive Konsistenz, d.h., sie unterstützen sich, andernfalls unterstützen sie sich nicht. Dies wird im Relaxationsprozess gesteuert, die Zuordnungswahrscheinlichkeiten ändern sich iterativ mit dem Maß der Unterstützung, wobei die Vordergrundpixel auf Wahrscheinlichkeit $p(i)=0$ und die Hintergrundpixel auf $p(i)=1$ hinarbeiten.

Zu Beginn werden die Wahrscheinlichkeiten aller Pixel initialisiert, z.B. als das Verhältnis der Grauwerte $g(i)$ zum maximalen Grauwert g_{max} : $p^{(0)}(i) = g(i) / g_{max}$. Die 8 Nachbarschaftswahrscheinlichkeiten $p(i_1) \dots p(i_8)$ sowie die Konsistenzwerte ergeben die Wahrscheinlichkeitsberechnung, die Wertigkeit des betrachteten Pixels wird erhöht (inkrementiert), wenn der Wahrscheinlichkeitswert des Nachbarn größer als 0,5 ist, andernfalls verringert. Dieses Inkrement q ist gegeben durch:

$$q(i_1) = c_p p(i_1) + c_n (1 - p(i_1))$$

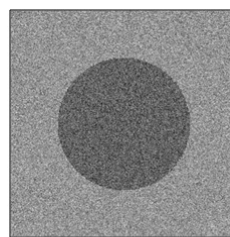
Die Gesamtwert der Inkremente $\Delta p(i)$ ergibt sich dann durch den Durchschnitt der Inkremente aller 8 Nachbarn. Danach wird der Vorgang iteriert:

$$p^{(r)}(i) \sim p^{(r-1)}(i)(1 + \Delta p(i))$$

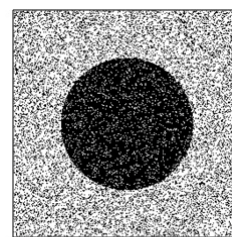
Die Wahrscheinlichkeiten müssen bei jedem Iterationsschritt noch normalisiert werden, um sie im Bereich $[0,1]$ zu halten:

$$p^{(r)}(i) = \frac{p^{(r-1)}(i)(1 + \Delta p(i))}{p^{(r-1)}(i)(1 + \Delta p(i)) + (1 - p^{(r-1)}(i))(1 - \Delta p(i))}$$

Relaxation Labelling löst das Segmentierungsproblem besser als punktorientierte Verfahren, wie das Beispiel rechts zeigt. Ein mit Rauschen versehenes Kreisbild soll segmentiert werden, in der Mitte das Ergebnis des optimalen Schwellwerts, rechts das Ergebnis von Relaxation Labelling nach 20 Iterationen.



High noise circle image



angepasst. Der Nachteil hierbei ist jedoch der erhöhte Rechenaufwand, da für jedes neu hinzugefügte Pixel ein neuer Region-Mittelwert berechnet werden muss.

Es wird nun zwischen einem **parallelen Wachstum** und einem **sequentiellen Wachstum** unterschieden. Beim parallelen Wachstum wachsen mehrere Regionen analog. Stoßen zwei untereinander annähernd homogene Regionen aufeinander, müssen diese miteinander verschmolzen werden. Beim sequentiellen Wachstum wächst immer nur eine Region, bis alle Nachbarn nicht mehr das Homogenitätskriterium erfüllen. Mit Hilfe einer bestimmten Methode zur Bestimmung eines neuen Saatpunktes wächst eine neue Region. Nach jedem Wachstum muss überprüft werden, ob evtl. angrenzende Regionen mit der zuletzt gewachsenen homogen sind. Ist dies der Fall, wird die Region mit der zuletzt gefundenen verschmolzen. Die neu entstandene Region wird wieder mit angrenzenden überprüft usw.

Split & Merge

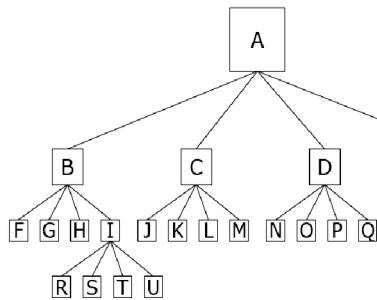
Das Split & Merge Verfahren versucht anhand einer einfachen Aufteilung des Eingabebildes eine vollständige Segmentierung zu erreichen. Als Eingabeparameter werden hier das Bild, ein Homogenitätskriterium (bspw.

Farbgleichheit oder Farbähnlichkeit) und eine Datenstruktur wie ein *Quadtree* erwartet. Ein Quadtree ist eine spezielle Baum-Struktur, in der jeder innere Knoten genau 4 Kinder hat. Diese Struktur wird hauptsächlich zur Organisation zweidimensionaler Daten eingesetzt. Die Wurzel des Baumes repräsentiert dabei eine quadratische Fläche. Diese wird rekursiv in je vier gleich große Quadranten zerlegt bis die gewünschte Auflösung erreicht ist und die Rekursion in einem Blatt endet. Durch rekursive Anwendung dieser Zerteilung kann die vom Wurzelknoten repräsentierte Fläche beliebig fein aufgelöst werden.

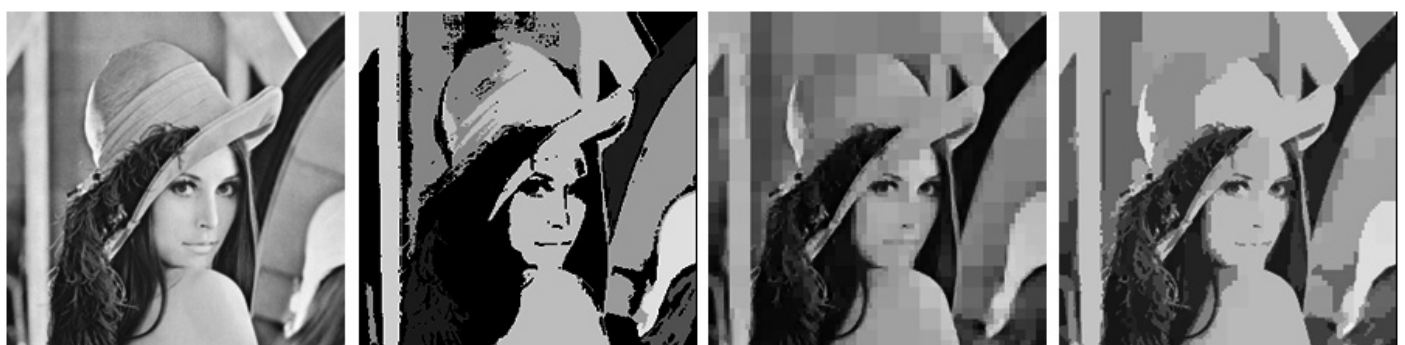
Split & Merge selbst gliedert sich in zwei Phasen: In der 1. Phase (*Split*) wird das Bild als Wurzel des Quadrees

angenommen. Erfüllt das Eingabebild nicht das Homogenitätskriterium, erfolgt eine Aufteilung des Bildes in 4 Regionen. Im weiteren Verlauf wird jede Region wieder auf das Homogenitätskriterium geprüft. Ist diese Bedingung nicht erfüllt, erfolgt eine weitere Teilung der Region.

Ist die 1. Phase abgeschlossen, folgt Phase 2 (*Merge*). Alle Blätter des Baumes repräsentieren eine Region. In der Merge-Phase kann also jede Baumstufe bzw. Zeile auf das Homogenitätskriterium hin geprüft werden. Ist dieses erfüllt, werden die betreffenden Regionen zusammengefügt. Die zusammengefassten Regionen werden ebenfalls in den Baum übertragen. Zum Abschluss werden alle Regionen, unabhängig von der jeweiligen Stufe des Baumes, ebenfalls auf das Homogenitätskriterium geprüft und zusammengefügt. Dabei wird jeder direkter Nachbar einer Region betrachtet. Die zusammengefügte Regionen bleiben zwar im Baum erhalten, jedoch erhält die Region, welche die anderen aufnimmt, einen Verweis, dass sie diese beinhaltet. Ebenfalls erhalten die hinzugefügten Regionen einen Vermerk, dass sie einer Nachbarregion hinzugefügt worden sind. Dadurch kann die so zusammengefügte Region mit den angrenzenden Nachbarn verglichen werden. Die folgende Abbildung zeigt das Ergebnis von Region Growing, Split sowie Split & Merge.



F	G	J	K
H	R	S	L
	T	U	M
N	O	E	
P	Q		



original

Region Growing

Split

Split & Merge